

# PIDE Policy Viewpoint

*PIDE Policy Viewpoints are carefully selected by the Director and Staff of the PIDE to convey important policy messages arising from PIDE research*

No. 4.

April 2007

---

## *Improving the Quality of Population Census 2008*

### **Introduction**

Census data generates information on total population count and its various socio-economic and demographic characteristics disaggregated by age and sex from national to provincial and down to district, tehsil and village level. The utility of these data is crucial for studying historical trends and for planning and formulating development programmes, especially for the recently devolved set up of administration. The next Population and Housing Census is likely to be carried out in 2008. The procedure of census data collection is lengthy and costly, and data compilation and its tabulation are subject to errors and inadequacies which have strong implications for policy formulation. Reporting errors during the data collection stage are also common, which need to be minimised.

### **Lessons Learned from Evaluation of 1998 Census**

It is important to learn from the experience of previous census of 1998 to improve the quality and methodology of next census data collection and make it more relevant to planning and development needs of the country. PIDE, in collaboration with the Population Census Organisation (PCO) and the UNFPA, undertook the initiative to evaluate the 1998 census data. A team of experts and demographers from PIDE and other relevant organisations was selected to undertake an evaluation of these data. Based on the analysis of micro-data of the 1998 census, the team produced 18 studies published by PIDE in an edited volume entitled: 'Population of Pakistan: An Analysis of 1998 Population and Housing Census'. The topics covered in the volume include population size and growth, its age and sex distribution, fertility, mortality, urbanisation and migration, ageing, housing, environment, and gender issues. The main findings are as follows:

- The 1998 census count is relatively better than the previous censuses in terms of its coverage and reliability, but there is evidence of under-enumeration of about 2 percent.
- The age and sex data continue to suffer from age misreporting, especially in the 0–4 year age group and for females and older population in rural areas.
- There are inconsistencies in defining urban areas in various censuses and urban population of large cities and migration levels are considerably under-estimated.
- The issues pertaining to the assessment of economically active population persist in the census data with a declining trend observed in economic activity which does not compare well with the Labour Force Survey data.
- Information on fertility, mortality, marital status and housing is of limited scope and needs to be used with caution.
- The data on literacy and education is inadequate in terms of measuring enrolment indicators and literacy rates are subject to reporting biases in age and being literate.

- The next census undertaking needs to take into account the adjustments and gaps identified in age distribution, reporting of children ever born, labour force, migration, urbanisation and literacy data through an improvement in census questionnaire, and in processing and methodology of data collection to obtain more reliable estimates of these parameters.

### **Census Data: Issues, Concerns and Suggestions for Improvement**

It is a welcome development that PCO has initiated the process of consultation with the stakeholders through holding a workshop in Islamabad on 'Census Methodology, its Issues and Concerns'. Among others, this workshop was also attended by those demographers of PIDE who were involved in the evaluation of 1998 census data. This workshop inspired them to share their viewpoint to improve the quality of next census data. They also held a series of internal meetings and discussions to draw some suggestions.

#### **1. Devolution Plan and Census Data**

The devolution process, initiated in 2001, has substantially restructured the government structure at sub-province levels (district, tehsil and union council). *To make the census data relevant to the devolution process, a new tier for union council may be introduced in the next census. Since tehsil is the main municipal unit in the new system for service delivery, all census data including sample-based information may be reported at the tehsil level with urban and rural breakdown.* It will make the census data more relevant for the devolution process.

#### **2. Post Enumeration Census Evaluation**

The post-census evaluation is essential to check the reliability of census data. This essential task was not performed in the 1998 population census. Thus, the accuracy in data collection and its coverage have been questioned, especially when population is counted on both *Defacto* and *Dejure* basis. The Census counts should clearly indicate the difference between the two types of population and how they are adjusted and reported in the published data. The reliability checks are normally performed by doing post-enumeration census evaluation. *PCO should plan for the post enumeration evaluation of the next census. A consistent and reliable sampling methodology may be evolved for this exercise. Involvement of third party in post-census enumeration would enhance the credibility of the census data.* The other consistency check on census data is that it should confirm with the demographic realities, such as fertility levels, known from other probability based surveys.

