Beyond and Below the Nation State: Challenges for Population Data Collection and Analysis

*We need to look both beyond and below
the nation state in our efforts to understand population change.*

By Terence H. Hull*

While the science of demography addresses the whole of the human population, substantive demographic research is most often focused on populations with common characteristics. For the last six decades the nation state has been the social unit that has dominated demographic research.

* Professor, Australian National University, Canberra, Australia, e-mail: Terry.Hull@anu.edu.au.

The reasons for this focus make perfect sense. Nations define their populations in terms of citizenship and define the ways in which people will be identified in any effort to count the numbers. They have the authority, the interest and the resources to carry out collections of information about members of these defined populations. As members of the United Nations they collaborate with other nations to develop the methodological and technical tools used to analyse national population numbers in ways that are relevant to state policies and actions. In short, the nation is the foundation unit for understanding human population composition and growth.

Global population numbers are estimated by compiling the information collected by nations. Interest in populations of units smaller than the nation also relies on national statistical collections and national definitions of component populations, but for most users of data the focus is on the nation, and not the units beyond or below that political entity.

The most obvious expression of the importance of the nation state is the production of “data sheets” by research organizations and United Nations agencies. One of the longest running of these efforts is the annual publication of the Population Reference Bureau’s (PRB) *World Population Data Sheet*.¹ For many years the PRB Data Sheet has been placed on the bulletin boards of libraries around the world and has been cited as the source of information in student papers, faculty lectures and even government policy documents. From time to time the development banks and various development assistance programmes emulated the PRB data sheet with wall charts designed to highlight different regions or measures of interest to their constituents. The World Bank’s annual World Development Report² and UNDP’s Human Development Report³ contain substantial data annexes with information on nations and often provide critical analyses of the quality of the information provided by national statistical offices.

The proliferation of data sheets waned in the new century after 2000 as computerized data collections became more readily available on the Internet and therefore easier to update. For those in the Asian and Pacific Region, ESCAP’s annual *Population Data Sheet* has long provided an authoritative collection of national statistics on a wide range of topics.⁴ Like so many other data compilations the *ESCAP Population Data Sheet* is now available online, thus greatly reducing the cost and increasing the accessibility of the material, and making it more useful to the broad range of potential users across Asia and the Pacific.

The common characteristic of all these data sheets is the reliance on the nation as the key data unit. Compilers rely on national statistical offices to report information on population numbers and characteristics, either directly or through

the United Nations Population Division or the United Nations Statistical Division. Over the years these organizations have gained great expertise in assessing the validity and reliability of national data. Adjustments can be made to any information that is obviously incorrect. These data compilations thus represent the best estimates of national populations and regional and global population structures and trends.

The term region is often confusing when applied to population studies. Sometimes it refers to groups of nations with shared boundaries and geographic similarities, such as the regions of South-East Asia, the Pacific or Asia. These supranational units are embedded in the various data sheets. Confusion arises from time to time when different organizations use different definitions for regions, such as the different collections of nations used by the World Health Organization and the World Bank to define regions in their operations. The historical reasons for such differences are understandable, but of little consequence to the policymakers and students who attempt to apply information about regions in their arguments about social and economic development. For them there would be great advantage to the use of standard regional units and consistent inclusion of individual nations in those standard regions. For a variety of historical and political reasons this is unlikely to be achieved in the short term.

Occasionally people write about regions with similar social characteristics such as the Islamic region, or the francophone regions of Africa. These regions are defined by common national characteristics, but pay little regard to differences within nations. This is undoubtedly a useful approach to marshal information above the national level, and it has the benefit of simplicity. To get a regional estimate you add up data from each component nation. The component nations are included if they have people with the selected characteristic, such as more than half the population adhering to a particular religion, or 80 per cent speaking a particular language. That is simple enough, but it is not likely to satisfy the concerns of the minority interests that are not recognized in this categorization, or the problem of nations changing categories over time as the populations with the selected characteristics become more or less numerous. For the time being, the definition and monitoring of such regions is likely to depend on the units of nations, rather than other geographic or social units.

