

*An Approach to Analysing and Disseminating Poverty Statistics at the level of Small Areas using Geographic Information Systems*

*By Edwin St. Catherine, Director of Statistics, St. Lucia*

## **1. Introduction**

Small areas as identified in the topic I am addressing refer generally to either statistical, planning, community based or politically defined geographic areas. These areas are generally smaller than the well know and defined administrative boundaries. Small areas are generally described as blocks/county's or wards, enumeration district boundaries, community or village boundaries and to a lesser extent electoral boundaries. In many countries of the Caribbean these small areas are defined in such a way that it is technically possible to aggregate the digital polygons from a GIS representation to the larger administrative subdivisions. In point of fact, generally, and I do not know of an instance where this does not hold true the aggregation of the ED polygons in a GIS can match perfectly up to the administrative region, suggesting the existence a priori of a spatial data model. Despite the preceding, when data is desired by a user on a small area, the user is almost exclusively interested in community boundaries. In some countries these can be very well defined as I believe is the case in Trinidad and Tobago, in others, example most of the OECS this is not the case. However, the community boundaries whether or not it is geographically well defined is the main means by which local planning institutions, community based organizations, NGOs, politicians and individuals develop plans and execute programmes at the sub-national level. Therefore, I will confine my use of the term "small areas" to communities whether notionally or geographically understood.

The Census is usually the only reliable source which the Statistical Office has of data on "small areas". Further, the question is, how can this information be massaged into a meaningful representation of welfare or living conditions to allow for the useful analysis of small areas. One well understood approach to this problem is the much more now discussed "Basic Needs" index. The idea is too ascribed to each household or characteristic of a person within that household a score based on the presence or absence of a condition which can be aggregated to the level of the household. This household based summary score can be normalized to a community score on the basis of the total number of households in the community. With each community allocated a score, all the communities within the country can be ranked and by extension a map of relative deprivation of communities derived when this community based indicator is combined with the "small area" GIS polygon to which I referred previously.

## 2. The Approach to Development of the “Basic Needs Index”

The following table describes the content of the index chosen in two studies, both completed this year by this presenter. The first of the two studies was done for the IDB and presented in a report entitled “*Trinidad and Tobago- Poverty Reduction and Social Development (TT-STR-COP)*”, using the CSSP (Continuous Sample Survey of Population 1995 to 2002) and the second was done using the Census 2001 database of St. Lucia.

The scores assigned to the variables were similar in both cases, in the case of the Census 2001 the variables identified in the attached table were associated with specific variables within the census database. Within the CSSP a combination of the person and housing modules for the questionnaire was combined to produce a living condition index classifying the CSSP dataset by quintile to examine the condition of the poorest quintile as the Trinidad and Tobago economy moved through its transition from 1995 to its current buoyant state and the impact of that transition on the “poorest” quintile. The Table which follows illustrates the variables used:

### Poverty Scoring System:

Item	Response	Score
1) Wall type	Brick/Block/Concrete	3
	Wood and Concrete	2
	Wood	1
	Wattle/Tapia/makeshift	0
2) Toilet type	WC to sewer/cess pit	1
	Pit latrine/None	0
3) Light Source	Electricity or gas	1
	Kerosene /none	0
4) Possessions***	TV/Telephone/Video/Stove/Fridge/ Washing machine	0.5 each
	Car/pick-up	1
5) No persons per bedroom	<1	3
	1-1.99	2
	2-3	1
	3.01 or more	0
6) Education of head (summary)	Tertiary/university	5
	Secondary complete	4
	Secondary incomplete	3
	Primary complete	2
	Primary incomplete	1
	None	0
7) No employed to total number of persons	1	3
	$x < 1, x > 0.49$	2
	$x < 0.5, x > 0.25$	1
	$x < 0.25$	0
Maximum Score		20

It can be argued that these indicators can be improved by incorporation of a larger and perhaps more granular set of computed variables. However, these capture the essence of what we are trying to demonstrate here. Each variable itself based on the category it assumes is ascribed a score, for example, each household possession gives the household a score of 0.5. whereas, the household with the value of a computed variable persons per bedroom which is 1.5 is assigned a score of 2. This implicitly states the relative position of possessions on aggregation when compared to the computed variable persons per bedroom. This is one of the main weaknesses of this approach; its intuitive scoring system can be subjective.

