DSDR | DATA SHARING FOR DEMOGRAPHIC RESEARCH

ICPSR 22626

India Human Development Survey (IHDS), 2005

Sonalde Desai
University of Maryland

Reeve Vanneman
University of Maryland

National Council of Applied Economic Research, New Delhi

Data Guide

Terms of Use

Information about Copyrighted Content

India Human Development Survey, 2005 (IHDS) Data Guide

Table of Contents

I.   Introduction

About the Guide

This Data Guide is an overview of the India Human Development Survey, 2005 (IHDS) and specific instructions for obtaining the IHDS datasets, which you can download to your own computer from DSDR. The Data Guide is also available as a webpage. IHDS users should refer to the User Guide (pdf), which provides greater detail on the topics discussed below. IHDS-II users should refer to the IHDS-II Data Guide.

About the Data

The India Human Development Survey, 2005 (IHDS) is a nationally representative, multi-topic survey of 41,554 households in 1,504 villages and 970 urban neighborhoods across India. Two one-hour interviews in each household covered topics concerning health, education, employment, economic status, marriage, fertility, gender relations, and social capital. Children aged 8-11 completed short reading, writing, and arithmetic tests. Additional village, school, and medical facility interviews are also available. A second round of IHDS (IHDS-II) re-interviewed most of these households in 2011-12 (N=42,152), and data for this wave can be found here.

The IHDS is a collaborative research program between researchers from the National Council of Applied Economic Research, New Delhi and the University of Maryland. The goal of IHDS is to document changes in the daily lives of Indian households in an era of rapid transformation. Additional information about the IHDS project is available on the India Human Development Survey website.

IHDS has four characteristics that make it unique among Indian surveys:
- breadth of topics including:;

- o caste & community
- o consumption and standard of living
- o energy use
- o income
- o agriculture
- o employment
- o government subsidies
- o education
- o social and cultural capital
- o household & family structure
- o marriage
- o gender relations
- o fertility
- o health
- o village infrastructure
- depth of human development indicators;
- a panel component; and
- a rich array of contextual measures.

All of these features make IHDS especially valuable for analyzing causal patterns underlying changes in human development. However, users who are primarily interested in descriptions of current levels of a particular human development indicator might prefer surveys with larger samples that are more narrowly focused on that topic (e.g., the National Sample Surveys, the Sample Registration System, or the National Family Health Survey). Users who want a state-level measure of a particular human development outcome are especially cautioned against relying exclusively on the smaller state samples of IHDS. IHDS's main purpose is to provide a means for gaining insight by analyzing the relationships among these human development outcomes and the connections between human development and its background causes.

IHDS and IHDS-II are designed to complement existing Indian surveys by bringing together a wide range of topics in a single survey. This breadth permits analyses of associations across a range of social and economic conditions. For example, studying children's outcomes (e.g., learning or immunizations) requires joint consideration of the role of poverty, family structure, gender relations, community context, and the availability of facilities. All of these are available in both IHDS surveys. A detailed survey topics and codebook indices list is available for IHDS here.

II.   Sample

IHDS was conducted in all states and union territories of India (with the exception of Andaman & Nicobar and Lakshadweep). The sample consists of 26,734 rural and 14,820 urban households. Of the 593 districts in India in 2001, 384 are included in IHDS. The sample is spread across 1,503 villages and 971 urban blocks. The IHDS household sample (N=41,554) is a composite of several separate subsamples that were each drawn somewhat differently. The basic division is between a re-interview sample of households previously interviewed in 1994-5 for the Human Development Profile of India (HDPI), N = 13,900, and new households, N= 27,654, but each of these divisions are themselves comprised of distinct components.[1]

---

[1] The numbers in this paragraph are found in the IHDS Technical Documentation (pdf), the IHDS Sample webpage, and in research publications referenced for the creation of this Data Guide.

Re-interview households:[2]

1. Re-interview households (N = 13,900)
   - 1a. Re-interview of HDPI (N=13,126)
     - 1a1. Original: total village (N=9,399)
     - 1a2. Original: partial village (N=2,165)
     - 1a3. Substitute: partial village (N=1,562)
   - 1b. Re-interview of Karnataka study (N= 774)

New households:

2. Replacement sample for lost village listings of HDPI (N = 4,051)
   - 2a. All district listings lost (N= 1,856)
   - 2b. Some village listings lost within a district (N= 2,195)
3. Refresher sample from HDPI districts (N = 7,843)
4. Extension sample from states and union territories not sampled in 1994 (N = 1,216)
5. Urban sample (N = 14,542)

See the User Guide (pdf) and Technical Documentation (pdf) for additional detailed information about sample design and implementation.

III. Data Elements

Data for IHDS are collected from multiple sources and made available in eight data sets (see Table 1 below). The questions fielded in IHDS pertaining to individuals and households were organized into two separate questionnaires; household and women. Each interview required between forty-five minutes and an hour and a half to complete. Because IHDS recognizes that all human development is nurtured within local and institutional contexts, separate questionnaires were developed to measure village characteristics and to assess the functioning of up to two schools and two medical facilities located within the selected villages. The survey was carried out in face-to-face interviews containing the following modules:

1. An interview with a knowledgeable informant – typically the head of the household – regarding socio-economic condition of the household including income, employment, educational status and consumption expenditure.

2. An interview with an ever-married woman aged 15-49 regarding health, education, fertility, family planning, marriage, and gender relations in the household and community.

3. Short reading, writing, and arithmetic knowledge tests were administered to all available children aged 8-11 in the household. These tests were developed in collaboration with researchers from PRATHAM, India, and were pretested to ensure comparability across languages.

4. Height and weight measurement of children under age 5, aged 8-11, and their mothers.

5. Facilities assessment of one government and one private primary school and primary health care facility in the community.

