Conceptual Framework and Territorial Definitions for Improving Rural Statistics

July 2016
(revised November 2016)

Working Paper No. 10

Global Strategy Working Papers

Global Strategy Working Papers present intermediary research outputs (e.g. literature reviews, gap analyses etc.) that contribute to the development of Technical Reports.

Technical Reports may contain high-level technical content and consolidate intermediary research products. They are reviewed by the Scientific Advisory Committee (SAC) and by peers prior to publication.

As the review process of Technical Reports may take several months, Working Papers are intended to share research results that are in high demand and should be made available at an earlier date and stage. They are reviewed by the Global Office and may undergo additional peer review before or during dedicated expert meetings.

The opinions expressed and the arguments employed herein do not necessarily reflect the official views of Global Strategy, but represent the author's view at this intermediate stage. The publication of this document has been authorized by the Global Office. Comments are welcome and may be sent to global.strategy@gsars.org.

Conceptual Framework and Territorial Definitions for Improving Rural Statistics

Drafted By

Susan Offutt

Table of Contents

Acronyms and Abbreviations

CSI	Crop Suitability Index
EC	European Commission
ERS	Economic Research Services
ESPON	European Spatial Planning Observation Network
EU	European Union
FAO	Food and Agriculture Organization of United Nations
GAO	Government Accountability Office
GLC-SHARE	Global Land Cover Share
GLCN	Global Land Cover Network
GSARS	Global Strategy to improve Agricultural and Rural Statistics
HDI	Human Development Index
ILO	International Labour Organization
MDGs	Millenium Development Goals
MSCD	Minimum Set of Core Data
NASA	National Aeronautics and Space Administration
NBER	National Bureau of Economic Research
NSO	National Statistical Office
OECD	The Organisation for Economic Co-operation and Development
SDGs	Sustainable Development Goals
SEEA	System of Integrated Environmental-Economic Accounting
SNA	Systems of National Accounts
Sourcebook	Sourcebook of Indicators for Monitoring and Evaluation
UMZ	Urban Morphological Zones
UN	United Nations
UNDP	United Nations Development Program
UNSC	United Nations Statistical Commission
US	United States
WB	World Bank
WDR	World Development Report
WDR 2008	World Development Report 2008
WDR 2009	World Development Report 2009
Wye2	The Wye Group Handbook - Second Edition

Executive Summary

The Global Strategy to Improve Agricultural and Rural Statistics (GSARS) seeks to provide information to guide national and multilateral agency decisions about designing, funding, and evaluating rural development policy. This research topic aims to assemble guidance for developing countries that wish to improve rural statistics. One focus will identify a set of indicators for measuring progress of policies that seek to reduce poverty and promote environmental sustainability. Another will present a set of definitions of what is rural to organize the collection and interpretation of the indicators. This working paper takes a comprehensive view of rural statistics, considers how indicators might be selected, and suggests alternative definitions of rural.

The relevant rural statistics are those that describe economic, social, and environmental conditions within a designated territory. Agricultural statistics, on the other hand, are those that focus on the economic activity of farming, that is, the production of commodities. Rural statistics therefore have a broader scope and involve data collection from farm and non-farm households as well as on natural resource use and quality.

The framework for collecting rural statistics involves the identification of relevant indicators for the geographical areas considered as rural. Statistical indicators provide information that is meaningful in judging the impact of a project or a policy on rural households and communities and the rural environment. The data that are used to construct these indicators are collected and then organized by geography, where boundaries are established to separate urban from rural areas and to distinguish among rural areas with varying character (e.g., sparsely inhabited forests or more densely settled farming communities). To support improvement of rural statistics available to developing countries, this research aims to provide guidance on selecting relevant indicators and on defining and categorizing areas that are considered rural.

Many national and multilateral agencies have considered how to choose and construct indicators that relate to poverty reduction and environmental sustainability. This extensive body of work provides a foundation for selecting indicators relevant to rural development, keeping in mind that the capabilities of statistical agencies in developing countries vary widely. Moreover, countries may have different rural policy priorities, though they share a commitment to

the broad aims of the Sustainable Development Goals. Further work is required to select a limited set of indicators that address economic, social, and environmental conditions from among the large number that have been proposed. An important consideration is the cost-effectiveness of collection in a developing country context.

To be useful in decision-making that targets policies to areas in greatest need, the indicators should relate to the territory that is considered to be rural. However, there is no internationally-accepted definition of rurality. That determination is by its nature multi-dimensional and subjective, depending on the cultural, historical, and economic circumstances in each country. There may be consensus that rural places have smaller populations at lower density of settlement than urban places, but, as a practical matter, more specificity is required. The purpose of having a definition of rurality (which may thought of as a continuum) is to target policy but also to allow comparison of "like" areas to "like" areas within a country but also across countries.

Defining what territory is rural requires sorting parcels of geography into different categories based on criteria that reflect key dimensions of rurality. The parcels of geography can be of irregular shape (e.g., a census tract, a state) or of uniform dimension (e.g., a cell in a grid with uniform cell area). The categories can be discrete as an urban/rural dichotomy or be conceived as a continuum that allows comparison of the degree of rurality across parcels. The criteria provide the rationale for sorting the parcels into categories or arraying them along a continuum. The criteria most frequently employed concern sparse settlement (the distribution of people over the land), remoteness (access to markets and to public services), and land cover (the presence of agriculture or forestry). A continuous definition has the advantages of flexibility and of the ability to reflect heterogeneity within and across areas. The idea of the continuum can be embodied in either a continuous index or in a typology with a sufficient number of discrete classes. Geographic parcels of uniform shape, as in a grid, permit comparison of indicators across countries, ensuring that variation is due to that in the characteristic of interest and not in the size of the area over which it was measured. However, within countries, irregularly shaped parcels, such as result from administrative boundaries, may be preferred for domestic decision making.

Implementing a definition of rural depends on having the appropriate data on geography, population, and land use. For any country, data may be sourced from domestic collection (as from a census) or from access to international data sets (such as observations from satellites). As this availability will vary, it will

be necessary to assess the feasibility of implementation across countries with different levels of statistical capability. Moreover, it must be emphasized that the definition of rural is subjective, so countries may have different views of the relative importance of the sparseness of settlement, remoteness, and land cover in determining rural character. However, it should be possible to advance a single definition that can be used to make international comparisons of progress toward poverty reduction and environmental sustainability.

1

Introduction

The Global Strategy for Improving Agricultural and Rural Statistics (GSARS) identifies the need to foster agricultural development, consistent with the Sustainable Development Goals (SDGs). The demand for information to guide policy makers, the allocation of resources for aid and investment in agriculture and rural areas and to measure progress towards these goals motivates the collection of relevant data. The use of comparable statistical information is key for more solid evidence-based policy making.

The research seeks to advance the quality and availability of statistics about rural people, activities and the environment. Its domain is construed to include agriculture but also the well-being of rural households (including those not engaged in farming) and the sustainability of the use of natural resources. The Wye Group Handbook Second Edition (Wye2) notes the distinction between two sets of related statistics, those on rural areas and those on agriculture households. Rural statistics are said to be *territorially based*, that is, addressing conditions within a rural area. On the other hand, agriculture statistics are concerned with the production of farm commodities and are said to be *activity based*. The handbook points out that the two terms are "orthogonal in the sense that they refer to concepts in statistics that are independent of one another" (p. 26). This research topic is focused on statistics on territory considered to be rural and not on the activity of farming. Defining what areas are rural is a necessary antecedent to collecting data on resident households and natural resources.

The demand for rural statistics arises *ex ante* as needed support for decision making but also *ex post* as measures of effective public policy. Accountability in the use of public resources requires demonstration of effectiveness, and, as stated in Wye2, "quantification is the by-word of accountability" (p.13). Here, both national governments and multilateral development organizations are expected to be accountable for the success of the aid and investment provided

rural areas. With respect to the Sustainable Development Goals, the objectives of national and international institutions are aligned, so there is a shared interest in gathering information about economic, social, and environmental conditions in rural areas.

Monitoring and evaluation of individual development projects requires information internal to project management but also measurement of results. As described in *Tracking results in agriculture and rural development in less-than-ideal conditions: A sourcebook of indicators for monitoring and evaluation*, the ability to assess results depends on information collected from the beneficiaries and targets of the project. Such data are likely to become available only after the project has closed and are typically collected by agencies other than those sponsoring the project. This research topic is concerned with identification of indicators that are meaningful in judging the ultimate impact of a project or, more likely, a collection of projects and policies, on rural households and communities and the rural environment.

While national and international decision makers share goals, they do not necessarily share frameworks for collecting and analyzing data with which to make resource commitments and to judge progress toward mutually held goals. National policy officials address the problems of their own country, and historically this focus has made it so that domestically-collected statistics served as relevant descriptors and measuring sticks. More recently, especially given the trans-boundary implications of economic, social, and environmental developments in rural areas, nation states seek to compare their own circumstances with those of others similarly situated. At the same time, multilateral institutions are deciding among projects or larger programs spread over countries. Accountability to their funders requires a justification for the choice and success of projects in one country versus another. For these reasons, the comparability of rural statistics based on international standards for sources, methods, and procedures would be advantageous. As one of its fundamental principles of official statistics, the United Nations Statistical Commission (UNSC) states, "The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels" (Principle 9). As the Wye Group Handbook - Second Edition (Wye2) observed, such harmonization does not exist in the context of developing countries. Achieving this consistency in pursuit of international comparability presents a considerable challenge to collecting rural statistics across developing countries, whose circumstances vary widely and whose national statistical systems are of uneven quality.

