

The views expressed in this paper are the views of the authors and do not necessarily reflect the views or policies of the Asian Development Bank (ADB), or its Board of Directors or the governments they represent. ADB makes no representation concerning and does not guarantee the source, originality, accuracy, completeness or reliability of any statement, information, data, finding, interpretation, advice, opinion, or view presented.

## ENHANCING SOCIAL AND GENDER STATISTICS

### **Gender and Social Statistics**

For purposes of this paper, social statistics is taken to refer to statistics in the areas of poverty, education, fertility and mortality, health and nutrition and the environment. Statistics are compiled from data collected principally from censuses, sample surveys, and administrative records. Statistics serve two main purposes, viz., they summarize information in such a manner as to make them more usable and they enable inductive conclusions or the making of generalizations about some population on the basis of a sample drawn from the particular population.

This paper discusses methods of enhancing readability and improving the explanatory capability of gender and social statistics that are, in general, currently available in the Asian Development Bank's developing member countries (DMCs). To enable this there is a need for collection of additional information, for information at the sub-national levels, for increased reliability in the information collected, for increased precision in statistical estimates, for increased timeliness in the statistics generated as well as for the statistics to be made available at regular intervals. With increasing globalization and development of communication technology, there is a pressing need for more timely information.

By its very nature, globalization implies both broadening and deepening. Phenomena that once affected a particular country or region now have broader implications and of necessity must include a larger set of countries. What were previously considered discrete problems can no longer be understood in isolation. Additionally, there is now a great overlap between the local and the global. What happens in one sphere inevitably spills into the other. The fact is that we live in a inter-connected world. This has further accelerated the need for more standardization in measurement concepts, definitions and measures and comparable development in information between countries.

Like race, ethnicity, and class, gender is a social category. It refers to those socially constructed roles, behaviours and expectations associated with being a male or female. There is increasing awareness that these gender differences cause policies to affect women and men differently. In fact, the importance of gender in the analytical framework of social data has been empirically proven. Policy makers have become acutely aware of the importance of integrating gender issues into all development issues and plans.

All societies experience gender asymmetries, that is, differences and disparities that are based on gender differences. While the magnitude of such differences and disparities varies between societies and over time, the content, direction and rate of these changes are influenced both by policy and socio-economic changes.

In the last century, especially in the second half, there were significant changes in gender differences. This was particularly so in the developing countries. The following are some examples of these changes:

1. Female education levels improved considerably. The primary school enrolment rates of girls about doubled in the South Asian countries, rising faster than boys' enrolment rates. This substantially reduced the large gender differences in schooling.
2. The life expectancy of women's in the developing countries increased by about 15 - 20 years. For the first time, in the 1990s, women in South Asia are on average, living longer than men.
3. More women have also joined the labour force. Since 1970, women's labour force participation has risen. On average it has risen by 15% points in East Asia. The growth was larger than that for men, again reducing the gender differences in employment. Gender differences in wages have also narrowed.

However, despite such changes, there still remain significant gender differences in rights, resources and voice in all developing countries. These inequalities have serious

implications for poverty reduction and sustainable development. Unfortunately, some of the achievements in these areas have been affected by the recent socio-economic crisis felt in many of the developing countries of Asia.

## **Development**

Improvement in the quality of people's lives is the ultimate development goal of all countries. The promotion of 'higher standards of living' is also set forth as a general goal of international economic and social activity in the Charter of the United Nations. In practical terms, this essentially implies the elimination of poverty and continuous efforts to raise the levels of living of all persons in a country.

For development to be sustainable, policies have to be based on a consideration of the interaction of economic, social and environmental factors. It has been proved empirically that economic development alone, or growth alone, will not be a sufficient condition for the elimination of poverty for all and sustainment of development. The causes of poverty are complex and go beyond economics. Political, cultural, social, and religious institutions and processes are also involved in the production and perpetuation of poverty. In fact, it has been proven that the social dimensions of development are important considerations for individual well being and institutional performance.

