India Human Development Survey - II

A public use database for informed public policy

February 2011

Goals

Design
questions to understand processes leading to human development.

Collect
high quality database for tracking changes in Indian society.

Understand
changes in Indian society in an era of social and economic transformation.

Disseminate
data and results for open and unbiased debate on policy relevant issues.

IHDS

India Human Development Surveys I (2004-5) and II (2011-12) form part of a collaborative research program between researchers from the National Council of Applied Economic Research and University of Maryland. The goal of this program is to document changes in daily lives of Indian households in a society undergoing rapid transition. In documenting the way they live, work, educate their children, care for their aged parents, and deal with ill health, we seek to infuse the development discourse with the lived experiences of ordinary people.

Building on NCAER’s prior work in the 1990s, these surveys provide a rich empirical data base that will be available free of charge to a wide range of researchers in India and abroad, allowing for informed policy debates.

The IHDS-I data are available free of charge to researchers worldwide and are being used by about 2,000 users worldwide.

Survey Description:

IHDS – I is a nationally representative survey of 41,554 households conducted in 2004-2005. IHDS –II will reinterview these households in 2011-12. The sample is spread across 33 states and union territories and covers urban as well as rural areas.
Substantive Focus of IHDS

Education, Health, Livelihoods, Family Processes and Social Embeddedness

India has a long and distinguished history of survey research starting with the 1950s. However, most national surveys are single focus surveys, making it difficult to study interrelationships between different aspects of human development. Moreover, these are cross-sectional surveys allowing for snapshots of society at different points in time. Repeatedly interviewing the same households allows for a richer understanding of which households are able to partake in the fruits of growth and the process through which they are incorporated in or left out of a growing economy.

IHDS-I (2004-5) collected extensive data on education, health, livelihoods, family processes and the way in which households are embedded in a broader social structure. Information was also collected about social and policy contexts through survey of village infrastructure and markets and one private and one government school and medical facility in each village/block. About 13,000 households were also part of an earlier survey conducted by NCAER in 1993-94.

In 2011-12, IHDS-II will re-interview all of these households as well as split households (if located within the same village or town) to trace changes in their lives.

Logistics permitting, it is our desire to continue to survey these households over the coming decade to provide a rich description of changes in Indian society.

Two sources of changes in household well-being are of interest: (1) Changes in broader society and their influence on households; and, (2) Life cycle changes as individuals age and face new roles and responsibilities.

The household panel of IHDS-II will primarily focus on the first. However, a special youth module added to IHDS-II (called IHDS-Y) will begin a panel of young people aged 15-18, to study ways in which early life experiences affect subsequent outcomes and personal, familial and social resources that allow some individuals to transcend their background.
The IHDS-I is a nationally representative survey of 41,554 urban and rural households. It covers all states and union territories of India – with the exception of Andaman/Nicobar and Lakshadweep. These households are spread across 33 states and union territories, 384 districts, 1503 villages and 971 urban blocks located in 276 towns and cities. Size of the urban sample is 14,544.

Villages and urban blocks formed the primary sampling unit (PSU) from which the households were selected. Urban and rural PSUs were selected using a different design. In order to draw a random sample of urban households, all urban areas in a state were listed in the order of their size with number of blocks drawn from each urban area allocated based on probability proportional to size. Once the numbers of blocks for each urban area were determined, the enumeration blocks were selected randomly with help from Registrar General of India. From these Census Enumeration Blocks of about 150-200 households, a complete household listing were conducted and household sample of 15 households was selected per block.

The rural sample contains about half the households that were interviewed initially by NCAER in 1993-94 in a survey titled Human Development Profile of India – HDPI and the other half of the samples were drawn from both districts surveyed in HDPI as well as from the districts located in the states and union territories not covered in HDPI. The original HDPI was a random sample of 33,230 households, located in 16 major states, 195 districts and 1,765 villages. In states where the 1993-94 survey was conducted and recontact details were available, 13,593 households were randomly selected for re-interview in 2005. After a gap of 11-12 years, about 82% of the households were contactable for re-interview resulting in a resurvey of 11,153 original households as well as 2,440 households which separated from these root households but were still living in the village.

