

I. INTRODUCTION

The 1993 Indonesian Family Life Survey (IFLS) consists of a household survey (IFLS-HH) and a community-facility survey (IFLS-CF). This Codebook describes the IFLS-HH database. The database consists of 112 subfiles that contain the information collected across the seven books of the household survey, plus seven supplemental subfiles that contain sampling weights and geographic codes. Subfile names reflect the book and the subsection of the household questionnaire that the data represent, with subfile names from a given book beginning with the same prefix. Table 1.1 lists the seven books and the subfile prefixes for the IFLS-HH data.

Table 1.1
IFLS-HH QUESTIONNAIRE BOOKS AND SUBFILE PREFIXES

IFLS-HH Book	Description of Book Contents	Subfile Prefix
CA	Anthropometry	BUKC
K	Control Information	BUKK
I	Household Roster, Characteristics and Consumption	BUK1
II	Household Economy	BUK2
III	Adult Information (ages 15 and over)	BUK3
IV	Ever-married Woman Information (ages 15-49)	BUK4
V	Child Information (ages 0-14)	BUK5

This Codebook describes the variables included in each of the 119 subfiles. The Codebook also presents special notes about many variables that help provide further insight into the observed responses.

This Codebook should be used in conjunction with the *User's Guide*¹, and the *Questionnaires and Interviewer Manual*². The User's Guide discusses the file structure of the IFLS-HH and IFLS-CF databases, addresses issues about identifying individuals of interest and linking the various subfile types to produce analytic files, and discusses some data quality issues. While the Codebook describes the contents of the IFLS-HH subfile data sets, the Interviewer Manual and Questionnaires provide the details regarding the questions asked (i.e., English translations of the original Bahasa Indonesia) and the motivation behind those questions. A fourth document, the *Overview and Field Report*³, provides detailed information on the sampling frame, field work, and response rates. This Codebook and the above documents comprise the basis of the IFLS-HH data documentation.

Given the size and complexity of the IFLS-HH and IFLS-CF databases and the available project resources, the preparation of the public use files required a strategy that would meet basic user needs and make the data available to the research community in a reasonable time frame. All data was 100%-verified at data entry (i.e., double entered) and the data entry program contained checks on valid ranges and skip patterns, so data entry errors were basically nil. The basic approach, then, was to concentrate on those data cleaning activities which required access to information that was privacy protected, such as respondent's names for identifier code verification. Efforts also focused on trying to provide as much translated material as possible. Users should be aware that similar information was sometimes collected in more than one section and sometimes from different individuals. One data preparation activity that was not able to be done in much detail before public release was the examination of inconsistencies in responses by household members to the same item or event, or by a given respondent to the same event asked about in more than one place. In general, the public use files do not include efforts to reconcile possible differences. After public release, subsequent data cleaning efforts

¹The 1993 Indonesian Family Life Survey: Appendix E, *User's Guide*, DRU-1195/6-NICHD/AID, RAND, 1995.

²The 1993 Indonesian Family Life Survey: Appendix A, *Household Questionnaires and Interviewer Manual*, DRU-1195/2-NICHD/AID, RAND, 1995 .

³The 1993 Indonesian Family Life Survey: *Overview and Field Report*, DRU-1195/1-NICHD/AID, RAND, 1995.

sponsored by RAND projects will continue and results of those efforts will be made available to the IFLS user community.

After reviewing the contents of this Codebook below, we present descriptions of a few key variables found throughout the subfiles. In addition, we discuss proxy respondents, text responses to open-ended questions, missing/don't know/not applicable codes, rupiah amount variables, privacy protection, data cleaning and suggested data fixes, and sampling weights. At the end of this introductory section is a multi-page table that provides a quick reference to the subfile structure of the IFLS-HH database.

CODEBOOK GENERATION

The IFLS-HH Codebook is an edited version of the codebook generated by ISSA (Integrated System for Survey Analysis)⁴, the data entry program used for the IFLS. The 20 character limit on variable labels in ISSA resulted in some creative abbreviations and phonetic spellings used in variable labels so as not to exceed that limit. The data entry program created for the IFLS applied value labels and ranges only to categorical variables. Due to cost constraints, codebook editing was limited to the addition of new variables, the renaming (in English) of variables whose original names were in Bahasa Indonesia, the rewording of misleading variable labels or value codes, the revision of column locations to match public use raw data files, and the deletion of name and address variables to ensure privacy protection for respondents to both the IFLS-HH and IFLS-CF surveys. Some minor editing of variable labels was also done, and ranges were added for newly created numeric versions of rupiah amount variables.

The headers used in the IFLS-HH Codebook are those generated by the ISSA data entry program. The column titled "CLASS" should be ignored, as its meaning is only relevant to the ISSA data entry program and does not relate to the final IFLS-HH Codebook content.