6. Village questionnaire assessing employment opportunities and infrastructure facilities in the village.

---

[2] The numbers in the Re-interview households and New households sections are found in the User Guide (pdf).

The survey instruments were translated into 13 Indian languages and were administered by local interviewers.

Users can access more specific technical information on instruments, assessments, and other data-related issues using the following links:

- Technical reports (including sample characteristics)

Table 1. List of Available Data Files

Part Number	File Name	Questionnaire (pdf)
DS1	Individual	Education-Health
DS2	Household	Household
DS3	Medical	Medical
DS4	Non-Resident	Household
DS5	Primary School	Primary
DS6	Birth History	Education-Health
DS7	Village	Village
DS8	Crops	Village

IV. Variable Names

All IHDS variables (except constructed variables) are named with a two letter code for the section of the questionnaire they are located, followed by the question's number (e.g., fm6a is from question 6a in the farm section of the data file).

Labels for variables record:

- on which questionnaire the variable is located;
- the page number of the questionnaire;
- the section number;
- the question number; and
- a very brief description.

For example, the variable label for id13, "HH3 1.13 Caste category", is from the household questionnaire ("HH"), page 3, section 1 ("identification"), question 13.

Information regarding constructed variables is available here. A crosswalk linking the variables in IHDS and IHDS-II is forthcoming.

V. Weight

The following is the weight variable and the datasets it appears in:

Table 2. Weight

Part Number	File Name	Weight	Weight Description
DS1	Individual	SWEIGHT	The IHDS sample is a complex combination of rural and urban samples (see the Sample section). To calculate population estimates for India or for individual states, SWEIGHT is needed as a design weight in all analyses. If doing individual cross sectional analyses, then use the appropriate individual survey weight, SWEIGHT for 2005.
DS2	Household	SWEIGHT	The IHDS sample is a complex combination of rural and urban samples (see the Sample section). To calculate population estimates for India or for individual states, SWEIGHT is needed as a design weight in all analyses. If doing individual cross sectional analyses, then use the appropriate individual survey weight, SWEIGHT.

VI. Merging Data Files

Within IHDS

Household and Individual Files

To merge the household and individual files, sort both in the following order by:

- STATEID
- DISTID
- PSUID
- HHID
- HHSPLITID

Alternatively, sort by IDHH, which is the 9-digit equivalent of the above variables. Merge or link these files using these sort variables. There are no household records without at least one individual record and no individual records without a household record.

School and Medical Facility Files

There are no direct linkages between the primary school file or the medical facility file and the household or individual files. A user can aggregate the school and medical facility files to the PSUID or DISTID level and integrate them with the household surveys to investigate the educational and medical context of the household.

Nonresident and Individual Files

The nonresident file can be appended to the individual file and sorted in the following order by:

- STATEID
- DISTID
- PSUID
- HHID
- HSPLITID
- PERSONID

(The nonresident PERSONID ranges from 50-54, beyond the range of the household members.). IHDS recommend renaming NR5 to RO3 (sex), NR6 to RO5 (age), NR7 to RO6 (marital status), and NR10 to ED5 (years of education).

Birth History File

The birth history file can be merged with the household file to add information about the household and the birth mother (the eligible woman variables EW3-EW9). Before merging, both files should be sorted in the following order by:

- STATEID
- DISTID
- PSUID
- HHID

- HHSPLITID

10,482 households have no birth history records. All birth history records will be matched to a household record, but there will be 270 birth history records from 57 households for whom the eligible woman is not identified on the household file (EW3 is missing). The merged birth history and household file can be merged with the individual file by:

1. Creating PERSONID=EW3;
2. Sorting by STATEID, DISTID, PSUID, HHID, HHSPLITID, and PERSONID;
3. Merging on those variables with the individual file.

The 270 birth history records with missing EW3 will not merge with any individual records. There will be 184,739 individual records that will not merge with any birth history records (because these individuals are not the eligible women who were interviewed).

Between IHDS and IHDS-II

IHDS and IHDS-II are panel surveys. IHDS-II re-interviewed about 83% of the IHDS households plus any split households that resided in the same community. Linking information is available at both the household and individual level. In order to link two rounds of data, you will require linking files which can be downloaded at the IHDS website. You will need to register in order to download these files. A Guide for Merging Files is also available.

Between IHDS and HDPI (Human Development Profile of India)
A linkage file is available through IHDS (ihdsinfo@gmail.com). It contains identification variables for the 1993-94 and 2004-2005 surveys, which allow for household level linkages for 13,081 households interviewed in both surveys.

VII. How to Obtain Data and Documentation Files

Downloading Data and Documentation from DSDR
Data from the India Human Development Survey 2005 (IHDS) are made available through DSDR, a data archive within ICPSR.

Researchers interested in downloading analysis-ready data and documentation files can do so free of charge through the DSDR website. Data are available in four formats: SAS, SPSS, STATA, and R. Raw ASCII data are also provided with accompanying setup (syntax) files. Documentation is provided in PDF format.

To download the IHDS data and/or documentation, researchers must agree to the Terms of Use. To download files, select the Data & Documentation tab. Click on the Download tab drop-down menu. Choose the file format you would like.

First Steps toward Obtaining Your Analytic File
Before downloading the data or beginning analysis, it is important for the user to become familiar with the IHDS User Guide (pdf) and Questionnaires (see Table 1).

VIII.   Learn More

Additional Resources
- IHDS Website
  - IHDS FAQ
- IHDS Bibliography
- ICPSR Abbreviated bibliography of IHDS-based Publications
- Questions regarding the using the IHDS project? Email IHDS.
- Questions regarding this Data Guide or IHDS data files? Email DSDR.

Acknowledgements