In order to check the representativeness of the sample, in each district where re-interviews were conducted, two fresh villages were randomly selected using a probability proportional to size technique. Comparing the panel sample with this randomly selected refresher sample allows us to determine whether this panel sample is overrepresented among certain segments of the society. The comparison between reinterview sample and fresh sample in target districts suggests that on most variables of interest such as caste, religion, education and economic status, the re-interviewed sample does not differ substantially from the fresh sample.

Additionally 3,993 rural households were randomly selected from the states where the 1993-94 survey was not conducted, or where re-contact information was not available. This approach to combining a randomly selected panel sample while refreshing it with another random sample has been used in a variety of surveys including the Panel Study of Income Dynamics in the U.S. and IHDS-II will be the first national panel survey since it will re-interview all households surveyed by IHDS-I.
Sample Design ...Continued

Malaysian Family Life Survey.

The questions finally fielded in IHDS were organised into two separate questionnaires, household and women. The household questionnaires were administered to the individual most knowledgeable about income and expenditure, frequently the male head of the household; the questionnaire for health and education was administered to a woman in the household – most often the spouse of the household head. Questions on fertility, marriage, and gender relations in the households were addressed to an ever-married woman between 15 and 49 in the household. If no household member fit these criteria, that portion of the questionnaire was skipped (about 19% of all households); if the household had more than one ever married woman between 15 and 49, one woman was selected randomly to answer those questions.

Fieldwork Organisation for IHDS I and II

Field work for IHDS-I was performed by 25 agencies throughout the country selected for their experience with administering large scale scientific surveys. The length and diversity of IHDS required more extensive training than is needed for single topic surveys. NCAER staff, assisted by researchers from the University of Maryland, organized eleven two-week training sessions across the country, each for 15-50 interviewers. Classroom reviews of each questionnaire section was combined with supervised field experience. In addition to written interviewer manuals, training films were developed.

Once trained, interviewers went into the field typically in teams of five: two pairs of male and female interviewers and a team leader. The team leader usually conducted the village, school, and medical facility interviews in addition to supervising the team.

Interviewers were supervised through random visits by supervisors and zonal co-ordinators working for NCAER and 2.5% of the households were partially re-interviewed.

IHDS-II will augment this by obtaining time/place stamp using a GPS enabled phone. Interviewers will be asked to upload these data on our central server daily in order to facilitate survey tracking and monitoring interviewer progress.
Comparison of IHDS estimates with other data sources