2

Conceptual Framework

The GSARS provides the conceptual framework for advancing proposals aimed at the betterment of rural statistics in developing countries. There are two key aspects of the framework that give structure to the task of this research topic; one is the extension of the domain of interest beyond agricultural activity, and the other is the establishment of the scope and coverage of data to be collected. Even with these parameters set out, there are still significant challenges to implementation with respect to defining rural areas and identifying core data in the less-than-ideal settings found in many developing countries.

2.1. COMPREHENSIVE VIEW OF WHAT IS RURAL

The World Development Report (WDR, 2008) emphasized the centrality of agriculture in promoting growth and attainment of the Millennium Development Goals (MDGs). Likewise, the key role of agriculture is recognized here and a comprehensive perspective on rural communities is also embraced. This research topic is concerned with the territorial aspect of rural statistics that connects agriculture to rural households and to the environment, a view that transcends an exclusive focus on the farm enterprise and farm households. There are three related dimensions:

- Economic, encompassing agricultural inputs (land, labor, and capital) and outputs (commodities but also income), as well as upstream (e.g. veterinary) and downstream (e.g. marketing and packaging) services;
- Environmental, including uses of natural resources (land and water); and
- Social, identifying farm and non-farm households in roles beyond those of producers and consumers of agricultural output to users of social services (e.g., health, education) and to members of communities and also providers of social services.

The intersection of these dimensions is discussed in the GSARS seminal document (WB & FAO, 2010).

2.2. SCOPE AND COVERAGE OF DATA

To move forward with rural statistics in the broad dimensionality of the GSARS, the scope and coverage of data collection must be identified. The GSARS is clear that the starting point is the structure established by the international Systems of National Accounts (SNA) and the satellite account, the System of Integrated Environmental-Economic Accounting (SEEA). Appeal to these internationally-recognized systems provides direction on collection and classification of data on agricultural activities as well as land and water use. There is not so clear cut system for structuring household data collection to produce social statistics, but the socioeconomic variables included in the SNA would provide guidance on which data are most relevant. The groundwork for comparisons across countries is laid by reliance on existing international standards.

Within this broad scope, the GSARS also establishes the basic units for data collection in each dimension. Setting aside agriculture, for social statistics the basic unit is the household and for environmental statistics it is a geographic land parcel or appropriately defined environmental segment. In the case of farm households, there may be a one-to-one correspondence with land parcels and so data on all dimensions may be collected from the basic unit of the household. Georeferenced data would allow placement of the household's farm into a mosaic of land use. However, for non-farm families, the linkages to the natural resource base may not be so clear and will have to be established in other ways.

2.3. SELECTING INDICATORS

Assuming the boundaries of a rural area and of sub-areas within it can be delineated, the question becomes what data to collect about households and land parcels in those areas. The data are the raw material for building indicators to be used in monitoring and evaluation. The GSARS identifies core items as ones whose data enter into a multitude of indicators. Designation of core data would provide the direction to countries about a minimum set to be collected and used in national analysis but also in international comparisons. The GSARS is fairly specific about what data constitute the core set for agriculture as an activity, but it is less prescriptive about which relate to the non-agricultural dimensions of rural statistics. Making a determination requires deciding which data would provide information relevant to the measurement of progress toward the shared goals of poverty reduction and environmental sustainability.

Finding a globally-applicable definition of rural and identifying a core data set to support the building of indicators is a tall order. As the Sourcebook (Tracking…, 2008) points out, the circumstances in developing countries are less than ideal. An ideal environment would exhibit strong demand for information, use evidence to inform decision making and measure progress, and have access to timely information. These conditions are not commonly observed in developing countries, especially with respect to the availability of relevant information. The Sourcebook emphasizes the challenges of dealing with partial information in these settings, and the current research topic's prescriptions have to accommodate the reality of less-than-ideal conditions.

2.4. DEFINING WHAT IS RURAL

The GSARS thus provides the high-level parameters for efforts to improve rural statistics and achieve international comparability. However, a practical scheme for organizing statistical data has to be devised. The first step required to move forward is setting a definition for rural. Definition in this context means deciding on a scheme to classify areas as rural and to make distinctions among areas of varying rural character. While Organization for Economic Cooperation & Development (OECD) has set forth a definition for its mostly developed member countries, there is no corresponding international standard for developing countries. Most often, rural is defined as the residual after what is urban is identified (and there are many ways this might be done), but that simple dichotomy obscures important differences within a rural area.

Going farther in making distinctions is complicated by the multi-dimensionality of rural character and the subjectivity that accompanies the selection and weighting of relevant dimensions. What area is considered rural is a social representation of space, depending on the perspective (resident, policy maker, research analyst), rural character can be associated with many different dimensions and observable traits, including proximity to urban areas, size of towns or villages, or the presence of agriculture or forestry. Correspondingly, there is no unique, universal definition. The choice of dimensions to include and their relative importance in creating a classification is unavoidably subjective. Clarity about the rationale for the choice is thus imperative.

3

Indicators

Assessment of progress toward poverty alleviation and sustainability calls for the measurement of outcomes and results at many levels of aggregation: household; regional; national; international. In pursuit of international comparability, it would be national results that are of interest, and here is where the identification of a core set of indicators collected to the same standard is required. The general question of the conceptual basis for creating and judging the quality of indicators has been addressed by Wye2 and OECD and elsewhere. At issue here is selection of core indicators for assessing poverty alleviation and sustainability as reflected in the wellbeing of rural households and in environmental quality.

3.1. CORE ITEMS

The first pillar of the GSARS concerns the identification of a Minimum Set of Core Data (MSCD) and the determination of national priorities. Specifically, the aim is to "select a minimum subset [of indicators] that countries will provide using common definitions and methodologies to ensure that measurements are internationally comparable" (p. 13). Recommendations for the frequency of required data collection and reporting are also required. The Wye2 handbook authors point out that there is no generally accepted definition of an indicator. The GSARS provides, for its part, a hierarchy: indicator (e.g. food production index); data item (e.g. maize, an item that enters into the index; and variables (e.g. for maize, these include area harvested, yield, production, etc). The core item is defined as "one whose data enter into a multitude of indicators needed to monitor and evaluate development policies, food security, and progress toward the MDGs" (p. 13). Further, "a core item should be the first to be included in the statistical system and the last to be removed as a result of budget shortfalls" (p. 13).

Six categories of core items are delineated, as listed below;

- Crop
- Forestry production
- Agricultural inputs
- Socioeconomic
- Land cover
- Public expenditures on subsidies, infrastructure, and health and education

While providing a good measure of specificity about components of the core with respect to the first three items on the list, the question of what data to include under other items is less settled. The MSCD includes some non-agriculture data. Economic variables refer to rural infrastructure with core items being the area equipped for irrigation, roads, railways and communications. Social variables are related to demographics, including such core data items as employment status by sex, household income, and housing conditions. The determination of the level of geographic detail and the frequency with which data are to be collected is left open, to be determined for each data item. The MSCD require that demographics data should be reported by urban and rural population.

3.2. CHOOSING RURAL INDICATORS

This research topic is aimed at providing guidance on constructing rural development indicators and making choices about those for domestic purposes that will enter into international comparisons. The policy goals are often broad, and it is likely that many indicators will be related to one another and to more than one policy objectives. Here, Wye2 considers that while all indicators should meet generic standards of quality, for rural development indicators there are three basic aspects that "any reasonable assessment of conditions in rural areas and trends must take into account" (p. 149). These three include;

- Territory because rural development is a spatial concept and is concerned with differences and similarities in performance and outcomes across geography.
- Themes because rural development is a multi-sectoral concept concerned with economic, social, and environmental issues.
- Time because rural development is a dynamic concept.

Considered as a set, the indicators should not be strongly correlated (as they would all "tell the same story"), thereby avoiding collecting data unnecessarily and overwhelming users with extraneous information. As regards construction of indicators in developing countries, Wye2 observes that the main policy goal is poverty reduction. It references as possible inspiration for selection of indicators the World Bank's (WB) identification of the key factors that drive gains in rural well-being: improvement in the rural economy; movement toward a sustainable natural resource base; creation of an "enabling" environment for broad-based and sustainable rural growth; and enhancement of social well-being, management and mitigation of risk, and reduction in vulnerability (2008, p. 153). The SDGs also consider the complicated inter-relationships among factors that promote improvements in conditions for rural people and the environment.

Because Wye2 addressed circumstances in both developed and developing countries, its presentation included the approach taken by OECD in generating statistics for rural development policy. The challenges identified therein apply equally to both sets of countries:

- Causality between indicators and outcomes is difficult to establish due to the multi-sectoral nature of rural development policies and the number of variables and places that might, in combination, provide insight. Further, strong theory on which to base selection of indicators is lacking.
- The diverse perspectives and large number of the users of rural statistics mean the indicators must be credible, and achieving this trust on the part of the user community imposes demands. Indicators must provide unambiguous measures of strengths or weaknesses of an area that are available across all units or regions being observed and do so in a timely fashion and in a time series. They must be complete with regard to the issues being evaluated. And there should be no controversy about the data or their reliability.

These challenges loom larger in the less-than-ideal circumstances of developing countries, but they are worth keeping in mind as guideposts in developing a core set of indicators.