It has also been established empirically that ensuring gender equality, that is, ensuring equal access to rights, opportunities, and productive resources is an important requirement for developmental progress. However, identifying the issues related to gender inequality, especially across the developing world, is difficult due to a lack of gender disaggregated or differentiated data. There is a need for a greater understanding of the links between gender and development for the formulation of effective policies. Hence the emphasis on gender disaggregated data.

The goals and targets to be achieved in the areas of gender and other social aspects are taken as those based on the United Nations Millenium Declaration. Our concern is with

public policy issues and specifically the process of communicating information to policy makers. The gender and social statistics that are collected are to be used to compute indicators for monitoring and measuring progress (or a lack of progress) in these areas. The indicators to be used are those measures recommended by the international bodies for these areas.

### **The Millenium Development Goals**

The first stage in the collection of any data is the determination of the purpose to be served by the data, that is, the questions to be answered. Countries and the international bodies have agreement on what should be the goals and targets to work towards. Appendix 1 lists these goals and targets.

The success in achieving these targets depends on the effectiveness of the government policies. These policies, in their turn, depend for their formulation on the availability of relevant, reliable and up-to-date information. Policies based on the wrong information will only compound the existing problems. The policies will be formulated on the basis of current statistics and statistical trends.

While it is necessary to determine goals and targets for the development of a country, the successful achievement of these goals and targets, requires that the progress in achieving these targets and hence the goals are closely monitored. Changes have to be detected early and their magnitudes measured. This requires implicit agreement on the measures to be used to identify and measure change and that these measures have to be quantifiable. As stated earlier, the international bodies have recommended specific indicators that should be used for this purpose. The choice of these indicators, among other considerations, has as its basis, empirical findings in different parts of the world.

## **Indicators**

Indicators are not overall measures. An indicator provides a clue or points to a matter of larger significance. As such, an indicator's significance extends beyond what is actually measured to a larger phenomenon of interest. It makes perceptible a trend or phenomena that is not immediately detectable. Indicators are not primary data but primary data translated into information. Increasingly, users are not interested in the data per se but in the information that they provide. Changes in the indicators will indicate the progress or lack of progress of the policies put in place. The determination of the indicators to be used to monitor progress is difficult.

Indicators have two main characteristics:

1. They quantify information so the significance of the information is readily apparent
2. They simplify information about complex phenomena to improve communication.

Indicators are distinct from statistics or primary data. Their position in the information system can be envisaged as follows: indices as being the top of the information pyramid, followed by indicators, then by statistics and primary data being at the base of the pyramid.

The availability of indicators is not a sufficient condition for their use. Indicators play a useful role only where decision making is responsive to information about new social issues or the effectiveness of current policies. The need for comparability in the way indicators are formulated and calculated has already been discussed.

Fortunately, we do not have the task of determining the indicators to be used in the areas of social concern. Our task is to collect the data to compile the indicators for our

respective countries. The implicit assumption being that the indicators will be based on accurate, reliable, representative, comparable information that is up-to-date. While historical data has its uses, as stated earlier, in the current era there is an increasing need for up-to-date and relevant information that identifies trends and measures the changes taking place.

The proposed indicators should be gender disaggregated and available for sub-national levels in addition to at national levels. This is not to imply that these are the only compilations that are required. On the contrary, the indicators should also be compiled for population groups that are required for individual country needs. The more information at sub-national levels, the more understanding there will be of the actual situation and its links. This will enable better identification of the target areas and groups and the formulation of much more effective policies and projects. The level and depth at which these indicators should be compiled is dependent on the particular needs of the country, the availability of reliable and representative data and the timeliness with which the statistics can be made available.

The issues involved in the collection of the data concerned and in the compilation of the indicators are discussed next.

### **Generation of Statistics**

Different countries of the region have different levels of statistical development and as such have available varying levels statistical data. Priority has usually been given to the collection and compilation of economic data. It is only of late that there is increased interest in gender and social statistics data. The development of adequate gender and social data in the developing countries will take a number of years and it will in some measure be dependent upon the general economic and social development of the country itself.

