

Survey of Living Conditions

UTTAR PRADESH AND BIHAR

DECEMBER 1997 - MARCH 1998

**INTRODUCTION TO THE
DATA SETS**

JUNE 1998

SURVEY OF LIVING CONDITIONS IN UTTAR PRADESH AND BIHAR

This note briefly outlines some background information likely to be of use to analysts and researchers interested in using data from the 1997-98 UP-Bihar Survey of Living Conditions.

1. Background:

A two-part study of rural poverty was carried out in 1997-98 in south and eastern Uttar Pradesh and north and central Bihar. This study utilized both qualitative methods – rapid rural appraisal (RRA) & participatory rural appraisal (PRA) methodologies, and semi-structured interviews – as well as quantitative methods drawing on data collected from household and community surveys modelled after the World Bank's Living Standards Measurement Study (LSMS) surveys.

The data being distributed are from the quantitative component of the study, field work for which was carried out between December 1997 and March 1998. Data were collected through household and village-level questionnaires in 120 villages drawn from a sample of 25 districts in UP and Bihar states; a total of 2,250 households were interviewed during the course of the survey (more details on distribution of the sample are provided in the sampling section of this note). Of the sample of 120 villages where the household and village surveys were conducted, 30 had been visited in the earlier qualitative component of the study, while the remaining 90 were drawn at random from the sample districts.

The household questionnaire comprised ten main sections, and collected information on (1) household demographics, (2) economic activities, (3) housing, (4) education, (5) health, (6) marriage and maternity history, (7) expenditures and durable goods, (8) vulnerability, (9) farming and livestock, and (10) remittances and transfers. The village questionnaire comprised seven sections covering (1) village characteristics, including size, caste composition and political structure, (2) access to facilities and services, (3) agriculture, irrigation and forestry, (4) employment and migration, (5) anti-poverty programs and organizations, and (6) changes over time, and (7) visits to facilities including the angawadi center, primary schools, health posts, and the PDS fair price shop.

2. Sampling information

Uttar Pradesh and Bihar, the two states selected for the study, are divided into 8 statistical regions: 5 in Uttar Pradesh (Himalayan, Western, Central, Eastern, and Southern) and 3 in Bihar (Southern, Northern, and Central).

Sampling Universe: The universe for the study comprised 4 statistical regions: 2 in Uttar Pradesh (Eastern and Southern), and 2 in Bihar (Northern and Central). Altogether, there were 55 districts in the area covered by the study: 24 districts in the 2 statistical regions in Uttar Pradesh, and 31 districts in the 2 statistical regions covered in Bihar.

In the first phase of the project, qualitative field work was carried out in 30 villages: 3 villages each from 4 districts in Bihar (Mungher, Jehanabad, Saharsa, and Vaishali), and 6 villages each from 3 districts in Uttar Pradesh (Banda, Allahabad, and Gorakhpur).

Sampling strategy: The sampling strategy followed for the quantitative study basically involved dividing the sample population into four main strata:

- | | | |
|----|---|---------------------|
| 1: | districts that were covered in the qualitative study in Bihar | (i.e. 4 districts) |
| 2: | districts that were covered in the qualitative study in Uttar Pradesh | (i.e. 3 districts) |
| 3: | remaining districts in the 2 selected regions of Bihar | (i.e. 27 districts) |
| 4: | remaining districts in the 2 selected regions of Uttar Pradesh | (i.e. 21 districts) |

All 12 villages in Stratum 1 that were covered in the qualitative study were included in the sample. Similarly, all 18 villages in Stratum 2 that were covered in the qualitative study were included in the sample. In each of these 30 villages, 30 households each were picked at random for the survey.

In stratum 3 and 4, 45 villages each were selected for the survey. A two-step procedure was used to select villages in these two strata: first, 9 districts were selected in each stratum using PPS. In each of the 9 districts, 5 villages were then selected at the second stage, again using PPS. In each of these 90 villages altogether, 15 households each were selected for the survey.