3.3. DEVELOPING COUNTRY CONTEXT

The Sourcebook considered how best to devise a plan for collecting data and creating indicators to monitor and evaluate agricultural and rural development projects and programs in developing countries. It also tackled the question of how to proceed in monitoring policy goals at the international level, where comparability requires standardization across countries. The Sourcebook stressed the need for commitment of all countries to maintaining the same indicators at the national level & highlighted that doing so should "not be too onerous a burden" because the same indicators are sought at both the national and international levels. These indicators were compiled in a list of nineteen, the majority of which relate to agricultural activity. Those that related to the territorial aspects of rural development include:

- prevalence of underweight children under five in rural areas and rural poor as a proportion of the total poor population;
- percentage of the rural population using financial services of formal banking institutions;
- withdrawal of water for agricultural as a percentage of total freshwater withdrawal, proportion of land area formally established as a protected area, and change in soil loss from watersheds; and,
- percentage of land area for which there is a legally recognized form of land tenure.

The Sourcebook notes that these indicators, on their own, are insufficient for monitoring policy progress but should be seen as an essential subset in each country's program of statistical data collection.

In the Rural Development Indicators Handbook (2000), the WB proposed five indicators to be used in constructing a "Rural Score Card", within which five outcomes and the corresponding indicators were detailed:

- The proportion of the rural population with incomes below the poverty level is represented by the indicator of the population below the poverty line expressed as a percentage of the total population.
- Improvement in social and physical well-being is represented by the indicator of the percentage of the rural population with access to safe water.
- Human development is represented by an indicator of the infant mortality rate, measured as deaths per 1000 live births.

- Increase in gender equality is presented in the indicator that is the ratio of male primary school enrollment (percent gross) to female primary school enrollment (percent gross).
- Food security enhancement is measured by the indicator showing the prevalence of malnutrition in children under 5.

In the compilation of country data and indicators actually available in 2000, many gaps existed. Some series were available only on an aggregated national basis, with supposed correlation with rural outcomes. In the context of the GSARS, the five indicators proposed by the WB would be included in the social dimension of rural statistics and indicators.

3.4. POLICY CONTEXT AND CRITERIA FOR SELECTION

This review demonstrates that there is no shortage of general guidance about creating indicators and that there are plenty of candidates for indicators. However, constructing a core set of rural indicators will require choosing which indicators should relate to poverty reduction and environmental sustainability. Moving from the relatively well-covered economic to the less-considered social and environmental dimensions outside agriculture is challenging. Moreover, matching the chosen indicators to definitions of rurality may present more methodological and practical challenge, as data collection and aggregation will have to be aligned with the boundaries of the designated rural areas. And, again, doing so in the less-than-ideal circumstances of developing countries will present significant constraints.

The next step is to winnow the multitude of indicators into a smaller set with relevance to rural development policy goals. A framework identifying the elements of rural development policy will provide structure for considering which indicators are likely to be of most value and use. Countries may have different emphases in goals at the domestic level and could tailor the selection of indicators accordingly. However, the research topic also aims to identify a minimum number of indicators for all countries to report on. In conjunction with a consistent definition of rural, valid international comparisons of progress toward policy goals could then be made. The ultimate output of the project is a set of guidelines that countries can use to collect raw data and construct indicators that are intended for use by domestic policy officials, civil society and by the international community.

The label of “rural development policy” may be applied to any number of government goals and interventions. To provide boundaries for the selection of

indicators in the GSARS context, the elements of rural development policy in this context must be identified.

For any country, the overarching development goal is balanced growth between urban and rural areas. In rural areas, improvements in agricultural productivity can lead to increased farm income but non-agricultural growth is required as a complement. So, the preferred policy approach seeks territorial rather than sectoral development, having the entire rural economy as its scope rather than only agriculture. In addition, countries may recognize distinctive cultural and geophysical aspects of rural territory, characteristics that again transcend a focus on farming activities.

Rural policy in developing countries would therefore generally include the following elements, as presented in the Wye2 handbook (p. 78).

- The role of a sustainable agriculture, which still represents the main source of livelihoods and food security for most of the rural people.

- The enhancement of the rural non-farm sector, in order to reduce farmers' vulnerability and improve their income and employment opportunities.

- The provision of basic social services, necessary condition for a broad-based rural development. Additionally, the creation of social safety nets may represent another important vehicle to tackle poverty.

- The lack of infrastructure and access to information, which are two of the most critical impediments for rural population development and economic growth.

- The promotion of easier access to credit, productive inputs and technologies, land, extension services, in particular to women and indigenous peoples.

In the time since the publication of the Wye2 Handbook, emphasis on the environment has continued to increase, reflecting a recognition of the need to address increasing resource scarcity and the underprovision of environmental services. Intensive agricultural systems may impair the environment through pollution of air and water, and extensive systems can result in deforestation, desertification, and loss of biodiversity. Considering these negative consequences, resource scarcity and degradation can adversely affect both the

sustainability of agriculture and public health. For rural farm and non-farm households, clean air and potable water are necessities beyond their application in agriculture.

In the context of GSARS focus on rural development, the domain of rural statistics excludes agricultural production (addressed elsewhere in the Strategy), but includes the more general circumstances of rural territory important to all households. Policy may also be concerned with the vitality of communities in which households reside and the effective social and economic linkages among villages, towns, and urban cities. Rural statistics and indicators would be required to inform decisions about policies in pursuit of the following objectives:

- sustainability of agriculture insofar as it affects natural resource qualities and quantities that are necessary to support public health and social sustainability;
- enhancement of non-farm employment opportunities;
- provision of basic social services such as education and health care and of social insurance;
- improvement in physical and telecommunications infrastructure; and,
- promotion of access to financial services outside agricultural credit.

So, although agriculture is a key component of rural development, here the domain of rural statistics and indicators will include non-agricultural aspects of household wellbeing and territorial development.

The focus of GSARS is on the household in the context of rural development. Indicators are sought that will enable assessment of the wellbeing of rural households, both agricultural and non-agricultural. Even with a delineation of rural policy that excludes agriculture, the scope for policy relevant indicators is broad. Here, attention will be given to selecting a relatively small set of key indicators that will be used in making international comparisons of wellbeing across countries once a definition of rural territory is consistently applied. These indicators will of course also have relevance to domestic policy goals, but would likely represent a subset, though an essential one, of all the indicators that might be appropriate to assess national rural policies.

These key indicators address the economic, social, and environmental dimensions of rural development from the perspective of household wellbeing. They are selected with reference to the broad aims of rural policy outlined earlier and with consideration of the basic needs perspective on poverty and also the Sen-Stiglitz Commission view of the components of quality of life.

- As described in the 1997 UN Human Development Report, the basic needs perspective takes poverty as the deprivation of material requirements for minimally acceptable fulfilment of human needs, including food. Insufficient income is one aspect, but also needs for basic health and education and employment. As such, it is more narrow than the capability perspective that underlies the human development indices.

- The Sen-Stiglitz-Fitoussi Commission (Commission on the Measurement of Economic Performance and Social Progress, 2009) identifies eight dimensions of the quality of life: health, education, personal activities, political voice, social connections, environmental conditions, and social and economic insecurity. For the GSARS context, the inclusion of environmental conditions alongside basic needs is significant.

It should also be noted that the approach taken to indicators here is less broad in scope that that of the rural livelihoods paradigm. A rural livelihood is defined as the capabilities, assets, and activities that rural people require for a means of living (FAO 2003). The range of associated indicators of improvement in rural livelihoods thus encompasses the different ways in which socio-economic and ecological systems and their governance contribute to determine income generation and distribution in rural areas (FAO 2016). So, while basic needs are included in the set, so are supportive aspects of development including, for example, the nature of land ownership and measures of agricultural productivity.

Accordingly, the list of key indicators given below is a subset, but an essential one, of all the indicators that would be useful in designing and assessing rural policy. With expansion of the capability of national statistical offices along the lines being pursued by GSARS, the addition of indicators to the key subset would be desirable. Additionally, attention should be paid to other requirements for indicators, especially in relation to the SDG, so as to allow use in several contexts. As it is, the list presented here is simply a place to start. The

emphasis of each indicator is indicated in a general way. The indicators are identified by associated dimension of rurality, although it will be seen that some will rightly belong in more than one.

- Economic: household monetary income from agricultural and non-agricultural sources vis-à-vis poverty thresholds
- Economic: employment in agricultural and non-agricultural jobs
- Economic: use of banking services
- Social/health: access to safe drinking water
- Social/health: access to adequate sanitation
- Social/health: access/ use of health services
- Social/health: food security
- Social/education: adult literacy
- Social/education: children in school
- Social/housing: overcrowding in dwelling
- Social/information: access/use of cell phones
- Environment: presence of degraded land
- Environment: dependence on single energy source

Possible additional indicators would include the following.

- Economic: access to road networks and public transport
- Social: subjective assessment of wellbeing
- Social: incidence of migration
- Environment: air quality

Whether or not these are included in the key set depends on the contribution of novel information (because, for example, access to road networks may be correlated with access to health or education services) and on the ease of measurement (because there may a lack of survey results or data from physical monitoring).

3.5. CONCLUSIONS AND NEXT STEPS

This set of key indicators is proposed in the context of the effort to develop a global definition of rural territory. As such, keeping the number to a minimum is intended to support assessment of the credibility and usefulness of the definition. What is sought is the ability to compare the status of household wellbeing vis-à-vis basic needs in a consistent way across countries.

In the next step, guidelines to support construction of these indicators will be written to identify raw data sources (censuses, household surveys, registers, other administrative data, remote sensing output), intermediate variables required for indicator computation, and specification of final indicators for reporting in aggregated form (i.e., as a percentage of all households in the unit of aggregation, say by country or sub-national administrative unit). Guidance on the use and interpretation of the indicators in a monitoring system will also be provided.