#### *Population Censuses/Registration Systems/Household Surveys*

Information for quite a number of the proposed indicators can be obtained from population census data. The topics canvassed in a population census, among other items, generally include fertility, mortality, employment, education, literacy and more recently income in some instances. The implicit assumption is that the census data is relatively reliable. That is, the overall under-enumeration is within acceptable limits, the under-enumeration is not skewed and the non-sampling errors are also within acceptable limits. The wide coverage of a census enables compilation of the indicators at both the national and sub-national levels and for specific populations for the census years.

One can, for the topics canvassed in a population census, envisage a data collection system where the census data is used for the census years and for the intercensal years, information obtained from registration systems and household sample surveys are used to fill gaps in statistical data availability and to update the census information. This again assumes that data from the registration and household sample surveys are relatively reliable.

#### *Fertility, Mortality, Health*

Most demographic data come from essentially three sources, that is, from population censuses, registration systems and household sample surveys. In some countries, for example in Japan, population registers are maintained and used as a source for demographic statistics.

Some population registers are completed at a point in time, others are repeated at periodic intervals, and yet others are cumulative. A cumulative population register can be updated by the inclusion of other events e.g. deaths, retirement from the labour force, migration etc. However, a universal population register must be distinguished from a partial one that is confined to sub-parts of the population, and is used for administrative purposes. These partial population registers only cover those persons affected by a specific program and as such would have a more limited use.

In some instances the information from the different sources can be matched to provide the required statistics. For example, to calculate an age specific rate, the numerator may come from registration data and the denominator from the population census. Rates and ratios compiled from a combination of all these sources provide very important kinds of demographic data.

The most typical demographic product of a registration system is statistics on vital events, more commonly referred to as vital statistics. In the absence of other sources of demographic data, vital statistics can be used to construct a population or parts of a population. Vital registration records have a variety of useful purposes for public policy and programs that require reliable information. Social programs can be safely designed, monitored and evaluated with the aid of reliable vital records. Because of its extensive coverage and the legal compulsion, the vital registration system is an important source for the identification of sub-groups of the population in need of medical and health programs, nutritional programs, family planning services, maternal and child health care programs and other social services. Vital statistics can also be collected in a population census or estimated from successive censuses.

No single source can provide all the demographic data needed for analysis. The use of different sources would be necessary both to collect all the data required and to check on the reliability of the data. In view of the broad scope of information collected and their large geographical coverage, governments assume direct responsibility for censuses and registration systems. Private agencies might also collect similar data but through small surveys or through censuses of limited areas.

A civil registration system generally records the occurrence of events of birth, death, marriage, divorce, annulment, separation, adoption, legitimization and recognition in accordance with the legal requirements of the country. These events relate to an individual's life to death and all the changes in civil status that may occur during the individual's life - time. Not all countries record all of these events. However, the

information from this source can be compiled to provide vital statistics pertaining to the frequency and characteristics of the vital events.

Vital registration systems are common. About 150 countries of the world have a vital registration system and the system in more than half of these countries covers the whole country. Births and deaths are the events commonly registered. Birth records generally include the personal particulars of the mother and child, the father and the medical personnel involved. Birth weight and immunization information are also generally available. Similarly, a death registration also includes personal particulars of the deceased. In addition to the personal particulars of the deceased, there will be included information on the cause of death and attendant medical personnel. The reliability of information from any vital registration system depends on the geographical coverage of the system and the legal requirements for compliance in the country concerned.

In general, the following indicators can be compiled from data from population censuses and registration systems and for both national and sub-national levels:

1. Under Five Mortality Rate
2. Infant Mortality Rate
3. Proportion of 1 Year Old Children Immunised against Measles
4. Maternal Mortality Rate
5. Proportion of Births Attended by Skilled Health Personnel
6. Prevalence and Deaths Associated with Malaria
7. Proportion of Population in Malaria Risk Areas using Effective Malaria Prevention and Treatment Measures
8. Prevalence and Death Rates associated with Tuberculosis
9. Proportion of Tuberculosis Cases Detected and Cured under DOTS (Directly Observed Treatment Short Cases)

Where the registration systems do not include the information required or the coverage of the registration system is inadequate, recourse should be made to household sample surveys and health administrative records.

The completeness of coverage of the vital registration system and the promptness with which the events are registered and the accuracy of the information in the registration records will determine the reliability of the indicators compiled.

