

ICPSR 36151

## **India Human Development Survey-II (IHDS-II), 2011-12**

Sonalde Desai

*University of Maryland, and National Council  
of Applied Economic Research*

Reeve Vanneman

*University of Maryland*

Data Guide

Inter-university Consortium for  
Political and Social Research  
P.O. Box 1248  
Ann Arbor, Michigan 48106  
[www.icpsr.umich.edu](http://www.icpsr.umich.edu)

## Terms of Use

The terms of use for this study can be found at:  
<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36151/terms>

## Information about Copyrighted Content

Some instruments administered as part of this study may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

### NOTICE

#### WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

# India Human Development Survey-II, 2011-12 (IHDS-II) Data Guide

## Table of Contents

**I. Introduction**

**II. Sample**

**III. Data Elements**

**IV. Variable Names**

**V. Weights**

**VI. Merging Data Files**

**VII. How to Obtain Data and Documentation Files**

**VIII. Learn More**

## I. Introduction

### About the Guide

This Data Guide is an overview of the India Human Development Survey-II, 2011-12 (IHDS-II) and specific instructions for obtaining the IHDS-II datasets, which you can download to your own computer from [DSDR](#). IHDS-II users should refer to the [User Guide](#) (pdf), which provides greater detail on the topics discussed below. IHDS users should refer to the [IHDS Data Guide](#).

### About the Data

The India Human Development Survey-II (IHDS-II), 2011-12 is a nationally representative, multi-topic survey of 42,152 households in 1,420 villages and 1,042 urban neighborhoods across India. These data are mostly re-interviews of households interviewed for IHDS ([ICPSR 22626](#)) in 2004-05. Both surveys cover all states and union territories of India with the exception of Andaman & Nicobar and Lakshadweep. Two one-hour interviews in each household covered topics concerning health, education, employment, economic status, marriage, fertility, gender relations, social capital, village infrastructure, wage levels, and panchayat composition. Children aged 8-11 completed short reading, writing and arithmetic tests. Additional youth, village, school, and medical facility interviews are also available. The initial round of [IHDS data collection can be found here](#).

The IHDS is a collaborative research program between researchers from the National Council of Applied Economic Research, New Delhi (NCAER) 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:
  - caste & community
  - consumption and standard of living
  - energy use
  - income
  - agriculture
  - employment
  - government subsidies
  - education
  - social and cultural capital
  - household & family structure
  - marriage
  - gender relations
  - fertility
  - health
  - village infrastructure;
- depth of human development indicators;
- a panel component; and
- a rich array of contextual measures.

In IHDS-II, one more module has been added for youth, 15-18, in the households to investigate transitions to adulthood in India. These youths are primarily the children who were administered the learning tests in IHDS.

All of these features make IHDS and IHDS-II 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 and IHDS-II. IHDS and IHDS-II'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-II on page 28 of the User Guide](#) (pdf).

## II. Sample<sup>1</sup>

---

<sup>1</sup> The numbers in this section are found in the IHDS [Technical Documentation](#) (pdf), [User Guide](#) (pdf), and research publications referenced for the creation of this Data Guide

IHDS is a nationally representative survey of 41,554 households conducted in 2004-2005. The initial IHDS sample consists of 26,734 rural and 14,820 urban households. The rural sample was drawn using stratified random sampling and contains 13,900 rural households who were interviewed in 1993-94 in a previous survey by NCAER (Human Development Profile of India (HDPI)) and 27,654 new households. The urban sample was a stratified sample of towns and cities within states (or groups of states) selected by probability proportional to population (PPP). 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.

IHDS-II re-interviewed 83% of the original households as well as split households residing within the village and an additional sample of 2,134 households. The final sample size for IHDS-II is 42,152 households; 27,579 rural and 14,573 urban. These households are spread across 33 states and union territories, 384 districts, 1420 villages, and 1,042 urban blocks.

For IHDS-II, some of the original IHDS households had to be replaced in some urban areas where interviewers were unable to locate the former households. In urban blocks and rural areas of northeastern states where 5 or more IHDS households were lost to attrition, the interviewers were asked to notify NCAER monitors of this loss. Once the loss was verified via physical check, a replacement household was randomly selected in the same neighborhood to refresh the sample. This has led to 2,134 new households being included in the IHDS-II sample. Replacement households are identified by a 9 in HHSPLITID.

**Table 1. Final Sample of Households<sup>2</sup>**

Households in IHDS	41,554
Households in IHDS-II	42,152
Households surveyed in both IHDS and IHDS-II	40,018
IHDS households lost to recontact for IHDS-II	6,911
IHDS-II households not included in IHDS	2,134

See the [User Guide](#) (pdf) for additional detailed information about sample design and implementation.

### III. Data Elements

Data for IHDS-II are collected from multiple sources and made available in fourteen data sets (see Table 2 below). The questions fielded in IHDS-II 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-II 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:

---

<sup>2</sup> The table includes numbers provided in the IHDS [Guide for Merging Files](#).