4

Definition of Rural

The purpose of having a definition of rural is to allow comparisons between rural and urban areas and among rural areas of different character and to be able to make these comparisons not only within a given country but also across countries. So, the aim is to be able to compare "like" areas to "like" areas, which is an ambitious goal given the heterogeneity in the nature of rural areas within and across countries. But capturing these distinctions is important because of the desire to target interventions for effective policy (see WDR 2008) and to understand how the context determines outcomes of interest, primarily poverty alleviation and environmental sustainability. The ambition is also challenged by the lack of a consensus definition of what is rural, as the concept of rurality is both multi-dimensional and subjective. There may be agreement at a high level, that places with small populations and low population density are more rural than places with larger, more densely situated populations. However, to make the concept useful and measurable, more specificity is required.

4.1. CONTEXTUAL SETTINGS

A fundamental issue is whether what is urban is defined first, making what is rural a residual, or whether a classification covers all of a country's territory. Wye2 counsels against use of the rural-as-residual approach because it limits flexibility and consistency in comparing and contrasting rural and non-rural areas. Nonetheless, the rural-as-residual approach is often employed even though it creates a sensitivity to the basis chosen for delineating what is urban. Cromartie and Buchholz (2008) demonstrate for a US city how the size of the population considered to be urban can vary with the demarcation of an urban area by its administrative boundaries (city limits), by its land use patterns (density of settlement), or by its functional economic relationships (labor market catchment area). This situation is not unique to the US or to developed countries but also applies to developing countries. The sensitivity of the

characterization of rural to the complementary concept of urban means that it is not advisable to make an arbitrary choice of definition simply for the sake of consistency. Experts with perspectives on both urban and rural settings must be involved in considering a definition that would be comparable across countries.

The key question, then, is how to choose a definition. As Wye2 and Cromartie (2012) point out, the answer depends on the purpose for which the definition is sought. The purpose in the current setting is to provide perspective that will help guide policy design and implementation and ultimately measure progress toward policy goals. That means that the definition of rural ought to be based on factors relevant in determining or explaining outcomes of interest, as opposed to referencing outcomes themselves. So, a definition of rural should not be based on stratification of areas by income if improving incomes is the policy goal. Instead, it should reflect the factors that condition income growth and that, ideally, are amenable to manipulation by policy interventions. So, for example, physical accessibility to labor markets would be appropriate to consider including in a definition if the goal is growth in wage employment for rural people. Accessibility could be improved by public road building. If ultimately the analytical purpose is to measure outcomes using indicators, then the definition of what is rural should not refer to outcomes.

What people consider to be a rural area is ultimately a subjective determination, conditioned by social and cultural context as much as by settlement patterns and landscape. Recognizing that subjectivity, no definition can be said or proved to be uniquely correct; rather, a definition can only be judged to be more or less appropriate for the use to which it will be put. In any context, rurality is likely to have several conceptual dimensions, and it is important to construct a definition that encompasses the aspects relevant for its purpose. A definition aimed at identifying what is rural for purposes of promoting recreation would presumably consider scenic beauty along with open space, while one that is focused on optimal land use would incorporate dimensions of resource quality and quantity. Here, the ultimate purpose is to facilitate assessment of the wellbeing of people and the quality of the environment in places in the developing world that are not densely settled. The definition should support understanding of the challenges of reducing the poverty that is widespread in rural areas and of promoting the sustainable use of natural resources in these areas. In this setting, the dispersion of the population over the land and the access the people have to basic requirements for living could be thought of as key dimensions of rurality. The characteristics and use of the available land and water resources are also relevant. The definition sought here is thus specific to

this context and to its use in policy analysis and decision making in pursuit of rural development within the framework of the SDGs.

One of the aims of this research is to inform the construction of a definition of rurality. Following from this aim, three aspects of a definition are particularly germane; first, it should provide a consistent basis for distinguishing what is rural from what is urban territory and also recognize different gradations of rurality across non-urban areas; second, given the focus on supporting effective rural policies in an international context, the definition should facilitate cross-country comparisons of progress toward policy goals and third, implementation of the definition must be cost-effective in the context of developing countries statistical capacity & capability.

4.2. EXISTING INTERNATIONAL AND NATIONAL DEFINITIONS

Over the years, definitions of what is rural have emerged in both domestic and international settings. While there is much to be learned from individual (mainly developed) country efforts, the international and multilateral definitions are most relevant to the task at hand because they have addressed the requirements peculiar to making cross-border comparisons. Differences in administrative structures and in the make-up of economies have to be acknowledged, along with the heterogeneity in social and environmental conditions that are present even within countries. Closer examination of these efforts sets the stage for constructing a definition.

The United Nations Statistical Commission (UNSC) has addressed the definition of rural. It observed that "because of national differences in the characteristics that distinguish urban from rural areas" no single definition of rural has emerged. The Commission suggests that the traditional urban/rural dichotomy be supplemented by constructing a continuum of localities based on density of settlement. Additional criteria might be added, for example, the portion of the workforce in farming or ease of access to medical care and education. In any event, the definition should not be too complicated in order that it be applied expeditiously and be understood by users.

The International Labor Organization (ILO) has recently undertaken a country-by-country inventory of urban and rural definitions to support its efforts promoting decent work in the rural economy. As the UNSC suggests is the case, ILO found that national definitions of rural and urban areas are highly

heterogeneous and frequently set rural territory as the residual of what is defined as urban. Majority of the countries based their definition on a single criterion: administrative area; population size; or settlement area. ILO also noted the use of multiple criteria, including the predominance of agricultural activities and the availability of infrastructure services. Regional differences were flagged, with administrative area being the main criterion in Latin America, Asia, and Europe, and population size in Africa. This heterogeneity suggests the considerable challenge in standardizing a definition for use in international comparisons.

In the WDR 2008, the WB emphasized the centrality of agriculture in promoting development and alleviating rural poverty. In making comparisons of agriculture growth and poverty across countries, the WDR uses the UN definition of rural as being that which is not urban. In UN population data, rural areas are as defined by National Statistical Offices (NSOs), where rural population is calculated as the difference between total and urban population. A listing of the basis for each national definition reveals a wide variety of object choice, types, and criteria. Consequently, this lack of consistency means that the rurality of the rural areas across countries is very different and is not based on any standardized methodology. Thus, the identification of a country's or the world's rural population will be sensitive to change in the underlying definitions of urban and the rural residual.

Acknowledging the heterogeneity across countries and within rural areas, the WDR devised a more differentiated scheme so as to enable comparison of "like with like." By plotting the contribution of agriculture to overall economic growth against the portion of the country's poor living in rural areas, the WDR placed each country into one of three "settings:" agriculture-based; transforming; and urbanized. Within each "setting," more and less favored areas for agricultural development were identified with respect to access to markets and to agro-climatic potential. These two criteria were used to sort rural areas into two types in each of three groupings of countries. The underlying objects were whatever unit a given country's national statistical system had used in making its initial delineation of urban versus non-urban. In this way, WDR 2008 devised a rural typology with the dichotomous urban/rural typology as its starting point. Here, the classification was used to support analysis of appropriate targeting of policy interventions to support development and poverty alleviation.

In the WDR 2009, Reshaping Economic Geography, a different definition of urban was used. It identified as urban all settlement above a certain minimum population size and minimum population density that are within a certain travel time by road. The residual areas considered non-urban or rural would thus be different from the 2008 WDR.

The OECD developed and implemented a definition of rural to apply to its mostly developed country membership. The objects to be sorted are administrative units specific to each country. The tiered sorting used two criteria, starting with classification by population density and then further by the size of urban centers within a region. The combined criteria produced three types of areas: predominately urban, intermediate, and predominantly rural. To recognize the heterogeneity of areas outside of the predominately urban, a further criterion of accessibility (driving time needed to reach a highly populated center) was applied to create more types. Eurostat employs a variant of this typology. Given its relative unimportance in these developed country rural areas, agriculture did not figure as a relevant dimension of rurality.

In pursuit of more effective territorial development policy, the European Union (EU) European Observation Network for Territorial Development and Cohesion (ESPON) devised an urban/rural typology and also presented one that differentiated among rural areas. The urban/rural split was based on the criteria of vicinity and remoteness, and it had five categories along a continuum (predominately urban, intermediate region close to a city, remote intermediate region, predominately rural region close to a city, and remote predominately rural region). Within the two categories of predominately rural, the areas were further divided into four classes. "Agrarian" areas were those characterized by primary sector (chiefly agriculture) dominance. Areas classed as "consumption countryside" were characterized by having visitors for recreational purposes. Secondary sector and private service sector areas were economically diversified and exhibited employment structures like those of urban areas. This structural typology makes delineations based on the characteristics of a region's economic activity.

4.3. CONSTRUCTING A DEFINITION

In defining what is rural, parcels of geography (e.g., provinces, counties, townships) have to be sorted into categories (e.g., rural, urban, suburban). In the absence of an international standard definition of what is urban or rural, different approaches have been deployed. The question is how to choose among

them. In the spring of 2015, the US National Academy of Sciences convened a workshop to consider how the US Department of Agriculture's Economic Research Service (ERS) should approach classifying rural areas. While the program was aimed at the US situation, one of the workshop presentations considered the question from a more general perspective (Waldorf and Kim 2015).