#### **3. Sample Survey**

For the sample-survey component of the census to administrator the long questionnaire, approximately 8–10 percent of the population is selected. This sample size is enough to make the results representative at the district-level by their rural-urban distributions.<sup>1</sup> As every district is self-selective, the use of one stage stratification sampling design is also theoretically consistent provided that it captures various socio-economic strata of population especially in the urban areas adequately. The worrying aspect, however, is the execution of the 8–10 percent sample in the field. The complete (100 percent) enumeration of a selected block, as has been done in the previous census, is against the spirit of sample surveys, which are designed in such a way that they capture maximum variance with minimum error at tolerable critical level. The presumed assumption in this complete enumeration is that large number of blocks (8–10 percent) to be selected at the national level would take care of between-the-blocks variance. While it may hold true at the national level, the other objective of making the survey results representative at the district or lower level may not

<sup>1</sup>This sample size also seems to be sufficient to make the census data representative even at the tehsil level.

be fully realised. *The suggestion is that the number of blocks may be increased and 20-30 households in each block may be selected, instead of enumerating all households in one block. The sample size in each district should be based on the probability proportional to size procedure.*<sup>2</sup>

#### **4. Age Reporting and Gender Statistics**

Age reporting has been a problem in all censuses. This problem was particularly observed for the 1998 census at the younger 0–4 and 5–9 age groups. This persistency in age misreporting is largely due to the fact that people do not remember their or their family member's exact age and would report to the nearest digit ending in fives or in zeros. There also appears to be a cultural problem in reporting of teenage girls. The system of national level of registration of vital events is recent and still incomplete to improve the age reporting problem. *Both recruitment of female enumerators and their proper training in probing techniques can reduce the misreporting of age.*

#### **5. Children Ever Born**

The 1998 population census of Pakistan included questions on number of children ever born (CEB) and on the number of births during the 12 months prior to the census date as well as the number of these births surviving at the time of the census. The evaluation studies first found that the 1998 census published data on CEB is defective, and the census births last year data contributes little to our knowledge of the level of fertility at the national level. Infant mortality rates are under-enumerated to the tune of 30-50% when derived from the census data. The value of census data lies mainly in its provision of estimates for sub-national units (district/tehsil) for which surveys either give unacceptably large sampling errors or no information at all. Unfortunately, assessing the variability of completeness of reporting of births for sub-national units in the census is problematic precisely because there are relatively few alternative sources on which to base assessments. There was a great concern on the reporting of births data among the participants of recent PCO workshop on census. It was suggested to drop all the birth-related questions from the census form. The concern is valid. But the question is: do we need data on CEB? The answer is definitely yes, particularly in the context of devolution process as well as achieving the health related MDG targets. *So, instead of dropping the questions, there is a need to improve the quality of reporting to the births to get reliable fertility and mortality estimates particularly at lower levels (districts or tehsils) to be used confidently by local governments. It is also suggested that data on more demographic variables may be collected such as addition of couple of new columns in the census form to enable estimation of maternal mortality using the time of death approach. Use of female enumerators, their better training and testing and revision, if necessary, of census questions on births could be effective tools to improve the reporting quality.*

#### **6. Urbanisation/Urban-Rural Classification**

Although the devolution plan has finished the urban-rural divide, ground realities are different. *PIDE recommends that the 2008 Population Census should be based on urban and rural classification. The use of administrative definition for defining an area 'urban' may be continued in the 2008 census but by combining it with the size-specific criterion. An 'urban' area may thus be defined as:*

1. *All metropolitan corporations, municipal committees, town committees or cantonments at the time of Devolution Plan.*<sup>3</sup>
2. *Adjacent areas of major cities may also be declared 'urban' by redefining their boundaries in collaboration of Tehsil Municipal Administration (TMA).*

<sup>2</sup>For districts with small population size, pre-determined minimum number of blocks may be selected to make the data representative at the district level.

<sup>3</sup>It is assumed that after the Plan no new locality has been declared urban.

3. *All rural localities with 10,000 or more population in 1998 may be declared 'urban' since most of these rural towns have the urban characteristics.*

*It is also suggested that to compute the urban density, area measurement of urban census blocks may be conducted and reported.*

### **7. Migration**

Internal migration in Pakistan, according to the census definition, refers to the inter-district movement of population. Intra-district movement, which could be the major migratory stream, is not recorded in the census. Information on the re-distribution of population within the district is particularly important for designing policies by district and tehsil governments under the new local government system. *Intra-district movement may be covered in the 2008 census.*

The 1998 census was probably the first census which has not generated information on the direction of movement e.g. rural-urban, urban-urban migration. *In census form, a sub-column to record the rural/urban division under the 'previous place (district) of residence' column may be added.*

Information on migration is collected through the long census form administered to 8-10 percent sample. *It is suggested that columns on migration may also be added in the short form to be administered to whole population.*

In 1981 census, information on overseas migration by country of destination was recorded. This information was missing from the 1998 census. *It is recommended that information on overseas emigration may be recorded in the 2008 census.*

### **8. Literacy and Education**

The census obtains information on literacy and educational attainment about each individual in the household through a few precise questions on the ability to read and write and the level of highest education completed. Learning to read and write from few years of schooling as well as from outside the school system raises questions about capturing reality and requires the use of consistent definition of literacy. Although the availability, quality and coverage of education statistics in the censuses have improved overtime, there are still limitations and anomalies in the age data and the related enrolment statistics.