Nations, though, are extremely varied in size, from the population billionaires of China and India, to the island microstates scattered across the oceans of the world. Among the larger nations there is great social, cultural and economic diversity. The use of regions defined as groupings of nations means that the true heterogeneity of the globe is understated. For this reason it is important to consider

regions smaller than nations, with characteristics that can be seen to cross national boundaries to reveal the realities of human variation.

The European Union has attempted to define subnational units for social analysis.⁵ Called the *Nomenclature of Territorial Units for Statistics* (NUTS) they are intended as units of comparable population size for regional comparisons within and across nations. The regulation establishing the system sets a hierarchy of three levels of NUTS units with minimum and maximum population sizes for each level. Thus NUTS 1 is a unit with a minimum population of three million and a maximum of seven million; NUTS 2 has a minimum of 800,000 and NUTS 3 a minimum of 150,000. These may be either administrative or non-administrative units according to the political structures of each member state. The terms used to describe them follow local languages and terminologies. In Germany a NUTS 1 unit is called a *Länder*, while in the United Kingdom it is a *Government Office Region*. The analysis of those subnational units has the dual purpose of focusing policy on specific needs of relatively homogenous populations, and allowing identification of regions that share characteristics across as well as within nations.

One recent innovation in the analysis and presentation of data at the subnational level is the Indonesian website⁶ titled *Improving Data for Decentralized Planning*. This website was developed collaboratively by staff from the BPS-Statistics Indonesia (*Badan Pusat Statistik*), the Demography Programme at the Australian National University and the Demographic Institute of the University of Indonesia. It was partially funded by the Australian Agency for International Development (AusAID). The design aimed to accomplish three main things. First, the website should be accessible to people working at the district (*Kabupaten/Kota*) level, as well as in the national and international research communities. To achieve this the site has table labels in Bahasa Indonesia and English, and there are training modules in Bahasa Indonesia to help local officials who are not familiar with demographic and social analysis techniques. Documentation of survey and census questionnaires is also available in two languages, and tables are downloadable in Microsoft Excel type spreadsheets for ease of analysis.

Second, the website aims to be comprehensive. Tables range from the basic demographic statistics, through a growing number of economic and social variables and indices, and even a set of population projections at the provincial level. Tables are regularly updated, and an effort is under way to include survey and census results stretching back to the 1970s.

Third, the website fosters cooperation. Analytical papers on relevant demographic topics are posted on the website, and users are invited to contact the webmaster and Statistics Indonesia staff with comments, questions or criticisms.

The Indonesian effort to look at district level data arose out of a change of governance that saw budgetary authority moving from the central to the district level of government. Even if this had not occurred, a nation of over 220 million people living on thousands of islands would have required data at the levels of small units to reveal the heterogeneity of Indonesian society and its economy. Decentralization made this all the more important because local governments need data for budgeting, planning and evaluation. Now it is likely that those same officials will come to recognize the importance of looking at data across districts and perhaps across national boundaries. For instance the annual problem of smoke haze is best understood by looking at patterns of subnational units across the entire region of Indonesia, Malaysia and the Philippines.

The same is true for most nations of Asia and the Pacific. We need to look both beyond and below the nation state in our efforts to understand population change. This will require a commitment by all nations to develop sharper tools to define hierarchies of units similar to the NUTS in Europe, and more open data distribution systems similar to the Indonesian website for decentralized planning. The key is not simply a system of geographical mapping, but rather a strong hierarchical system of units defined for data collection and analysis. This will only be achieved if nations make their statistical systems more open and accessible and commit themselves to discussions that go beyond narrow national interests.

Endnotes

1. <http://www.prb.org/pdf06/06WorldDataSheet.pdf>
2. <http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTWDRS/0,,contentMDK:20227703~pagePK:478093~piPK:477627~theSitePK:477624,00.html>
3. <http://hdr.undp.org/>
4. http://www.unescap.org/esid/psis/population/database/data_sheet/index.asp
5. <http://europa.eu/scadplus/leg/en/lvb/g24218.htm> or
http://en.wikipedia.org/wiki/Nomenclature_of_Territorial_Units_for_Statistics
6. <http://www.datastatistik-indonesia.com>