South Africa - Project for Statistics on Living Standards and Development 1993

Southern Africa Labour and Development Research Unit

Report generated on: November 3, 2020

Visit our data catalog at: https://microdata.fao.org/index.php

Overview

Identification

ID NUMBER
ZAF_1993_PSLSD_v01_EN_M_v01_A_OCS

Overview

ABSTRACT
The Project for Statistics on Living standards and Development was a countrywide World Bank Living Standards Measurement Survey. It covered approximately 9000 households, drawn from a representative sample of South African households. The fieldwork was undertaken during the nine months leading up to the country's first democratic elections at the end of April 1994. The purpose of the survey was to collect statistical information about the conditions under which South Africans live in order to provide policymakers with the data necessary for planning strategies. This data would aid the implementation of goals such as those outlined in the Government of National Unity's Reconstruction and Development Programme.

KIND OF DATA
Sample survey data [ssd]

UNITS OF ANALYSIS
Households

Scope

NOTES
The scope of the study was:

(a) HOUSEHOLD QUESTIONNAIRE

Household Roster

Household Services

Food Spending and Consumption

Non-Food Spending

Education

Remittances and Marital Maintenance

Land Access and Use

Employment Status

Transport

Livestock

Health

Anthropometry

(b) COMMUNITY QUESTIONNAIRE

Demographic information

Economy and infrastructure

Education

Health

Agriculture

Recreational facilities

Shops and commodity prices

Literacy

TOPICS

Topic	Vocabulary	URI
Agriculture & Rural Development	FAO	
Food (production, crisis)	FAO	
Land (policy, resource management)	FAO	
Labor	FAO	
Livestock	FAO	
Nutrition	FAO	
Financial Sector	FAO	
Access to Finance	FAO	
Payment Systems	FAO	
Infrastructure	FAO	
Prices statistics	FAO	

Coverage

GEOGRAPHIC COVERAGE
National

UNIVERSE
All Household members. Individuals in hospitals, old age homes, hotels and hostels of educational institutions were not included in the sample. Migrant labour hostels were included. In addition to those that turned up in the selected ESDs, a sample of three hostels was chosen from a national list provided by the Human Sciences Research Council and within each of these hostels a representative sample was drawn on a similar basis as described above for the households in ESDs.

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
Southern Africa Labour and Development Research Unit	University of Cape Town

OTHER PRODUCER(S)

Name	Affiliation	Role

The World Bank		Technical assistance

FUNDING

Name	Abbreviation	Role
Government of Denmark		Financing the survey
Government of the Netherlands		Financing the survey
Government of Norway		Financing the survey

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Office of Chief Statistician	OCS	Food and Agriculture Organization	Adoption of metadata for FAM
DataFirst		University of Cape Town	DDI Producer

DDI DOCUMENT VERSION
ZAF_1993_PSLSD_v01_EN_M_v01_A_OCS_v01

DDI DOCUMENT ID
DDI_ZAF_1993_PSLSD_v01_EN_M_v01_A_OCS_FAO

Sampling

Sampling Procedure

(a) SAMPLING DESIGN

Sample size is 9,000 households. The sample design adopted for the study was a two-stage self-weighting design in which the first stage units were Census Enumerator Subdistricts (ESDs, or their equivalent) and the second stage were households. The advantage of using such a design is that it provides a representative sample that need not be based on accurate census population distribution in the case of South Africa, the sample will automatically include many poor people, without the need to go beyond this and oversample the poor. Proportionate sampling as in such a self-weighting sample design offers the simplest possible data files for further analysis, as weights do not have to be added. However, in the end this advantage could not be retained, and weights had to be added.

(b) SAMPLE FRAME

The sampling frame was drawn up on the basis of small, clearly demarcated area units, each with a population estimate. The nature of the self-weighting procedure adopted ensured that this population estimate was not important for determining the final sample, however. For most of the country, census ESDs were used. Where some ESDs comprised relatively large populations as for instance in some black townships such as Soweto, aerial photographs were used to divide the areas into blocks of approximately equal population size. In other instances, particularly in some of the former homelands, the area units were not ESDs but villages or village groups. In the sample design chosen, the area stage units (generally ESDs) were selected with probability proportional to size, based on the census population. Systematic sampling was used throughout that is, sampling at fixed interval in a list of ESDs, starting at a randomly selected starting point. Given that sampling was self-weighting, the impact of stratification was expected to be modest. The main objective was to ensure that the racial and geographic breakdown approximated the national population distribution. This was done by listing the area stage units (ESDs) by statistical region and then within the statistical region by urban or rural. Within these sub-statistical regions, the ESDs were then listed in order of percentage African. The sampling interval for the selection of the ESDs was obtained by dividing the 1991 census population of 38,120,853 by the 300 clusters to be selected. This yielded 105,800. Starting at a randomly selected point, every 105,800th person down the cluster list was selected. This ensured both geographic and racial diversity (ESDs were ordered by statistical sub-region and proportion of the population African). In three or four instances, the ESD chosen was judged inaccessible and replaced with a similar one. In the second sampling stage the unit of analysis was the household. In each selected ESD a listing or enumeration of households was carried out by means of a field operation. From the households listed in an ESD a sample of households was selected by systematic sampling. Even though the ultimate enumeration unit was the household, in most cases "stands" were used as enumeration units. However, when a stand was chosen as the enumeration unit all households on that stand had to be interviewed.

Weighting

A self-weighting sample design should in principle eliminate the need for weighting. A number of factors intervened, however, which made it essential to use weights after all. Amongst these was violence, which prevented survey teams from conducting interviews in two clusters on the East Rand; failure to continue interviewing in a cluster until the required take had been interviewed; and systematic under-representation of whites in the sample. This last problem resulted both from systematic non-response (whites were found to be more likely to refuse to be interviewed, or to be absent than other groups) and from sampling problems themselves. As a final comment on weights, the data provided for the user contains weights to correct for the enumeration difficulties discussed above as well as census based weights. If the user of the data wishes to use these weights, they are found in the data file named "weight02". The variable name for the enumeration-based weight is "rsweight" and the name for the census-based weight is "rcweight". (Do not use the "sweight" and "cweight" variables.)

Questionnaires

No content available

Data Collection

Data Collection Dates

Start	End	Cycle
1993-08	1993-12	N/A

Data Collection Mode

Face-to-face [f2f]

Data Processing

Data Editing

All the questionnaires were checked when received. Where information was incomplete or appeared contradictory, the questionnaire was sent back to the relevant survey organization. As soon as the data was available, it was captured using local development platform ADE. This was completed in February 1994. Following this, a series of exploratory programs were written to highlight inconsistencies and outlier. For example, all person level files were linked together to ensure that the same person code reported in different sections of the questionnaire corresponded to the same person. The error reports from these programs were compared to the questionnaires and the necessary alterations made. This was a lengthy process, as several files were checked more than once, and completed at the beginning of August 1994. In some cases, questionnaires would contain missing values, or comments that the respondent did not know, or refused to answer a question.

These responses are coded in the data files with the following values: VALUE MEANING
-1 : The data was not available on the questionnaire or form
-2 : The field is not applicable
-3 : Respondent refused to answer
-4 : Respondent did not know answer to question

Data Appraisal

Other forms of Data Appraisal

The data collected in clusters 217 and 218 should be viewed as highly unreliable and therefore removed from the data set. The data currently available on the web site has been revised to remove the data from these clusters. Researchers who have downloaded the data in the past should revise their data sets. For information on the data in those clusters, contact SALDRU http://www.saldru.uct.ac.za/.