According to their exposition, the traditional design of a typology or classification for places or regions consists of three components: objects to be assigned to discrete types; types; and criteria used to assign objects to types.

- Objects would be spatial units, such as an administrative area or a census tract of irregular area or a cell in a grid with uniform cell area.

- Types would be the number of different categories to which an object could be assigned; urban and non-urban would be a two type or binary classification. As opposed to such a discrete typology or classification with a number of discrete, mutually exclusive types of objects, a continuous measure would allow comparison of the degree of rurality across spatial units.

- Criteria used to choose the underlying dimensions of rurality, such as density or size of settlement, would relate to the purpose of the definition. Criteria provide the rationale for sorting objects and for delineating among types. Variables are chosen to represent each dimension, dependent on data availability.

For each component, there are choices to be made in constructing a definition. In what follows, these choices are examined in turn by component. Then, alternative definitions can be formed by combinations of different choices for each component. From these possible alternatives, a subset is proposed for consideration.

4.3.1. CHOICE OF SPATIAL UNIT

In the context of territorial statistics, the spatial unit (or object in Waldorf and Kim's term) is the basis for arraying geographic areas along a continuum of rurality or placing them into discrete categories in a typology. Typically, political or administrative boundaries (such as provinces or census tracts) define the units to be placed in categories. Obviously, these units vary in size within and across countries, complicating interpretation of comparison of statistics.

One would like to ensure that areas with the same characteristics would end up in the same category. But when the units have markedly different geographic area, this is not likely. Units might end up in different categories or be assigned different degrees of rurality due to variation in the size of the unit and not in the dimension of interest.

Traditional approaches to data collection rely on political boundaries or specify administrative areas to construct spatial units. Political jurisdictions can be aggregated within a country, and administrative units, such as census tracts, can be related to political boundaries or other useful groupings, such as mail codes. Statistics derived from these data are aligned with delivery of public services and legislative edicts, and therefore can provide a good basis for analyzing policy interventions. For such reasons, irregularly shaped spatial units are the norm in national statistics.

An alternative to adopting spatial units set by political or administrative boundaries is to define units of equal area, as in grids with cells of square kilometers or miles or other resolutions. Then, comparisons across units are free of the distortions introduced by size differences. For a definition of rural, equal size spatial units provide a consistent basis for comparison across and within countries of variables representing key dimensions of rurality. Other advantages include stability over time and ease of integration with other spatial and georeferenced data (Eurostat). As Eurostat explains, "Statistical grid data are statistics geographically referenced to a system of (usually squared) grid cells in a grid net with Cartesian coordinates." For a population grid, these data can be derived from georeferenced point data sets based on building, business, and address registers in a "bottom-up" approach. When such data are not available, a "top-down" modelling approach combines data from the smallest available administrative unit with data on land use and cover to distribute population across cells within the administrative unit's boundaries. Information from other administrative records or spatial datasets could also be used to refine the modelling results. Other than for population, grid data exist that represent observations from satellite and aerial imagery about land cover and also meteorology, as for example in the global dataset collected by NASA's CERES (Clouds and the Earth's Radiant Energy Systems).

While the use of gridded data has clear advantages in principle, in practice the availability and accuracy of such data, especially for developing countries, can be problematic. Many, if not most, national statistical systems have yet to implement geocoding across all census and survey data collections. The

"bottom-up" approach to building a grid dataset is not therefore universally feasible. Instead, resort is made to "top down" approaches that use modelling techniques to populate grid cells. However, modelling results are necessarily a simplification of the real world and may not correspond to what could be observed on the ground. Whether such inaccuracy disqualifies a grid as the basis for defining a spatial unit depends on the use to which the data will be put, as will be discussed later. "Top down" population and physical data with global coverage are available as a result of multi-lateral initiatives, as for example, from the Center for International Earth Science Information Network at Columbia University's Earth Institute.

4.3.2. CHOICE OF TYPES

The delineation of what areas are rural can be made either using discrete categories of types or by eliding the types into a continuum. The two type classification, urban/rural, is a dichotomous typology. A typology can be constructed with any number of categories. For example, using access to urban areas as a criterion, a more remote place might be designated as being more than 100 miles from population center, a less remote place as 50 miles, and an adjacent place as within 50 miles. Multiple criteria could be used in a hierarchical way to construct more complicated typologies, as for example using density and size of settlement. Thresholds could be chosen to allocate objects among types (for example by establishing cutoffs for settlement population size). Similarities among objects could also be determined by non-parametric statistical techniques such as cluster analysis, most appropriate in the absence of a priori reasoning about the factors that should define groupings.

An alternative to the typology or classification with a number of discrete, mutually exclusive types of objects is a continuous measure that allows comparison of the degree of rurality. To introduce this measure, novel in this setting, consider that there are four steps in constructing a continuous aggregate index. Waldorf and Kim enumerate them as listed:

- identifying the dimensions of rurality;
- selecting variables;
- rescaling of variables;
- choosing a link function.

In the current context, the dimensions of rurality of interest are potentially the density and size of population settlement, the size of population, access to urban concentrations, and land use or character. Variables are chosen to represent

each dimension, and, as Waldorf and Kim note, are dependent on data availability. Rescaling is necessary because, of course, variables are expressed in different units. Waldorf and Kim propose a bounded scale, where zero would be the most urban and one would be the most rural. Variables are transformed by expressing each value relative to its maximum as a proportion of the difference between its maximum and minimum value. The scaling can reflect either a positive or negative relationship to rurality, where settlement density, for example, would be negatively related to rurality while distance to urban concentrations would be positively related. Finally, a "link function" expresses the way the dimensions "jointly determine the rurality of a place" (Waldorf and Kim). A geometric or an arithmetic mean could be chosen. See appendix for detail on these calculations.

Waldorf and Kim observe that a continuous aggregate index produces a one-dimensional measuring stick that allows comparison among and between spatial units. The United Nations Development Program (UNDP) uses a continuous aggregate index to measure countries' level of development with the Human Development Index (HDI). The HDI is a "summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and have a decent standard of living" (UNDP Human Development Report 2015). The variables used to represent each dimension are, respectively, life expectancy at birth, mean years of schooling, expected years of schooling, and Gross National Income per capita. Such a continuous aggregate index would be generally applicable to any choice of dimensions of rurality of interest. Waldorf and Kim also note that, because the index is continuous, it is sensitive to changes in the underlying variables, which would allow, in this context, the value of the index to vary for a particular place as it becomes more or less densely settled, for example. Given the dynamic nature of rural to urban migration in developing countries, this responsiveness could be an asset.

4.3.3. CHOICE OF DIMENSIONS

Deciding on the dimensions of rurality to be represented is fundamental in constructing a definition. As the Wye Group Handbook asserts, "rural is about people and territory" (p.112). Accordingly, places with small population and low population density are generally viewed as more rural than places with larger, more densely settled populations. So, one important dimension of rurality concerns the size and distribution of the population over the land. Another dimension of significance could be the relative isolation or remoteness

of the population from urban areas. Lack of access to employment and goods and services can be a determinant of key outcomes or indicators of interest, including income, health status, and educational attainment. Given the importance of agriculture in the rural areas, the use of the land in farming is also a key dimension of rurality. These dimensions associated with sparse settlement, remoteness, and land cover are the ones that are most frequently incorporated into definitions of rurality (see Improving Rural Statistics Technical Report 1 and Wye Group Handbook, p. 109).

Sparse settlement: What is rural is a relative concept, based on observation of variation in the number and distribution of people over the land. The most populous and densely settled parts of a country are its urban areas. Rural areas are relatively less densely populated and are seen as different in degree, if not in kind. Chomitz et al note that the "opportunities, constraints, and living conditions in these two locales are obviously different" (p.1). The motivation of the Global Strategy lies in understanding the circumstances peculiar to rural areas as a means of informing policies aimed at alleviating poverty and promoting sustainability. The ability to characterize rural areas appropriately is key to designing policy and monitoring progress toward goals of policy.

When it comes to characterizing rurality with respect to population size and distribution, the question of what is urban cannot be avoided. Typically, urban areas are defined relative to thresholds regarding settlement size and density so that territory that is not urban is considered to be rural. But, as many have argued, the urban/rural dichotomy is a false one, there is in reality a continuum, where the thresholds set the line between urban and rural is an arbitrary choice, made with considerations peculiar to each country. For this reason, the characterization of what is rural should always recognize the need to consider the country's entire territory (Wye Group Handbook, p. 111). In this way, the idea and reality of an urban/rural continuum is preserved, and inconsistency between the definitions of what is urban and what is rural is avoided.

The sparse settlement dimension of rurality can be represented by variables that measure or reflect population size and distribution. Population density is frequently used in characterizing the urban/rural continuum. As the Wye2 points out, it is "both intuitive for users and simple for providers of rural indicators to calculate," and "rural areas will always have a lower population density than urban areas" (p. 112). Moreover, using density as opposed to population size "neutralizes some of the distorting effects" of differences in the size of spatial units defined by political or administrative boundaries (p. 112).

Population density may be correlated with other variables that could also be used to represent sparse settlement. According to the Wye Group Handbook, "Population density reflects characteristics of settlement, distance, and even intensity of communication and land use" (p. 112). So, for example, a measure of built-up area could be mapped onto the urban/rural continuum as a proxy for a direct measure of population density. Other formulations of variation in land cover might also be used. The European Union has devised Urban Morphological Zones (UMZ) using land cover data that identify continuously built up areas that are relatively close to one another. The Food & Agriculture Organization of United Nations (FAO) Global Land Cover Network (GLCN) includes a category of "artificial surfaces" that might correspond to the UMZ concept of built up areas. The full spectrum of land cover categories in GLCN extends to cropland and grassland and other vegetative or geological cover that could be correlated with less dense populations. In many developing countries, the presence of agriculture is associated with more sparsely settled territory.