#### *Contraceptive use*

Data on the use of contraception is available in most of the developing countries, in fact more so than in the developed countries. The information is available from household sample surveys of knowledge, attitudes and practice of family planning, from the records of family planning and maternal and child health clinics. A combination of the clinic records and census and population estimates should enable regular compilation of the indicators and at sub-national levels.

#### *HIV/AIDS, Malaria, Tuberculosis*

While the registration of deaths in most countries includes information on the cause of death, records maintained at hospitals, doctors' clinics and health departments will have to be used to provide data on treatment of particular diseases. This treatment information might have to be used as a measure of the prevalence of the disease though the coverage would not be comprehensive. Some countries have placed emphasis on the identification and treatment of HIV/AIDS. The records compiled would also be another source of information on HIV/AIDS.

#### *Education, Literacy*

Education is one of the social characteristics of persons canvassed in population censuses and household sample surveys and occasionally in registration systems. In most countries, educational statistics from school systems are collected by the national

government through a reporting process in which individual schools report to local governmental units which in turn report to local government units which in turn report through a hierarchy of government units until it reaches the national government. The information so gathered generally refers to enrollments, graduates, teachers and other staff, educational finances, and schools and classrooms. As the data is by school, statistics at sub-national levels can be compiled. Gender and age are also collected in this process.

Enrolment data from school systems vary in quality depending on the attention given to statistical collection and reporting systems in countries and the adequacy of the number and skills of the personnel assigned to record the data.

Population censuses of many countries include statistics on the educational characteristics of the population. Most typically they include enrolments, illiteracy, and level of school completed. Since the time span between censuses may be long, household sample surveys that ask for educational characteristics as well as other population characteristics may be used to provide data on current educational characteristics. Use may also be made of ad hoc household surveys to collect current educational data. However, comparability in data will be affected if the populations surveyed are different, if definitions are not consistent and there are differences in the completeness and accuracy of the data reported.

Data on educational progression provides a basis for seeing to what extent population groups persist in school and to what extent continuation in school is related to normal grade progression. Measures of school retention are generally based on enrollment data. Census data are not useful sources unless they are taken at frequent intervals of time. The most frequently found source of data are the reports of school systems that give annual distributions of enrollment by grade and on the number graduating from school at different levels. Unlike enrolment rates, the data has to relate to the same cohort of persons.

Most countries obtain their information on literacy from population census data and from household sample surveys. These latter should enable estimates by gender, urban/rural and regions. To the extent there is misreporting on this topic there is an understatement of illiteracy. In a country with a high rate of illiteracy there is presumably no problem in identifying oneself as illiterate. On the other hand, in a country with a high rate of literacy, most people will be reluctant to classify themselves as illiterate. One check of the reliability of the data might be to analyse the illiteracy in a sequence of censuses or surveys. Cross tabulations against related variables is another way of checking on the reliability of the data.

The following indicators might be canvassed from data from these sources, that is from population censuses and school enrollment records and household sample surveys:

1. Net Enrolment Ratio in Primary Education
2. Proportion of Pupils Starting Grade 1 who reach Grade 5
3. Literacy Rate of 15 – 24 year olds
4. Ratio of Girls to Boys in primary, Secondary and Tertiary Education
5. Ratio of Literate Females to Males of 15 – 24 year olds

In the absence of other information, the denominator for computation of rates can also be estimated by ageing the population with the aid of estimated survival rates to an adjusted population census data.

### *Economic Activity*

Information on economic activity is collected from households in population censuses, household sample surveys and from establishments through regular reporting in certain kinds of administrative programs or through special sample surveys. The data that is generally available from sources other than households refer to employment and unemployment but not to the economically inactive persons.

Most of the national statistics on economic activity for households and their members are derived from population censuses and household sample surveys. Non-household sources of statistics include establishments, compulsory social insurance schemes, labour registrations and records of unemployment offices. Generally these sources do not cover the total economically active population and, in some, cover only a very small part of the whole. Groups generally excluded from these types of statistics include employers and own account workers, unpaid family workers and seasonal workers.