Please note: the ranges noted in the Codebook are those from the original IFLS-HH data before any cleaning was done. All cleaning was done

⁴Survey data entry program provided and maintained by Macro International. ISSA is the data entry package used for the Demographic and Health Surveys. Original development was by the Institute for Resource Development at Westinghouse through the Demographic and Health Surveys Project funded by USAID.

to SAS® file versions and not to the original ISSA raw data files. No revisions were done to the ranges outputted by ISSA when the codebook was originally generated, except for newly added variables (e.g., numeric rupiah amounts). If cleaning changed the minimum or maximum value, the ranges in the IFLS-HH Codebook will be misleading. The ranges should only be used as a guide to the maximum number of digits for the variable and not used to check whether raw data versions have been read in correctly.

CODEBOOK CONTENTS

The IFLS-HH database comes in two forms: rectangular files and SAS® transport files. For each subfile, the Codebook provides variable locations for use with raw data versions and variable names for use with the SAS data versions. In most cases, the variable names reflect the question number represented by the variable, providing an easy link between the questionnaire and the codebook variables. Note that question numbers help the user relate the discussions in the Interviewer Manual to a given variable. Where relevant, special notes are presented with the description of a given variable. These notes include more detailed information on selected variables and/or their values plus citations for specific cases which may be of interest to the user.

The Codebook contains four appendices. Appendix A contains the occupation code descriptions that apply to the two-digit occupation codes added to the Book III employment history data. Appendix B presents a list of the enumeration areas completed before the second round retraining (see the *Overview and Field Report* for details). Households in those 32 enumeration areas may have an higher than expected rate of interviewer errors and miscodings. Users may want to delete households from those areas or examine those households more closely with respect to the variables of interest in the given analysis. Appendix C contains province identifier codes used in the IFLS and Appendix D contains such codes for kabupaten.

CODEBOOK ORGANIZATION

Within this Codebook, the subfiles are grouped by IFLS-HH questionnaire books. Each questionnaire book represents a section of this Codebook. The IFLS household questionnaire consists of seven books or sections. The contents of these books are presented in Table 1.2.

Table 1.2
1993 IFLS HOUSEHOLD QUESTIONNAIRE BOOKS AND MODULES

BOOK		MODULE SECTION AND CONTENTS
CA	Anthropometry	CA: Height and weight
K	Control Book	SC: Sampling/enumeration record IK: Recontact information PS: Within-hhld sample selection FP: Questionnaire tracking form
I	Household Roster	AR: Household member roster KR: Household characteristics KS: Consumption PP: Outpatient care provider knowledge
II	Household Economy	UT: Farm business NT: Non-farm business PH: Labor and nonlabor income HR: Household assets GE: Household economic shocks AK: Health insurance
III	Adult Information	DL: Education history TK: Employment history AW: Time allocation KW: Marital history BR: Birth summary (women 50+) MG: Migration history SR: Circular migration history KM: Tobacco smoking KK: Health condition MA: Acute morbidity PS: Self-treatment RJ: Outpatient utilization RN: Inpatient utilization BA: Noncoresident family roster and transfers TF: Other transfers HI: Individual assets, nonlabor income
IV	Ever-married Woman Information	KW: Marital history BR: Pregnancy summary CH: Pregnancy and infant feeding history CX: Contraceptive knowledge/use KL: Contraceptive calendar
V	Child Information	DLA: Child education history MAA: Child acute morbidity PSA: Child self-treatment RJA: Child outpatient utilization RNA: Child inpatient utilization

Individual measures of height and weight for interviewed adults and children were recorded in a single anthropometric record for each household (Book CA). Three sections of the questionnaire collected information at the household level: a control book (Book K) completed by the interviewer and household head or spouse of head and Books I (Household Roster) and II (Household Economy) administered to a knowledgeable household member (in most cases the household head or spouse of head). The three remaining questionnaire books collected individual-level data from adult respondents age 15 and above (Book III), ever-married women respondents aged 15 to 49 (Book IV), and, by proxy, young children aged 0 to 14 (Book V).

Table 1.8 at the end of this introductory section provides a quick reference to the subfiles described in this Codebook.

HOUSEHOLD IDENTIFIER VARIABLES

The key identification variables across all IFLS-HH household survey subfiles are: CASE and PERSON. Household-level subfiles contain only CASE; respondent-level subfiles contain CASE and PERSON. Below we will describe each variable rather than repeat the descriptions throughout the Codebook. Identifier variables in the IFLS-HH data which link to the IFLS-CF (community-facility data) are discussed in the next section.

CASE: Household Identifier

The variable CASE is a 9-digit numeric variable which contains the household identifier number. The household identifier number consists of four parts:

Province:	digits 1-2	(BPS codes, see Appendix C)
Kabupaten:	digits 3-4	(BPS codes, see Appendix D)
Enumeration Area:	digits 5-7	(IFLS codes 1-321)
Household number:	digits 8-9	

PERSON: Person Number.

The variable PERSON is a numeric variable which represents the individual's person number from the household roster. The values run from 1 to the number of people listed in the roster. The variables CASE and PERSON, then, uniquely define an individual across all households.