The distribution of the sample by state and district is as follows:

State:	District	Number of villages	Number of households
Bihar:	1. Jehanabad 2. Vaishali 3. Saharsa 4. Mungher	3 villages each district	30 households each village
	1. Bhojpur 2. Gaya 3. Saran 4. W. Champaran 5. Muzaffarpur 6. Samastipur 7. Madhepur 8. Baghalpur 9. Araria	5 villages each district	15 households each village
Total:	13 districts	57 villages	1,035 households
Uttar Pradesh:	1. Banda 2. Allahabad 3. Gorakhpur	6 villages each district	30 households each village
	1. Hamirpur 2. Bahraich 3. Faizabad 4. Siddharthnagar 5. Basti 6. Mau 7. Jaunpur 8. Ghazipur 9. Mirzapur	5 villages each district	15 households each village
Total:	12 districts	63 villages	1,215 households
OVERALL:	25 districts	120 villages	2,250 households

3. Data documentation:

The documentation accompanying the data sets comprises the following:

- (1) Introduction to the data sets
- (2) Field Manual
- (3) Household Questionnaire
- (4) Village Questionnaire
- (5) 2 Data Diskettes

4. Data diskettes:

In addition to electronic copies of the above documents, the diskettes contain 3 main zip files:

- (a) `househld.zip` – this file contains the 30 data sets corresponding to the household questionnaire
- (b) `village.zip` – this file contains the 19 data sets corresponding to the village questionnaire
- (c) `gendata.zip` – this file contains 2 generated data sets containing useful information pertaining to the sample

Note that all these data files are in Stata format; however, these data sets can be converted into files read by other statistical packages using any standard format conversion software such as `dbmscopy`, `stat-transfer`, etc.

Household data sets: Data collected through the household level in the Survey of Living Conditions is stored in 30 stata data sets called `sect01a` - `sect10b`. Details pertaining to the contents of each of these data sets, such as the correspondence between sections of the household questionnaire used in the survey and each of these data sets, the level of observation of each of these data sets, the name(s) of the variable that uniquely identifies each observation, is summarized in the table below.

Community data sets: In addition, data collected at the community level in the Survey of Living Conditions is stored in 19 stata data sets called `vill01a1`-`vill07d`. The contents of these data files is also described in the table below.

Generated Stata data sets: Two generated data sets containing information that is likely to be of use to analysts and researcher have also been included in the data diskettes (e.g. information on district codes, sampling weights, etc.). The generated data sets are named `HHLIST` and `PSULIST`.