These other sources of statistics are primarily of use in evaluating the quality of data obtained from population censuses and household sample surveys. In this context it would be important to be aware of the coverage differences in the two sources of data. In developing countries, which often have a large proportion of workers in small - scale enterprises, one would not expect employment figures based on household and establishment surveys to agree. Nor would one necessarily expect trends in employment shown by the two series to be similar if a significant proportion of the employment is excluded by the establishment size cut-off for inclusion in the survey. The important shortcoming of these establishment statistics is that they do not include the social characteristics of the population that are required for social analysis. The other important difference is in the geographic classification in that the establishment statistics refer to the place of work while household survey statistics refer to the place of residence.

The indicators that can be compiled from these sources include:

1. Share of Women in Wage Employment in the Non – Agricultural Sector
2. Unemployment Rate of 15 – 24 year olds

Like in the case of the other characteristics, the data is subject to various errors and checks should be applied to the data. The conceptual problems of coverage, definitions etc. however continue to be a problem in a number of countries, in fact, more so than response errors.

## *Poverty*

Income statistics are of direct value to economists and others interested in wealth, the distribution and sources of consumer income, wages, salary rates, and the effective employment of manpower. Income is one of the best measures of economic well-being and it vies with educational attainment and occupation as a measure of socio-economic status. Income is a variable that has been introduced into population censuses only of late. Income data are much more widely collected through household sample surveys especially on sample surveys of household expenditures. A shortcoming of such surveys is that they are relatively small samples and as such do not facilitate analysis at sub-national levels. However, it should be possible with some of the sample sizes to compute indicators at rural/urban, by region and for the larger population and age groups.

Data on personal and family incomes are also available from establishment records e.g. income tax records, social security records. However, these records contain very few demographic and social characteristics for analysis. The data is confined to segments of the population that are not defined in demographic terms. Then again, not all recipients of income might be included in the coverage of the survey e.g. income below a cut-off point is not liable for taxation. The usefulness of these data are in evaluating the reliability of that obtained from the household sample surveys.

The recommended poverty indicators are:

1. Proportion of Population Below US\$ 1 per Day (PPP-values)
2. Poverty Gap Ratio
3. Share of Poorest Quintile in National Consumption
4. Prevalence of Underweight Children (under 5 years of age)
5. Proportion of Population below minimum level of dietary energy consumption

The most important purpose of a poverty measure is to enable poverty comparisons. For a more meaningful comparison it should be based on Purchasing Power Parities (PPP). The

objective of the comparison is to see the underlying trend as closely as possible. The PPPs refer to the ratio between the amount of national currency and the amount of a national currency needed to buy the same bundle of consumption goods in the two countries. Exchange rates reflect not only the prices of consumption goods but a host of other factors, including international movements of capital, interest rate differentials and government interventions. As a consequence, exchange rates exhibit greater variations over time than PPPs.

The Poverty Gap Ratio is a measure of the number of the poor relative to the poverty line. It measures the depth of poverty and as such is sensitive to changes in those already below the poverty line. A decline in the Poverty Gap Ratio would be reflective of an improvement in the current situation.

A poverty measure is a summary statistic on the economic welfare of the poor in a society. Unfortunately, there is no one universally accepted single measure of poverty. A number of different approaches can be used. However, for computing the poverty gap, estimates of individual welfare and the poverty line are required.

As stated earlier, there are a number of different approaches to measuring welfare. The approaches differ in terms of the importance attached to the individual's judgement of well-being as against a concept of welfare decided upon by others. The former approach would focus on measuring an individual's consumption of a bundle of goods and services. An example of the latter would be the defining welfare by the level of nutritional intake (though people do not live with only food or make food choices based solely on considerations of nutrition). The relative availability of the data required will also influence the approach used.

Developing countries generally place a higher weight on nutritional attainments. This emphasis is consistent with the behaviour of poor people in a specific society. The more common measure used is that of expenditure on total consumption, that is of all expenditure on all goods and services consumed, including the consumption of own

produce. The use of income data has been less frequent. This is also in part due to the difficulties in collecting income data and in measuring income. This has already been discussed. Current consumption is also more likely to give a better indication than current income of a household's typical, long-term economic welfare; income may fluctuate greatly over time, particularly in rural economies.