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>26</td>
<td>31</td>
<td>25.32</td>
<td>27.78</td>
</tr>
<tr>
<td>% literate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 5+</td>
<td>66.62</td>
<td>67.35</td>
<td>66.28</td>
<td>NA</td>
</tr>
<tr>
<td>Age 7+</td>
<td>68.08</td>
<td>68.62</td>
<td>67.33</td>
<td>64.8</td>
</tr>
<tr>
<td>Caste</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other backward classes</td>
<td>41.79</td>
<td>39.6</td>
<td>40.96</td>
<td>NA</td>
</tr>
<tr>
<td>Scheduled Castes</td>
<td>21.14</td>
<td>19.2</td>
<td>19.59</td>
<td>16.2</td>
</tr>
<tr>
<td>Scheduled Tribes</td>
<td>7.06</td>
<td>8.4</td>
<td>8.64</td>
<td>8.2</td>
</tr>
<tr>
<td>Other</td>
<td>30.01</td>
<td>31.9</td>
<td>30.81</td>
<td>NA</td>
</tr>
<tr>
<td>Religion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hindu</td>
<td>80.29</td>
<td>81.6</td>
<td>82.16</td>
<td>80.5</td>
</tr>
<tr>
<td>Muslim</td>
<td>14.01</td>
<td>12.5</td>
<td>12.66</td>
<td>13.4</td>
</tr>
<tr>
<td>Christian</td>
<td>2.16</td>
<td>2.7</td>
<td>2.09</td>
<td>2.3</td>
</tr>
<tr>
<td>Sikh</td>
<td>1.53</td>
<td>1.6</td>
<td>1.88</td>
<td>1.9</td>
</tr>
<tr>
<td>Buddhist</td>
<td>0.65</td>
<td>0.9</td>
<td>0.64</td>
<td>0.8</td>
</tr>
<tr>
<td>Jain</td>
<td>0.27</td>
<td>0.3</td>
<td>0.28</td>
<td>0.4</td>
</tr>
<tr>
<td>Others</td>
<td>1.11</td>
<td>0.4</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>% currently in school (age 5-14)</td>
<td>80.37</td>
<td>NA</td>
<td>82.58</td>
<td>NA</td>
</tr>
<tr>
<td>Knowledge of AIDS (women)</td>
<td>53.64</td>
<td>61</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>work participation rate for males</td>
<td>52.7</td>
<td>NA</td>
<td>54.7</td>
<td>51.7</td>
</tr>
<tr>
<td>work participation rate for females</td>
<td>31.9</td>
<td>NA</td>
<td>28.7</td>
<td>25.6</td>
</tr>
<tr>
<td>Average family size</td>
<td>5.2</td>
<td>4.8</td>
<td>4.69</td>
<td>5.3</td>
</tr>
<tr>
<td># children ever born to women (age 40 - 44)</td>
<td>3.7</td>
<td>3.9</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td># children ever born to women (age 45 - 49)</td>
<td>4.02</td>
<td>4.14</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>% women married (age 15 - 49)</td>
<td>73.17</td>
<td>75.23</td>
<td>75.78</td>
<td>76.61</td>
</tr>
<tr>
<td>% women married (all ages)</td>
<td>47.52</td>
<td>47.22</td>
<td>48.24</td>
<td>47.7</td>
</tr>
<tr>
<td>% electricity</td>
<td>71.85</td>
<td>67.9</td>
<td>65.2</td>
<td>55.8</td>
</tr>
<tr>
<td>% piped water</td>
<td>40.4</td>
<td>24.5</td>
<td>40.5</td>
<td>36.7</td>
</tr>
<tr>
<td>TV ownership (color or b/w)</td>
<td></td>
<td>(Colour)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPG use</td>
<td>33.3</td>
<td>24.7</td>
<td>21.9</td>
<td>17.5</td>
</tr>
<tr>
<td>% flush toilets</td>
<td>22.7</td>
<td>NA</td>
<td>19.2</td>
<td>18</td>
</tr>
<tr>
<td>% poor</td>
<td>25.73</td>
<td>NA</td>
<td>27.47</td>
<td>NA</td>
</tr>
</tbody>
</table>
Questions about Household:

- Income by source
- Employment and wages
- Remittances
- Education – quality and quantity, educational expenditure.
- Morbidity and treatment, medical expenditure
- Household consumption expenditure
- Intra-household relationships
- Social networks and organisational memberships
- Access to social safety nets
- Reading, writing and arithmetic skill tests for children aged 8-11 designed by PRATHAM
- Height and weight measurements for children and mothers

Institutional Modules:

Village Survey:

- Infrastructure availability
- Employment availability
- Prices and wage rates

Medical Survey

- One government and one private medical facility for treatment of minor illnesses
- Physical infrastructure
- Medicine stock
- Staff composition, training and presence

Primary School Survey:

- One government and one private school per PSU
- Physical infrastructure
- Staff composition, training and presence
- Mid-day meal availability

IHDS-II Plans:

Household questionnaire will be very similar to IHDS-I to allow for comparisons.

Some tweaking of institutional modules to capture new program.

“If you want to study change, do not change questions” – IHDS-II design principle.