Population density is one variable that could be used to represent the sparse settlement dimension of rurality. By itself, however, it may not adequately capture the concept. A relatively densely populated area could be found in a town surrounded by a less densely settled rural region. That town could be classified as urban if no reference were made to its otherwise rural context. For this reason, urban/rural definitions may include a variable measuring the size of a population in a densely settled cluster. For example, the OECD considers an area otherwise deemed rural by virtue of low population density to be "intermediate" between urban and rural if it contains an urban center of more than 200,000 representing at least a quarter of the population. The threshold that moves an area out of the rural category is a choice that depends heavily on context. The OECD sets an "intermediate" threshold of 500,000 people for the generally more densely populated countries of Japan and Korea. In any event, both population density and population size may be included as variables representing the dimension of sparse settlement.

Remoteness: Remoteness as a dimension of rurality represents the opportunity, or rather lack of opportunity, people have to gain access to markets and to public services. Considering the relevant dimensions of rurality in the context of Latin America and the Caribbean, Chomitz et al. select population density and remoteness from large cities;

We argue that these criteria constitute important gradients along with economic behavior and appropriate development interventions might vary substantially.

Where population densities are low, markets of all kinds are thin, and unit costs of delivering most social services and many types of infrastructure are high. Where large urban areas are distant, farmgate (or factory-gate) prices of outputs will be low and prices of inputs will be high, and it will be difficult to recruit skilled personnel to public service or private enterprises. Thus remoteness and low population density together define a set of rural areas that face special challenges in development. (p.1)

Following this lead, the World Development Report 2009 (WDR) emphasized the importance to development of agglomeration economies, that is, the benefits that come when firms and people locate near one another together in cities and industrial clusters, effectively lowering the costs of transporting goods and sharing knowledge (NBER). Remote populations will find it more costly to take advantage of the benefits associated with urbanization. While there are surely negative aspects of urbanized areas, such as pollution and overcrowding, they also present unique opportunities to advance economic wellbeing.

Identification of a variable to represent the dimension of remoteness depends on a country's circumstances. Intuitively, a remote area is far from an urban area in terms of distance, the time it takes to travel physically from one place to another. The mode and speed of transportation, however, would be expected to vary depending on topology and on the presence or absence of infrastructure. Travel by road or train might be most usual in one place, but travel by water or foot in another. While the variables chosen would be different across countries, or even within countries, the underlying supposition is that physical presence in an urban area is key, however it is achieved. Here, it should be acknowledged that some degree of remoteness could be mitigated by telecommunication or internet services, as for example with the provision of health care services through satellite video. However, remoteness in terms of travel time is likely to be the most expedient approach in selecting a variable.

Remoteness is a concept defined relative to the proximity of an urban area offering the benefit of agglomeration economies. The question of how to define urban thus arises and again is seen to depend on a country's circumstances. The threshold for a city in Egypt, with Cairo one of the largest in the world, will be different from that in Mali, which has only Bamako with a population over one million. In the context of remoteness, these absolute differences in thresholds across countries matter less because the function of the variable is to represent relative difficulty of access to the closest urbanized areas. An alternative to country-specific information on access to urbanized areas is provided in a map

created for the WDR 2009 by the European Commission and the WB. It uses an urban/rural gradient to show travel time to large cities, set as those with 50,000 or more inhabitants. Whether or not this threshold is acceptable requires a judgment about relevance to each country's distribution of concentrated urban settlement and also regarding evidence about the minimum size of cities required for significant agglomeration economies.

Land cover: The variable that represents land cover is necessarily specific to geographic location. Moreover, presence of agriculture matters because of its importance & contributions to the rural livelihoods & ecosystem, requiring high-quality data on farms and farm households. While likely available in some countries by political or administrative unit, land cover grid data may also be found in datasets created and maintained by international institutions. The FAO Global Land Cover-SHARE (GLC-Share) initiative "…brings global land cover data under one roof for the first time and represents the most reliable global view of planetary land cover assembled to date" (www.glcn.org). The final classification of the GLC-SHARE database contains eleven land cover classes, as listed below, that accord with the System of Environmental-Economic Accounting (SEEA).

- Artificial surfaces
- Croplands
- Tree covered areas
- Shrub covered areas
- Herbaceous vegetation
- Mangroves
- Sparse vegetation
- Bare soils
- Snow and glaciers
- Inland water bodies

The GLC-SHARE grid contains cells of one square kilometer. High resolution national or regional data sets are available for some places (for example, east and southern Africa), but for others cell values are extracted from global databases.

The use of a land cover variable in a definition has to be tailored to the situation. Low-density settlement associated with rurality may be present outside of urban and cropped areas, and the nature of the associated land cover would need to be reflected in the definition. While every country has people

and cities and (presumably) cropland, not all have glaciers or mangroves. Another way to distinguish cropland from all other land cover might be to use an index of suitability for cultivation, as is available from national mapping or from a global database such as that compiled by FAO in its Crop Suitability Index (CSI).

4.4. EVALUATING OPTIONS FOR DEFINITIONS

Considering the number of ways in which a definition could be constructed, it is useful to have a framework that sorts them by key similarities and differences. Such a framework will be presented, then, having created set of generic options, examples of their real world application will help illuminate their value in the context of the statistical system.

4.4.1. FOUR OPTIONS

Different combinations of choices regarding spatial units, measures, and dimensions yield alternatives could be used to define what is rural. The choices to be made are identified below.

- Spatial units may be of equal size as in a grid or of irregular size as with political or administrative boundaries.
- A measure could be a typology with discrete categories or a continuous index.
- Dimensions could include one or more of sparse settlement, remoteness, and land cover.

So potentially there are two dozen combinations available as candidates for a definition. Setting aside the issue of which dimensions are chosen, there are four candidates based on choice of spatial unit and of discrete versus continuous measure.

S.No.	Spatial units	Measure	Dimensions
I	Irregular	Typology	Sparse Settlement, Remoteness, Land Cover or Use
II	Irregular	Continuous	Sparse Settlement, Remoteness, Land Cover or Use
III	Equal	Typology	Sparse Settlement, Remoteness, Land Cover or Use
IV	Equal	Continuous	Sparse Settlement, Remoteness, Land Cover or Use

4.4.2. EXAMPLES

Examples of each type can be found in research literature and in practice, with some variation in the dimensions included in each:

- *TYPE I:* The OECD Regional Typology classifies regions of member countries into predominately urban, intermediate, or predominately rural, based on population density and the size of urban centers within a region. The spatial or "local" units are administrative entities. A unit is considered rural if its population is below 150 inhabitants per square kilometer. These local units are aggregated into larger areas using a consistent scheme. Then a predominately urban area is one in which fewer than 15 percent of the population lives in rural local units, intermediate if between 15 and 50 percent, and predominately rural if more than 50 percent. The final step adjusts these classifications depending on the presence and relative size of urban centers within an aggregated area. An extension of this methodology adds the dimension of remoteness, represented by the driving time needed to reach a highly populated center, and results in five types of regions rather than three.

- *TYPE II:* Waldorf and Kim construct for the United States an Index of Relative Rurality with variables incorporating all three dimensions of rurality as represented by population size and density, remoteness, and built-up area. The continuous aggregate index employed has a value of zero for an extreme urban unit and one for extreme rural. The spatial units of irregular size are counties, though they note that "the index approach to capture degrees of rurality can also be applied to groups of counties… as well as to smaller scales such as townships or census tracts" (p. 11).

- *TYPE III:* The European Commission has produced a harmonized definition of cities and rural areas that presents a "new degree of urbanization" (Dijkstra and Poelman, 2014). It creates a three category typology using the settlement dimension: densely populated area (cities); intermediate density area (towns and suburbs); and thinly populated area (rural area). The spatial units of one square kilometer each are contained in a population grid based on population registers or other detailed sources of where people live. For countries that have not constructed such a grid, the EC disaggregated larger population data sets according to land use or land cover information.

- *TYPE IV:* The WB, WDR 2009 presented a continuous classification that arrayed developing countries according to settlement concentrations. This "agglomeration index" used a global population grid to define equal size spatial units. The index itself was constructed as the proportion of a country's population living in areas that met thresholds for population density and size and for travel time to urban areas. Though the Report focused on identifying urban areas and populations, the logic of the classification necessarily classified rural areas as non-urban based on the same criteria.

4.4.3. EVALUATION

The suitability of the candidate definition types I, II, III, and IV can be assessed based on the intended use discussed earlier. As enumerated earlier, the relevant considerations include the distinction between urban and rural and among degrees of rurality, the validity of making cross-country comparisons using the definition, and the cost-effectiveness of implementation in the context of developing country statistical capacity.