One major limitation of census data on enrolment is that it does not provide current school attendance level of education, thus making it difficult to estimate level specific gross or net enrolment rates. Instead, it provides data on student population currently attending educational institutions by age and the highest passed without specifying their current enrolment/attendance at a certain level. *The suggestion to improve these data and enhance its utility for planning and development purposes is to add information on students' current level of attendance in addition to their highest grade completed.*

### **9. Economic Activity**

Labour Force Survey (LFS) and census data are not comparable mainly due to the different approaches and concepts adopted in these two sources. This shows lack of coordination between PCO and Federal Bureau of Statistics (FBS). The unemployment rates obtained from the 1998 census were very high as compared to 1981 census. It was mainly due to recording of a significant portion of 'unpaid family helpers' as 'unemployed'. *Information on economic activity as recorded in the population census should be consistent with the definition used in the LFS as well as Pakistan Socio-economic Living Standard Measurement (PSLM) survey, which has generated data on employment representative at the district level for the 2004-05 period. PCO and FBS may work together to make the Census, LFS and PSLM consistent in terms of generating the data on economic activity. Definition of adult population to record the economic activity may also be same in these data sources.*

## 10. Housing Census

In 1981 census, housing questions were canvassed separately as an independent census undertaken for the first time along with house-listing in December 1980. It provided detailed information on the country's housing stock. During the 1998 census, housing information was collected differently—being part of the population census. *It is suggested that the housing census may be conducted at the time of house-listing.*

*The 1998 census covered most of the core topics but missed some important questions like sewage disposal, and solid waste (garbage) disposal as both are considered important for assessing the impact on environment. Similarly, due to some improvement in provision of piped natural gas in both urban and rural areas, questions like heating of housing unit, availability of hot water, and availability of piped gas in housing units can be added.*

*Many sections of the housing questionnaire need thorough revision e.g. instead of collecting information on watching TV or listening radio, access to mobile phones can be added.*

*The information about the collective living quarters is not available in the published census data. Pakistan being a developing country is progressing fastly and future requirements need to have information about the collective living quarters. It may be added in the next census.*

## 11. Improving the Census Questionnaires

The census forms (or questionnaires) are the main instruments to generate good quality data. The questionnaires included in the 1998 census forms need thorough revision in the context of shortcomings of the previous forms used and socio-economic changes which took place during the current intercensal period. *It is suggested that PCO should constitute a committee for the construction/revision of census forms, consisting of demographers and other social scientists based at PCO, PIDE, NIPS, Population Council and FBS.* It is worth noting that for the PSLM survey, FBS constituted this type of committee which helped in producing a short but good questionnaire. *The census draft questionnaires may be disseminated widely to get comments from stakeholders for improvement.*

## 12. Recruitment of Field Staff and Training

Recruitment of field staff and their training to administer census forms are vital activities. The success of the census largely depends on these activities.

*Female enumerators may be recruited to gather information from female respondents.*

In the past enumerators were not paid according to their efforts in the field, resulting less motivation for work. *Their salaries and allowances should be increased.*

There is a need to make strong supervisory structure because it will help in improving the quality of the census data. *It is proposed that headmasters of the primary and middle schools, lecturers of colleges and other staff in different government departments may be hired as supervisors.*

*For the training of the field staff, especially the supervisory staff, experts in various fields such as demographers, economists, statisticians and related fields may be invited to share their experiences during training.*

*At the end of training, pre-testing of census forms may be arranged to judge the interviewing quality of enumerators. The checking of questionnaire filled during the pre-testing exercise may be ensured in the presence of enumerators.*

*Each enumerator and their supervisor should be provided a bag for carrying the material along with stationery and other materials. It will ensure the security of the field census forms.*

*The transportation facilities should also be provided to the field staff to perform their duties.*

In sum, this PIDE policy viewpoint is an attempt to contribute to the consultative process initiated by the PCO to enhance our understanding on concerns and issues relating to census data collection, its methodology and accuracy of data. There is a need to take steps in right direction to improve the quality of the next census data.

*This Policy Viewpoint has been prepared by  
Naushin Mahmood, Ghulam Yasin Soomro, G. M. Arif,  
M. Framurz Khan Kiani, and Khalid Hameed Sheikh.*

References and other material are available on PIDE website: [www.pide.org.pk](http://www.pide.org.pk)