Poverty lines can be defined in a number of ways:

1. Absolute Poverty Lines

This is fixed in terms of the living standard indicator being used, be it based on the consumption approach or the nutrition approach. The poverty line is the same for all. That is, the same level of economic welfare is used to measure and compare poverty across the whole population. The poverty line itself may change, but only because of differences in the cost of a given level of welfare. Such poverty lines are more commonly used by the developing countries.

2. Relative Poverty Lines

The Relative Poverty Line approach is more commonly used by the developed countries. The poverty line is pre-determined. It is set, for example, in relation to income e.g. 50% of the national mean income. Then, if all incomes rise by the same proportion then there would be no changes in the relative inequalities. There will however be a rise in the poverty line, suggesting that people are better off. This however, in some instances, could be misleading. It has to be ascertained that the poor are in fact better off.

Empirical work has demonstrated that with economic development, poverty lines will go up but more slowly in the case of the poorer countries. As such, absolute poverty lines seem to be more appropriate for the lower income countries.

As discussed earlier, the most important source of the data required for these compilations are the household sample surveys of the national statistical offices. About two thirds of the developing countries have conducted such surveys, the information of which are representative nationally. Some of the samples are large enough to enable urban/rural compilations and in some cases for the larger regions of the country.

Measures of poverty focus on the population who are at the bottom of the income distribution. This in its turn requires information about the mean level of income as well as its distribution. The economic welfare need is measured by the Share in National Consumption of the Poorest Quintile.

Although quartiles are frequently used to describe income distributions because of their close relationship to the median (the median is the second quartile point) quintiles have come to be used extensively in describing family income distributions. Attention focuses on the lowest fifth and the share of the nation's aggregate income that they receive.

The importance of the data suggests that the data should be collected at regular intervals and at least to distinguish between gender and urban and rural populations. Cross tabulations by income reveal very interesting differentials in vital rates, social indices etc. information that further reinforces the need for close monitoring of poverty.

To compute the prevalence of underweight children under five years of age, there is a need for information on the number of children under five years of age weighed and the number of children under five with weight for age within the national reference values. This data are routinely collected by Ministries of Health at the national and sub-national levels for most countries.

To compute the proportion of the population below minimum level of dietary energy consumption would require information on food consumption energy and standards against which the levels are to be compared. Food consumption sample surveys and project records might be the sources for this information.

## *Environment*

The recommended indicators in this area are:

1. Proportion of Land Area covered by Forest
2. Land Area Protected to Maintain Biological Diversity
3. GDP per Unit of Energy Use
4. Carbon Dioxide Emissions (per capita)
5. Ozone Depletion
6. Accumulation of Global Warming Gases
7. Proportion of people with Sustainable Access to an Improved Water Source
8. Proportion of People with Access to Improved Sanitation
9. Proportion of People with Access to Secure Tenure

Data on land areas are obtainable from national forest inventories, cadastral maps which are compiled at periodic intervals, from remote sensing records, from land cover information, from agricultural census information and from administrative records.

National data on protected areas are available for most countries. Sub-national data are available for many. This data is generally available with the national agencies responsible for protected areas.

Information on access to water sources, sanitation facilities and tenure can be obtained from housing censuses. In some instances, supplementary information on the quality of water provided, the type of system used to provide the supply or sanitation facility may have to be obtained from the agencies supplying the facilities e.g. the national water authorities.

To compute energy consumption, data on the annual consumption of energy in all production have to be collected. The different energy inputs have to then be converted

into a common unit of measure such as terajoules (TJ) and aggregated as total energy. Using the annual GDP compiled estimates per unit GDP consumption can be calculated.

The current data on energy consumption is considered to be not very reliable. Special surveys could generate such data but it would be an additional expense and may not be one of the priority areas for statistical data collection of the developing countries. In addition, the computation of this indicator would be dependent on the availability of GDP estimates. GDP estimates are more commonly available and generally are more reliable. Unless both GDP and energy consumption estimates are available at sub-regional levels, it will not be possible to compute this indicator at levels below the national levels.