Effort to capture NRHM and MNREGA use/impact.
IHDS-Y Panel
Survey of Youths Aged 15-18

IHDS-Y will be administered to 17,700 youths who were aged 8-11 in 2004-5. At that time, about 12,400 youths were administered short skill tests developed by PRATHAM. Extensive school history including education expenditures was also collected.

Thus, we have considerable prior data on these young people to begin a new youth panel to study processes through which transition to adulthood is shaped and personal, familial and social factors that facilitate or inhibit social mobility.

As inclusive growth emerges as a dominant challenge to a rapidly developing nation, equality of opportunity by caste, religion, gender and class is likely to emerge as a key policy challenge.

In February-May 2010 about 40 qualitative interviews were conducted by IHDS staff to provide guidance for development of the youth questionnaire. Based on these interviews as well as analyses of IHDS-I data, heterogeneity of educational quality across regions and social classes has emerged as a key theme to be explored in IHDS-Y.

Approximately 10% of the original sample of girls is likely to be married. Contact information for them will be collected for a possible phone or in-person survey in September 2012.

Focus of IHDS-Y

- Educational history
- Employment history
- Skill assessment to be developed in collaboration with PRATHAM
- Social networks and channels of information available to youth and parents about opportunities
- Marriage history/plans
- Educational and employment expectations and aspirations
- Engagement in risk taking behaviors such as cheating in examinations, smoking, drinking and engagement in sex. Information collected via sealed envelope to enhance reporting and confidentiality.

Major Innovation of IHDS-II is addition of a youth module administered to about 17,700 youths aged 15-18. A majority were also interviewed as 8-11 year olds in IHDS-I.
IHDS-I
A Unique Public Resource

IHDS-I was placed in public domain in July 2008 and availability of these data was advertised at professional association meetings as well as through advertisement in the Economic and Political Weekly. These data are now being used by over 2,000 users worldwide and have been used for diverse purposes such as analyses of returns to English language skills, calculations of intergenerational transfers in a global project on National Transfer Accounts, and studies of social safety nets.

In 2009, IHDS-I was ranked 7th in the frequency of download from ICPSR archive of over 7000 studies. It was the top most download for surveys outside of the United States. This reflects both growing interest in India as well as the breadth of IHDS coverage.

IHDS-I data are freely downloadable from the website of Interuniversity Consortium for Political and Social Research (ICPSR) located at the University of Michigan. Users are required to register so that they can be informed of new data availability or updates, however registration is free of charge. Data files are well documented and data are available in formats compatible with a variety of statistical packages. Most experienced users are able to download and use IHDS-I data in a matter of 3-4 hours.

Linking the 2004-2005 IHDS data with the NCAER survey of 1993-94 requires completing a confidentiality agreement and recognizing lack of support and detailed documentation.

It is expected that the IHDS-II data will be made publicly available in a similar fashion. However, to protect the privacy of the respondents, no information below the district level identifiers will be ever placed in public domain.

IHDS-I data are also available for online tabulation through Survey Documentation and Analysis system developed by researchers from University of California, Berkeley. This allows users to tabulate the IHDS-I data on-line without requiring special software.

More information about survey and research results in:

Funding for IHDS-I

IHDS-I was funded by grants from the U.S. National Institute of Child Health and Human Development. Supplementary funding was provided by The World Bank.

Funding for IHDS-II

Bulk of the funding for IHDS-II is provided by the U.S. National Institute of Child Health and Human Development, with supplementary funding by The Ford Foundation.

Principal Investigators – IHDS-II

- Sonalde Desai, Professor of Sociology, University of Maryland and Senior Fellow, National Council of Applied Economic Research
- Amaresh Dubey, Professor of Economics, Jawaharlal Nehru University and Senior Consultant, NCAER
- Reeve Vanneman, Professor of Sociology, University of Maryland

For further Information about IHDS, contact:
Program Manager, IHDS.
Email: ihdsinfo@gmail.com

Survey design, data access information and questionnaires at:
www.ihds.umd.edu