SUMMARY OF STATA HOUSEHOLD DATA SETS

SECTION	PART	STATA DATA SET	UNIT OF OBSERVATION	IDENTIFIER
1. HOUSEHOLD INFORMATION	A. HOUSEHOLD ROSTER B. SOURCES OF LIVELIHOOD	SECT01A SECT01B	INDIVIDUALS HOUSEHOLDS	HHCODE & IDCODE HHCODE
2. ACTIVITIES	A. ACTIVITIES B. CASUAL WAGE LABOR C. LONG-TERM EMPLOYMENT IN AGRICULTURE D. SALARIED EMPLOYMENT E. BUSINESS / TRADE / MANUFACTURING	SECT02AD SECT02E	ACTIVITIES OF INDIVIDUALS HOUSEHOLD ENTERPRISES	HHCODE & ACTCODE HHCODE & ENTCODE
3. HOUSING AND ACCESS TO FACILITIES	A. HOUSING B. UTILITIES C. ACCESS TO FACILITIES	SECT03AB SECT03C	HOUSEHOLDS FACILITIES	HHCODE HHCODE & FCODE
4. EDUCATION	A. CHILD DEVELOPMENT B. FORMAL SCHOOLING	SECT04A SECT04B	CHILDREN AGED 0 - 6 YRS CHILDREN AGED 6 - 19 YRS	HHCODE & IDCODE HHCODE & IDCODE
5. HEALTH	A. ILLNESSES AND INJURIES B. IMMUNIZATION & DIARRHEA	SECT05A SECT05B	INDIVIDUALS CHILDREN AGED 0 - 5 YRS	HHCODE & IDCODE HHCODE & IDCODE
6. MARRIAGE AND MATERNITY HISTORY	A. MATERNITY HISTORY B. PRE- AND POST-NATAL CARE C. WOMEN'S ROLES	SECT06A SECT06B SECT06C	INDIVIDUALS CHILDREN AGED 0 - 5 YRS ONE WOMAN / HOUSEHOLD AGED 15 - 49 YRS	HHCODE & IDCODE HHCODE & IDCODE HHCODE & IDCODE
7. EXPENDITURES AND DURABLE GOODS	A. WORKSHEET B. FOOD EXPENSES AND HOME PRODUCTION C. NON-FOOD EXPENDITURES D. INVENTORY OF DURABLE GOODS	SECT07A SECT07B SECT07C SECT07D	CROPS CULTIVATED FOOD ITEMS NON-FOOD ITEMS DURABLE GOODS	HHCODE & CROPCODE HHCODE & FCODE HHCODE & NFCODE HHCODE & ITEMCODE
8. VULNERABILITY	A. FOOD AVAILABILITY B. LOANS C. SAFETY NETS	SECT08A SECT08B SECT08C	HOUSEHOLDS HOUSEHOLDS TRANSFER LINE ITEMS	HHCODE HHCODE HHCODE & TRANCODE
9. FARMING AND LIVESTOCK	A. LANDHOLDING B1. CROP PRODUCTION:- Q. 1 - 4 B2. FERTILIZER USE:- Q. 5 B3. FERTILIZER USE:- Q. 6 - 10 C1. LIVESTOCK:- Q. 1 C2. LIVESTOCK:- Q. 2 - 4 D1. FARMING ASSETS:- Q. 1 D2. FARMING ASSETS:- Q. 2 - 4	SECT09A SECT09B1 SECT09B2 SECT09B3 SECT09C1 SECT09C2 SECT09D1 SECT09D2	HOUSEHOLDS CROPS CULTIVATED HOUSEHOLDS FERTILIZERS USED HOUSEHOLDS LIVESTOCK OWNED HOUSEHOLDS FARMING ASSETS OWNED	HHCODE HHCODE & CROPCODE HHCODE HHCODE & FERTCODE HHCODE HHCODE & ANIMCODE HHCODE HHCODE & ASSTCODE
10. REMITANCES AND TRANSFERS	A. QUESTION 1 B. QUESTIONS 2 - 6	SECT10A SECT10B	HOUSEHOLDS TRANSFER PAYMENTS	HHCODE HHCODE & LNUMBER

SUMMARY OF STATA VILLAGE DATA SETS

SECTION	PART	STATA DATA SET	UNIT OF OBSERVATION	IDENTIFIER
1. VILLAGE CHARACTERISTICS AND INFRASTRUCTURE	A. SIZE, CAST COMPOSITION ...	VILL01A1 VILL01A2	VILLAGES TOLAS WITHIN VILLAGES	VILLAGE VILLAGE
	B. ECONOMY & INFRASTRUCTURE	VILL01B		
2. ACCESS TO FACILITIES	A. ACCESS TO FACILITIES	VILL02A	VILLAGES	VCODE
	B. ACCESS TO EDUCATION	VILL02B	VILLAGES	VCODE
	C. ACCESS TO HEALTH	VILL02C	VILLAGES	VCODE
3. AGRICULTURE AND FORESTRY	A. LAND AND IRRIGATION	VILL03A	VILLAGES	VCODE
	B. FORESTRY	VILL03B	VILLAGES	VCODE
4. EMPLOYMENT AND MIGRATION	A. PREVAILING WAGES	VILL04A	VILLAGES	VCODE
	B. MIGRATION	VILL04B1	VILLAGES	VCODE
		VILL04B2	VILLAGES	VCODE
5. PROGRAMS AND ORGANISATIONS		VILL05A	VILLAGES	VILLAGE
		VILL05B	VILLAGES	VILLAGE
6. CHANGES OVER TIME		VILL06	VILLAGES	VILLAGE
7. VISITS	A. ANGANWADI CENTER	VILL07A	VILLAGES	VILLAGE
	B. PUBLIC PRIMARY SCHOOL	VILL07B	VILLAGES	VILLAGE
	C. PRIMARY HEALTH POST	VILL07C	VILLAGES	VCODE
		VILL07C1	VILLAGES	VCODE
	D. PDS SHOP	VILL07D	VILLAGES	VILLAGE

SUMMARY OF STATA GENERATED DATA SETS

STATA DATA SET	UNIT OF OBSERVATION	IDENTIFIER
HHLIST	HOUSEHOLDS	HHCODE
PSULIST	VILLAGES	VILLAGE