Data on carbon dioxide emissions are estimated based on activity data from fuel combustion, fuel emissions, industrial processes, solvent use, agriculture, land use change, forestry and waste. Emissions are calculated using emission factors associated with emissions of each gas for relevant activities. A greater degree of international comparability has been achieved by using default emission factors proposed by the Intergovernmental Panel on Climate Change (IPCC). National emission factors have been used, whenever available, which has resulted in increased precision of the estimates. There is generally annual detailed information available because of agreements by countries under the Montreal Protocol.

The data on sulphur oxide emissions is not as generally available as that on carbon dioxide. The need here would be to make the data available on a more frequent, preferably annual basis. Generally nitrogen oxide emissions are calculated with the help of emission factors that reflect the presence of nitrogen compounds in different types of fuels and other products. These emissions should be disaggregated by fuels, facilities or economic sectors. National emission factors should be used whenever available.

Ozone depleting substances are those organic substances containing halogenated substances that destroy the stratospheric ozone layer. The indicator is that of the consumption of ozone depleting substances. The data needed are that of production,

imports and exports of the substances. The data are available in most countries, on a national level, on a regular annual basis, as part of their reporting to the Montreal Protocol.

An increasing proportion of the population of developing countries live in urban areas. The majority of pollution sources tend to be found in or close to urban areas. As a result, the greatest potential for human exposure to adverse environmental conditions and subsequent health problems occurs in urban areas. Improving air quality is a significant aspect of promoting sustainable human settlements. Knowledge of air pollution concentrations is needed to define areas of non-attainment of air quality standards and to determine appropriate control measures on pollution sources. The synergistic effects among these pollutants may increase the potential for adverse health effects. As such, the data should at least be made available at urban/rural levels.

Air pollution monitoring is based on representative random sampling and chemical analysis of air samples. These samples are then compared for their statistical local parameters such as means and percentiles with the standards.

In discussing the statistics needed to compile the indicators recommended, the sources of the data and the limitations of the data have also been discussed. Following are some further considerations specifically on analysis of gender and social statistics data and on intercensal estimates.

### **Analysis of Data**

The first step in analytical work would be to check the data for errors. Then several interrelated procedures are performed to summarize and rearrange the data followed by computations and interpretations. Analysis of the data is however dependent on the data available, its reliability and representativeness.

The advantage of census data is its wide coverage. The use of census data will enable analysis at sub-national levels and the wide scope of data collected will enable links with a number of variables. The disadvantage of this data source is that in most instances it is only available once every ten years. The data collected will not be detailed as censuses are not subject specific. The data will also be subject to non-sampling errors and these are difficult to measure.

Household sample surveys, as against censuses, have the advantage of being cheaper and providing for better control and measurement of errors. Sampling theory will enable the calculation of sampling errors. However, the samples, in most instances, have the disadvantage of being only large enough to permit analysis at the national level. In the larger samples, it might be possible to derive urban/rural and regional estimates.

Data from registration systems suffer from limitations in the scope of the data available and the coverage of the system. The reliability of the data will be influenced by the pressure to register, the speed with which the event is registered and the level of training of the registration staff. The method of storage of the records will affect accessibility to the data. The data from this source while providing information on the geographical location or population covered might not enable generalizations to the whole of the country if it does not cover the whole country. Data from a part of the population represents the whole only if probability theory and practice had been used in the design and selection of the sample of the population. Only then will the information generated enable concluding to the whole. Reports based on studies of a part of the population should always have explicit statements on the nature of the methods used to collect the data and indicate to what extent the data may be regarded as representative of the whole country, if at all. These comments would also apply to data available from project records and general administration. It follows that data from any data source needs to be evaluated for its reliability and coverage before further analysis of the data is done.

While the focus of this paper has been on the compilation of the recommended indicators, it does not suggest they are the only measures to be compiled. The statistics that would be

required for the specific country needs should also be compiled. The statistics recommended as the minimum requirements in different areas should also be compiled. Examples are birth and death rates and life expectancies.