The list of variables built into the summary “basic needs” index at the household level consists of three types.

- 1) There are strictly household based or derived variables such as wall type, toilet type, light source and possessions which emanate from questions asked about housing conditions, this is the dominant set of variables used in this particular version of the index.
- 2) There is the education of head variable which ascribes a score to the household based on the level of education achieved by the head of the household. It is to be noted at this point that the education of head is not a household based variable but it is a variable derived from a person characteristic of a household member who happens to be the head of the household. The generation of this variable requires that the education of head variable be generated in the person dataset and via the relational link between the person and the household dataset it is transferred to the household dataset.
- 3) The remaining variables in the group 5) and 7) specifically, are a cross fertilization of the household and the person variables, that is, number of person per bedroom, which is an indicator of “overcrowding” and number of employed persons to the total number of persons which is the employment rate in the household, respectively, brings together both person based and household based characteristics to derive a score for the household.

### **3. Aggregation of the Index to the Community Level**

The seven scores are generated within the household database of the St. Lucia Census 2001 and the Trinidad and Tobago CSSP dataset. The sum of the scores represented the score for the household. To allow for the dissemination of this data in tabular or GIS form it is necessary to classify the household score. This is done by first ordering the records of the dataset on the basis of this living condition or “basic needs index” from the lowest score to the highest score and dividing it into five equal parts. Each quintile must have an equal number of records give or take one. This is the first most crucial step for the generation of variables suitable for display in a database.

In order to transfer the data into a GIS one has to determine what GIS shape files are available to allow the display of the poverty index described in the previous paragraphs. Typically, the administrative district boundaries are readily available and are pretty well established, however for local and community based planning activities this is highly inadequate for the presentation of the index via a map since the dynamics of the communities within the administrative boundaries are masked by the display of the data at the administrative level in a GIS.

At the level of the GIS, ED, ward/county or community boundaries are more suitable for this purpose and are defined and maintained at the polygon level, or at a more granular level of themes of buildings defined using a coordinate system to uniquely identify the building based on an index which can be readily linked to the Census dataset. The process we have described so far involves the generation of poverty scores at the household level. Since the GIS map requires the availability of the poverty index at a level that is physically identifiable, the household level computations described thus far is not suitable to allow the link to be made with physical features represented in the GIS. Therefore, a process of aggregation from the household poverty/basic needs index scores to the building level, in its most granular form at the very least is needed. A similar aggregation can also be made to transform the household dataset “basic needs index” scores to the ED or the community level. However, for the aggregation to be meaningful the household scores summed to the community level must be normalized by division by the total number of household for the community. Upon completion of this aggregation the derived score at the community level is the weighted average of the scores for all households within that community and in our specific case this score would be a number under 20 for all the communities identified in the GIS dataset.

At this point since the communities have a unique identifier in the GIS database and this same identifier in the poverty scores database, it is possible to display the poverty scores on the map for each community and to so color code the community map that the score attained influences the concentration of the specific color (“graduated”) on the map that identifies the poorer communities more intensely than the more “well-off” communities. In other words, a darker shade of brown will illustrate communities which are worse off than communities with a lighter shade of brown for example.

This score which allows us to classify the map at the ED or the community level also allows us to rank them and therefore give a critical indication of which are the least “well-off” communities or the “better-off” communities with respect to the basic needs index. This gives a very powerful indication to institutions involved in poverty reduction through geographically targeted interventions an evidenced based approach to the allocation of their resources to areas where they are most needed.