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, consumption expenditure, and social capital.
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. Those ever-married women who were interviewed in the initial IHDS wave, but were no longer eligible (i.e. older than 49 years of age), have also been interviewed.
3. An interview with youth in the households aged 15-18 years regarding education, employment, marriage, life skills, future planning, friendship and risky confidential behaviors.
4. 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.
5. Height and weight measurement of children under age 5, aged 8-11, their mothers, and other available household members.
6. Facilities assessment of one government and one private primary school and primary health care facility in the community.
7. Village questionnaire assessing employment opportunities and infrastructure facilities in the village.

The survey instruments were translated into 13 Indian languages and were administered by local interviewers. Please note, while IHDS-II follows the same general pattern as IHDS and many of the questions are identical, some questions were changed based on the experience with IHDS. Moreover, question numbers and variable names have also changed. Users are urged to consult the two sets of questionnaires and compare the question wording before trying to interpret the results.

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

**Table 2. List of Available Data Files**

<b>Part Number</b>	<b>File Name</b>	<b>Questionnaire (pdf)</b>
DS1	Individual	<a href="#">Education and Health</a>
DS2	Household	<a href="#">Income &amp; Social Capital</a>
DS3	Eligible Women	<a href="#">Education and Health</a>
DS4	Birth History	<a href="#">Education and Health</a>
DS5	Medical Staff	<a href="#">Medical Facility</a>
DS6	Medical Facilities	<a href="#">Medical Facility</a>
DS7	Non Resident	<a href="#">Income &amp; Social Capital</a>
DS8	School Staff	<a href="#">Primary School</a>
DS9	School Facilities	<a href="#">Primary School</a>
DS10	Wage and Salary	<a href="#">Income &amp; Social Capital</a>
DS11	Tracking	<a href="#">Tracking Sheet</a>
DS12	Village	<a href="#">Village</a>
DS13	Village Panchayat	<a href="#">Village</a>
DS14	Village Respondent	<a href="#">Village</a>

## IV. Variable Names

All IHDS-II 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., fm14a is from question 14a in the farm section).

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, “IS3 1.13 Caste category”, the variable label for id13, is from the Income & Social Capital questionnaire, page 3, section 1 (“identification”), question 13.

Information regarding [constructed variables is available starting on page 10 of the User Guide](#) (pdf). A crosswalk linking the variables in IHDS and IHDS-II is forthcoming.

## V. Weights

The following are the weight variables and the datasets they appear in:

**Table 3. Weights**

Part Number	File Name	Weight	Weight Description
DS1	Individual	WT	Sample weight for the household; most useful and usually used in almost all analyses
		FWT	Integer weight (truncated from WT) for STATA routines that require integer weight
DS2	Household	WT	Sample weight for the household; most useful and usually used in almost all analyses
		FWT	Integer weight (truncated from WT) for STATA routines that require integer weight
		INDWT	WT * NPERSONS - this represents the number of individuals in the household for analyses that require individual specific weights (e.g. Head Count Ratio for Poverty) when using the household-level file

		INDFWT	Integer value of INDWT
DS3	Eligible Women	WTEW	Weight for eligible women 15-49
		FWTEW	Integer value of WTEW
DS4	Birth History	WTEW	Weight for eligible women 15-49

## Weight Selection

Weights are a complex issue. IHDS recommends the following:

- If doing individual cross sectional analyses, then use the appropriate individual survey weight (WT for 2012).
- If doing a panel analysis, best approximation is to use the weights for 2005 (SWEIGHT) rather than 2012.

# VI. Merging Data Files

## Within IHDS-II

### *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 the 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.

### *Eligible Women and Individual Files*

Merging the eligible woman file with her individual information from the individual file should also be straightforward. The variable PERSONID is the eligible woman's person id on the eligible woman file and should merge exactly with the individual file.

### *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

Some households have no birth history records. All birth history records will be matched to a household record, but there will be some birth history records from 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; and
3. Merging on those variables with the individual file.

The birth history records with missing EW3 will not merge with any individual records. There will be some 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.

## **VII. How to Obtain Data and Documentation Files**

### **Downloading Data and Documentation from DSDR**

Data from the India Human Development Survey-II, 2011-12 (IHDS-II) 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-II 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-II [User Guide](#) (pdf) and Questionnaires (see Table 2).

## **VIII. Learn More**

### **Additional Resources**

- [IHDS Website](#)
  - [IHDS FAQ](#)
- [IHDS Bibliography](#)
- [Abbreviated bibliography of IHDS-II-based Publications](#)
- Questions regarding the using the IHDS project? [Email IHDS](#).
- Questions regarding this Data Guide or IHDS data files? [Email DSDR](#).

### **Acknowledgements**

This Data Guide was prepared by Sara C. Britt using IHDS-II documentation. It was developed for the Data Sharing for Demographic Research (DSDR), a project supported by the Population Dynamics Branch (PDB) of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (U24 HD048404). DSDR is housed within the Inter-university Consortium for Political and Social Research (ICPSR).