The consistency of the definition of rural with the definition of urban should be sought. Often analyses will define the urban boundaries first; the area outside is by default considered rural. Rural is then a residual, and also an undifferentiated one, whereas rural character is often considered to vary from one place to another. However, rurality may be seen as a gradient rather than a discrete condition. Chomitz et al. propose that population density and remoteness from large cities "constitute important gradients along which economic behavior and appropriate development interventions might vary substantially" (p. 1). This view is consistent with that of Cromartie and Bucholz, who assert the need to recognize that "any simple dichotomy hides a complex rural-urban continuum, with very gentle gradations from one level to the next." Definitions I and II with spatial units of irregular size rely on political or administrative boundaries that are unique to each country; they cannot deliver consistency across countries in a delineation of what is urban versus what is rural. Definitions II and III employ discrete categories for sorting spatial units; conceptually, enough categories could be defined to approximate a gradient of rurality but the number might be unwieldy. It would be more straightforward to use the continuous measures in II and IV if fine distinctions among rural areas were sought. Because definitions III and IV provide consistent granularity in segmenting territory across a country, they could accommodate a continuum that would include the most urban places on one end and the most rural on the

other. To provide a consistent rural/urban differentiation, the entire territory of a country should be characterized by the dimensions in the rural definition. That is, every spatial unit should be included in constructing the empirical categories or index values.

Comparability across countries is achieved only with definitions that use spatial units of equal size as in III and IV. In order to ensure that "like" areas are compared to "like," the distorting effects of variation in size of political and administrative units within and between countries have to be removed. With equally sized spatial units, either a typology or a continuous measure could be employed.

Adoption of a Type III or IV definition would be novel, so it is worth reviewing the essential steps in construction.

- Assume the dimensions of rurality of interest are the density of population settlement, the size of population, access to urban concentrations, and land use or character, represented, respectively by population size and density, travel time to an urban concentration, and one or more of land under cultivation, a suitability index for agriculture, extent of forest cover or desertification or other resource characteristic. Rescaling is necessary because variables are expressed in different units. A bounded scale, like that used in the HDI, could be chosen where zero would be the most urban and one would be the most rural. Variables are transformed by expressing each value relative to its maximum as a proportion of the difference between and minimum value.

- Data that represent each dimension's variable must be available. Because the domain under consideration is that of territorial statistics georeferenced data are indispensable. Linkage of each place to its relevant variable values would be facilitated by application of GIS. Ultimately, indicator values would be linked to each place, as well. Here it should be possible to draw on global data sets that have been compiled and georeferenced. The foundation for the data set to be used with the index is the population grid. Working from individual country population census data, analysts have used different techniques to distribute population into standardized grids. Each grid is assigned a population count, from which density can be computed with knowledge of the assigned area. These gridded

population maps are available from several sources and cover the globe. One such dataset was used in the construction of the WDR 2009 agglomeration index. The WDR 2009 used a complementary dataset to calculate travel distance by road from the center of each grid to the nearest urban concentration. To represent characteristics of land use, there are a number of options, including FAO's land cover data bases.

- A "link function" expresses the way the dimensions (represented by the variables) "jointly determine the rurality of a place" (Waldorf and Kim). The HDI "links" its dimensions using a geometric mean; Waldorf and Kim propose an unweighted arithmetic mean. The choice depends largely on the independence of the factors in determining rurality, see the discussion in the Appendix.

- With an index value for each grid in a country, the distribution of index values for the country can be graphed. Each country's distribution will be unique, but it would be expected that there would similarities in shapes of the distribution. For example, a highly urbanized country would have relatively large numbers of grids with values close to zero, while a very rural country's distribution would be skewed toward one. And, of course, a global distribution could be conceived, as well. With these distributions, the characterization of what is urban and rural can be done in a way that suits its user. A country such as Mali with small cities might consider that urban concentrations occur at lower densities and sizes than would a country like Egypt with very large metropolitan centers. For a multilateral organization like FAO, distinctions among degrees of rurality could be made simply by quartering the distribution, as is done with the HDI. Alternatively, policy relevant thresholds could be defined, as for example might be done to focus on the upper ten percent of the grid distribution representing the most rural and isolated places in a country or globally. Thus, the index distribution for a country would be identical for each user, but the way in which users characterized or distinguished among degrees of rurality could vary, as is appropriate given the subjective nature of the definition.

More detail on the calculation of the index can be found in the Appendix.

As for feasibility of implementation within a developing country context, the availability of gridded versus traditional boundary data is the key consideration. Certainly, domestic data collection would more likely support the use of traditional spatial units of irregular size as in definitions I and II. Gridded country data are available via the internet in global data sets compiled centrally by research and multilateral institutions. To construct a definition using a grid, however, a country would have to have the capability to manipulate these geo-referenced data sets. It is unknown how universal that capability might be. However, institutions outside developing countries with the requisite expertise and computing power could be enlisted to empiricize definitions III or IV. Statisticians in individual countries would not therefore be required to collect or transform data beyond their current practice; instead, they would likely require training in how to manipulate the typology groupings or index values to support delineations between urban and rural areas and among rural areas appropriate to their own circumstances. Cost effectiveness depends on the extent to which the existing capacity of a country would be taxed or require additional resources in constructing a definition. Both domestically collected and global databases are considered accessible. Capacity to manipulate data is also a consideration.

A further consideration is the specific use to which a definition might be put. One type of application would be stratification of samples in survey design, which would require creating grouping of potential survey respondents, which would be facilitated by a discrete typology. Another use would be policy design and analysis. In presenting and interpreting statistics and indicators for policy decision makers, a discrete typology might be preferred because it is simpler to explain than a continuous index. The opacity of the index and the complexity of its weighting scheme and variable transformations will not likely be accessible to most lay people, including policy officials. However, the continuous index could be very useful in empirical policy analysis, where statistical techniques such as econometrics are used to try to isolate the effect of rurality on progress toward policy goals. There are advantages in model building to adopting a continuous index of rurality as opposed to dummy variables associated with typology classes.

Returning to the question of appropriate dimensions to be included in a definition, here the subjectivity of the selection should be recalled. Either type of spatial unit or measure can accommodate any number of dimensions, as embodied in the variables chosen to represent each one. That is, there is no computational constraint on the number of dimensions to be included; more important are judgments about which dimensions are relevant. Avoiding

inclusion of an outcome of interest in a definition's dimensions is advisable, but, beyond that, there are no strictures. Dimensions that appear relevant for rural definition were listed above as sparse settlement, remoteness, and land cover or use. One or more dimensions could be chosen depending on the availability of variable data and relevance. Empirical implementation of alternative combinations across countries would help reveal the sensitivity of the typology categories and/or index values to the choice of dimensions and associated variables.

The choice of a definition for identifying areas as rural and to what degree involves tradeoffs among considerations of consistency in concept, cross country comparability, and feasibility of implementation. A fundamental choice concerns which dimensions of rurality to include in a definition and is conditioned on whether there are data available to appropriately represent each. Use of georeferenced data would facilitate the association of indicators with households located in each spatial unit, but whether that is feasible is a question to be answered. Location-specific labels for data is a potent tool, but it is not known how widespread the practice is among developing countries. While it might be expedient and ensure comparability of a definition across countries to have a central organization conduct the empirical implementation, those advantages might be mitigated by a desire to involve country statistical agencies more directly.

4.4.4. SELECTING DEFINITIONS

The conceptual framework for construction of a definition, augmented by considerations regarding the availability of data, provides the basis for proposing a set of definition of rural areas. With the foregoing identification and evaluation of options in mind, either of Type III or IV definitions using gridded data would be preferred to Type I or II using administrative units mainly because of the desire to make consistent comparisons across countries. Between Type III (discrete typology) and Type IV (continuous index), Type III would be preferred for statistical purposes when computing and comparing indicators, while Type IV would be most useful in empirical policy analysis. It should be noted that the same underlying gridded data are required for both types, so that both a discrete typology and a continuous index could be constructed from the same set. It is also the case that typology and/or index values for each grid cell can be aggregated when administrative or political unit boundaries are overlaid on a map grid.

As for the dimensions to be represented, the availability of data to construct appropriate variables is key. For any country, there are two potential sources of data. Domestic statistical or administrative collections (such as a census or household survey) are one source, and datasets compiled by external collection and synthesis (such as those available from the Earth Institute) are another. Which source yields the most appropriate data? When there are no country-generated data, then external sources may provide the only recourse. Such might be the case for comprehensive land cover data, for example. Or, domestic data may exist, but not in the required format. Population density might have to be modified in the absence of suitably georeferenced observations in spatial units. This modelling could be done in-country or results could be taken from population grids prepared by international institutions. In any event, judgements have to be made as to the reliability of the data in its proposed use: to create a definition of rural with which to associate policy relevant indicators. Reliability means that the data are sufficiently complete and accurate to be convincing for its purpose and context (US GAO 2009, p. 5). Modelled population density might not be accurate enough to plan postal delivery, but be sufficient for definition an area to be considered rural. Scale therefore is an important consideration.

The dimensions of rurality to be initially included in a typology (or an index) are sparse settlement and land cover. Remoteness is a third dimension that has been discussed, where it is meant to represent distance to an urban center that may restrict access to demanded goods and services. However, remoteness is not proposed here as an element for two reasons.

- The measurement of remoteness can be problematic. Road network data may not be available in all countries to the desired level of accuracy, and it is not clear how travel time is necessarily related to the existence of roads or even public transport. At the same time, the concept of remoteness is changing, where delivery of services (as with education and health) may be done to a certain extent through the internet, obviating the need for at least some physical travel.

- The purpose of the rural definition is to facilitate comparisons of indicators across "like" areas, and a number of proposed indicators measure access to important services, which can be a function of remoteness (e.g., lack of access to health care may be that it is too distant from the household's living place). Including remoteness in the

definition of rural would incorporate a policy target, and that, as discussed, is to be avoided.

Consider the construction of a discrete typology that combines a global population grid with land cover data, representing two dimensions of rurality.