In understanding poverty, the poverty profile would be important. This profile should be constructed, where possible, for various sub-groups of the population such as for regions, educational levels, the sexes, ethnic groups, and employment sector. A good poverty profile can reveal a number of aspects of poverty reduction policies, for example, the regional or sectoral policies for spending. Another important analysis in this area would be comparisons over time. Poverty comparisons over time should be made to assess overall performance over time. Poverty measures are also linked to other measures, for example, to the adult literacy rate and to GDP per Capita. These links to other measures should also be made both to verify the data and to establish relationships.

Personal income data is always disaggregated by gender. Income also varies markedly with age, urban-rural residence, size of place, and in some countries with ethnicity. There is also interest in specific occupations and industries. Recently there has been great interest in the relationship between education and income especially life-time income. Studies of income also include consideration of the characteristics of the head (age, sex, marital status etc.) and of the family itself (size, composition, number of earners, number of income recipients etc.). When changes in personal or family income over time are considered, allowances should be made for changes in the cost of living. The information is generally shown in constant dollars or there will be an exaggeration of incomes if the values are shown in current dollars.

Mortality data shows significant variations in relation to characteristics of the deceased and certain characteristics of the event. In view of the very close relation between age and the risk of death, age may be considered the most important variable in the analysis of mortality. Other important characteristics are sex and the usual place of residence. Elements of primary importance characterizing the death are cause of death and place of occurrence of the death.

Also important in the analysis of mortality are marital status, socio-economic status, hospitalization and urban/rural residence. Other factors that would permit analysis of the significant social and economic factors in mortality are ethnicity and citizenship.

Mortality tends to vary with the characteristics of the community and the physical environment. These would include environmental conditions, degree of pollution, place of occurrence or usual place of residence in terms of specific geographic sub-divisions.

Data on education is linked with age, region, gender, unemployment, occupation and poverty. It has been found empirically that gender differences in education and health are often greatest among the poor. A recent study of boys and girls school enrolments in 41 countries by the World Bank indicates that within countries gender disparities in school enrolment rates are commonly greater among the poor than among the non-poor. Similar patterns across poor and non-poor households were seen with respect to boys and girls mortality rates for children under five years of age.

These patterns also emerged when poor and non-poor countries were compared. While gender equality in education and health has increased noticeably over the past 30 years in to-day's low income countries, disparities between males and females and males in school enrollments are still greater in those countries than in middle income and high - income countries.

The above are only some examples of possible gender and social statistics analyses.

### **Intercensal Data**

The emphasis is on sub-national statistics. As censuses with their wide coverage are the main source for such data, and, as censuses are conducted only between long intervals of time, the concern is with intercensal estimates. Ideally, most of these estimates should be available on an annual basis.

However, in the search for intercensal data sources it is important to note that changes in most demographic and social phenomena, even under induced development programmes, proceed relatively slowly. The precision of available measurement techniques is in many cases is also inadequate to the task of measuring the small changes which may be expected to occur within a single year. The impact of seasonal, cyclical and secular change (long term) may also cloud understanding of annual change.

It is also recognized that much of the data required for monitoring and evaluating progress are actually available through the administrative process – for determination of policy, administration of programmes and measurement of progress in respective spheres of operation. This information coupled with information from registration systems and household sample surveys should enable the compilation of at least reliable estimates for urban/rural and the large regions of a country at relatively frequent intervals.

#### References:

1. United Nations: Indicators of Sustainable Development, 1996
2. World Bank Institute: Introduction to Poverty Analysis; January 2002
3. World Resources Institute; Environmental Indicators, May 1995
4. World Bank Policy Research Report: Engendering Development, 2001
5. United Nations: International Definitions and Measurement of Standards and Levels of Living
6. Henry Shyrock, Jacob Siegel & Associates: The methods and Materials of Demography; October 1971, U.S. Department of Commerce
7. SIAP/UNDP: Statistics for Human Development Reports
8. ESCAP: Statistical Compendium on Women in Asia and the Pacific
9. UNDP: Urban Governance-Why Gender Matters
10. UNDP: Women's rights and Status-Questions of Analysis and Measurement
11. UNDP: Gender Analysis-Alternative Paradigms
12. UNDP: Gendered Governance-An Agenda for Change