- The starting point would be deciding on a definition of urban. As has been discussed, the decision should be made in consultation with experts who are working on urban definitions. At present, a collaboration among the WB, OECD, and the EU has considered global application of an EU definition of urban, periurban, and rural, with boundaries established by thresholds with respect to settlement population density and size. The "rural" areas of interest could be both the periurban and rural, subdivided further by density and size thresholds. (The intermediate periurban area is worth including because will contain both "near urban" and "near rural" places, which may challenge the typology).

- Once the density and size subdivisions are established, each class can be further split into land cover types, from artificial surfaces typical of urban areas to glaciers found in rural regions.

- Application of GIS techniques would assign each grid cell to a particular class in the typology (settlement density and size by land cover).

Administrative boundaries can then be imposed on this global grid at the national and/or subnational level. In this way, countries can see how the typology identifies rural areas and assess how this categorization aligns with domestic definitions and concepts.

When international comparisons are required, indicators can be constructed using grid cell data from the same typology class in individual countries. In a future when all economic, social, and environmental data are georeferenced, indicator values can be constructed by grid cell at any level of aggregation. In today's setting, however, the data that are to be used in the construction of indicators may not be georeferenced and may only be identified as coming from a particular administrative unit. In this case, administrative units will have to be characterized by the predominant typology class exhibited (i.e., the typology class with the largest number of grid cells within the administrative unit would give the unit the identity of that typology class).

As with the construction of indicators, guidelines will be prepared that provide the necessary detail on the datasets available for use and the procedures for specifying the typology with respect to population parameters and land cover. Provision will be made for sensitivity analysis to understand the stability of the categorization when thresholds are varied. As well, the mapping of administrative boundaries onto the grid cells will be discussed.

4.5. CONCLUSIONS AND NEXT STEPS

As has been emphasized, the definition of what is rural is unavoidably a subjective determination. For historical, economic, and cultural reasons, countries may differ in their views. There is no correct or incorrect definition, the important thing is that the definition is fit for purpose. A single definition will be put forward as a basis for a consistent comparison of policy relevant indicators across countries. An important aspect of the work going forward will be trying alternative approaches and considering their usefulness and feasibility in both domestic and international settings. The adoption of an international definition based on a grid cell typology or index will not in any way preclude individual countries from using existing definitions of rural and may encourage creation of a new one using the international template.

As with the construction of indicators, the next step is to prepare guidelines that provide the necessary detail on the datasets available for use and the procedures for specifying the typology with respect to population parameters and land cover. Provision will be made for sensitivity analysis to understand the stability of the categorization when thresholds are varied. As well, the mapping of administrative boundaries onto the grid cells will be discussed.

References

Brezzi, M., L. Dijkstra and V. Ruiz. 2011, "OECD Extended Regional Typology: The Economic Performance of Remote Rural Regions," *OECD Regional Development Working Papers*, 2011/06, OECD Publishing, Paris.

Chomitz, K.M., P. Buys, and T.S. Thomas, 2005. "Quantifying the Rural-Urban Gradient in Latin America and the Caribbean," World Bank Policy Research Working Paper 3634, The World Bank, Washington DC.

Cromartie, J., 2012. "Rural Definitions," www.ers.usda.gov.

Cromartie, J. and S. Bucholz, 2008. "Defining the 'Rural' in Rural America," Amber Waves, USDA Economic Research Service, Washington, DC.

Dijkstra, L. and H. Poelman, 2014. A harmonized definition of cities and rural areas: the new degree of urbanization. European Commission Regional Working Paper 01.

WB and FAO, 2010. Global Strategy to Improve Agricultural and Rural Statistics. Food and Agriculture Organization, Rome.

FAO, WB and UNSC, 2012. Action Plan of the Global Strategy to Improve Agricultural and Rural Statistics: for food security, sustainable agriculture and rural development. Food and Agriculture Organization, Rome.

FAO, 2014. Global Land Cover Network, Global Land Cover – SHARE. www.glnc.org, accessed 18/05/16.

European Union, 2014. ESPON Atlas – Mapping European Territorial Structures and Dynamics, European Commission, Brussels.

Eurostat. "Populations Grids." Statistics Explained. http://ec.europa.eu/eurostat/statistics-explained/ accessed 16/05/16.

International Labor Organization, Department of Statistics, 2016. Definitions of rural/urban areas for use in labour statistics: Analysis of the most common criteria used, Powerpoint presentation.

National Academies, 2015. Rationalizing Rural Area Classifications for the Economic Research Service: A Workshop Summary, National Academies Press, Washington, DC.

National Bureau of Economic Research, 2010. Agglomeration Economics, University of Chicago Press, Chicago.

OECD, 2011. OECD Regional Typology, Directorate for Governance and Territorial Development, OECD Publishing, Paris.

Robles, P.S., H. Kashef, and M. Castillo, 2016. "Definitions of rural/urban areas for use in labour statistics: Analysis of the most common criteria used," International Labour Organization, Department of Statistics, Geneva.

Tracking results in agriculture and rural development in less-than-ideal conditions: a sourcebook of indicators for monitoring and evaluation.
GDPRD, FAO and World Bank, 2008.
http://www.fao.org/fileadmin/templates/ess/documents/Sourcebook-Web-Version.pdf

UN, 1994. Fundamental Principles of Official Statistics, UN Statistics Division, New York.

UN, 1997. Human Development Report. UN Development Program, Geneva.

UN, 2014. *World Urbanization Prospects*, Population Division, New York.

UN, 2016. Principles and Recommendations for Population and Housing Censuses. UN Statistical Commission, New York.

Waldorf, B. and A. Kim, 2015. Defining and Measuring Rurality in the US: From Typologies to Continuous Indices, Commissioned Paper for National Academies Workshop on Rationalizing Rural Area Classification, National Academies, Washington, DC.

World Bank, 2000. Rural Development Indicators Handbook, Office of the Publisher, The World Bank, Washington, DC.

World Bank, 2009. World Development Report, Reshaping Economic Geography, Office of the Publisher, The World Bank, Washington, DC.

World Bank, 2008a. World Development Report, Agriculture for Development, Office of the Publisher, The World Bank, Washington, DC.

World Bank, 2008b. *Tracking Results in Agriculture and Rural Development in Less-than-ideal Conditions: A Sourcebook of Indicators for Monitoring and Evaluation*, the Global Donor Platform for Rural Development, FAO and World Bank, Sector Manager, Agricultural and Rural Development, The World Bank, Washington, DC.

Appendix

To translate the concept of the urban/rural continuum into an empirical measure, variables must be chosen to represent the desired dimensions of rurality, the variables must be transformed to a unit-free scale, and the resulting values must be combined to a single index number. Understanding the arithmetic that these steps entail helps guide the choices to be made.

In addition to a judgment regarding aptness in representing the underlying dimension of rurality, variable selection has to take into the distributional properties of the chosen data. This consideration arises because the values of the variable may not be evenly distributed along the axis between its minimum and maximum values. For example, Waldorf and Kim note an abundance of places with small populations and densities but few places with large populations and high densities and accordingly adopt the logarithms of population size and density as the forms the variables take. This transformation corrects for the skewness in the distributions that would otherwise weight the variables' scaled values toward the concentration in one part of the distribution. It is also likely that the data for land cover will exhibit skewness, or at least be non-normal, so a logarithmic transformation may be indicated in that case, too.

With a given set of values for each variable, the data are then transformed to a unit free scale. This transformation yields scale compatibility across the chosen dimensions of rurality as well as independence from the initial units of measurement (Waldorf and Kim, p. 10). As with the Human Development Index, Waldorf and Kim chose a bounded scale, where 0 represents the most urban object and 1 the most rural. They present the transformation, as below (p. 11).

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

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

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:

$$X_i \rightarrow 1 - (X_{max} - X_i) / (X_{max} - X_{min})) \; E[0,1]$$

Thus, with rescaling, the variable value for each object (here, a grid unit) will range from zero to one.

Once the scaled variable values for each grid unit are obtained, they are then combined to yield a single index value. According to Waldorf and Kim, this "link function" should "reflect how the four dimensions jointly determine the rurality of a place" (p. 11). What weight should each dimension have in determining the unit's index value and so its relative rurality, compared to all the other units' values? Waldorf and Kim choose the unweighted average (arithmetic mean) of the four rescaled variables as the link function, which they feel is appropriate in "the absence of any theoretical guidance." In contrast, the Human Development link function is now the geometric mean, having been changed from the arithmetic mean with the 2010 Human Development Report. The explanation is given below.

Poor performance in any dimension is directly reflected in the geometric mean. That is to say, a low achievement in one dimension is not anymore linearly compensated for by high achievement in another dimension. The geometric mean reduces the level of substitutability between dimensions and at the same time ensures that a 1 percent decline in index of, say, life expectancy has the same impact on the HDI as a 1 percent decline in education or income index. Thus, as a basis for comparisons of achievements, this method is also more respectful of the intrinsic differences across the dimensions than a simple average.

In the context of the GSARS, the choice between the arithmetic and the geometric mean seems to turn on the purpose of the index. The HDI is a continuum where low values of the index are indicative of relatively poorer wellbeing, which implies that better wellbeing ought to reflect good performance on all dimensions of the index. The concept of the continuum in the context of GSARS is not meant to impart any judgment about whether being an urban place is better than being a rural place. So, the idea that a high score on presence of agriculture should offset a low score on remoteness, for example, is incompatible with the notion that they are not substitutes for one another in determining rurality.