COMMUNITY AND FACILITY IDENTIFIER VARIABLES

Communities in the IFLS-HH data are represented by their IFLS Enumeration Area number (EA). Health facilities and schools mentioned in the household data are represented a facility code (FACCODE).

SC07, CASE: IFLS Enumeration Area

The EA for a household can be found in two ways. First, the variable SC07 in the Book K (Control Information) subfile BUKKSC1 contains the IFLS enumeration area code (sometimes referred to as the SAKERTI code). The variable SC07 can be merged onto any household survey subfile using CASE. Second, the household identifier number, CASE, has the EA code embedded within it: the 5th, 6th and 7th digits

The variable EA found on all the IFLS-CF subfiles corresponds to the household survey variable SC07.

FACCODE: Facility code that matches IFLS-CF data

Table 1.3 lists the IFLS-HH subfiles which contain references to facilities, many of which were interviewed as part of the Community-Facility Survey. The variable FACCODE is on each file. FACCODE links to the variable FACCODE on the IFLS-CF subfiles. Please note that not all facilities mentioned in the household survey responses were interviewed as part of the IFLS-CF; likewise, not all facilities interviewed appear among the household responses.

Table 1.3

IFLS-HH SUBFILES WITH FACILITY CODES FOR LINKING TO IFLS-CF DATA

IFLS-HH Subfile	Subfile content	Type of Facility
BUKKAR3	Roster education	School
BUK1PP1	Health Provider Knowledge	Health
BUK3DL2	Education	School
BUK3RJ2	Outpatient utilization	Health
BUK3RN2	Inpatient utilization	Health
BUK4CX1	Contraceptive Use	Health
BUK4KL4B	Contraceptive Calendar	Health
BUK5DL1	Education	School
BUK5RJ2	Outpatient utilization	Health
BUK5RN2	Inpatient utilization	Health

NOTE: File BUK4KL4B has three facility code identifiers: FACCODE1, FACCODE4, FACCODE6.

The variable FACCODE is a 7-digit character variable. The first three digits contain the EA code. The fourth digit identifies the type of facility:

- 1: Government Health Center (Puskesmas)
- 2: Private Doctor/Clinic
- 3: Nurse/Midwife/Paramedic
- 4: Community Health and Family Planning Post (Posyandu)
- 5: Traditional Healer
- 6: Primary School
- 7: Junior Secondary School
- 8: Senior Secondary School

The remaining digits uniquely identify the facility within the EA and facility type.

GEOGRAPHIC LOCATION CODES

When respondents were asked to provide location information, geographic information was obtained for four levels of administrative units in Indonesia (from largest to smallest): province, kabupaten (regency/municipality.), kecamatan (subdistrict), and desa (village). Names of these administrative units were collected. At the time of data entry, geographic codes were assigned to a subset of geographic location variables, mainly those in the migration sections, for which codes were noted in the questionnaire. Other locations, such as school and facility locations, were not assigned geographic codes. To ensure the privacy of IFLS-HH and IFLS-CF respondents, desa names and codes were dropped. In the migration history, the desa codes were replaced with a scrambled version to facilitate analyses of moves. Appendix C contains the province codes used and Appendix D lists the kabupaten codes. Kecamatan codes are too numerous to list in the Codebook appendix. Kecamatan code values can be found in the geographic location subfiles of the IFLS-HH database (province and kabupaten subfiles exist as well).

The IFLS-HH database contains three supplementary subfiles which contain listings of geographic codes and the names of the associated administrative unit: province (PROV), kabupaten (KAB), and kecamatan (KEC). If desired, users may try to merge on geographic codes by names of the locations for location variables without numeric codes. However, this is not a simple task for two reasons. First, because geographic codes for a given unit are not unique and neither are place names (with the exception of

province), users must first assign the province code, then the kabupaten code, and then the kecamatan code. Second, and most important, the spellings used in the three subfiles above will often not identically match the spellings used by interviewers. Many abbreviations were used by interviewers and some places have names that can be spelled in various ways; in addition, some interviewers did not know how to spell a given place name so they wrote down something that was phonetically close. The differences in spellings makes matching by names a very difficult process.

PROXY RESPONDENTS

If a selected respondent could not be interviewed, a proxy respondent was selected. If a proxy respondent was not used, the proxy respondent identifier code is the same as the book respondent code. If a proxy respondent was used, the proxy respondent identifier code on the given questionnaire book cover is different from the respondent identifier code on that same cover. Table 1.4 provides a list of the respondent and proxy respondent identifier codes for each book:

Table 1.4
RESPONDENT AND PROXY RESPONDENT IDS BY BOOK

IFLS-HH Book	Respondent ID	Proxy ID
Book I (Household Roster/Characteristics)	RESP1_1	RESP2_1
Book II (Household Economy)	RESP1_2	RESP2_2
Book III (Adult Information)	RESP1_3	RESP2_3
Book IV (EMW Information)	RESP1_4	RESP2_4
Book V (Child Information)	RESP1_5	RESP2_5

NOTE: Book K has no proxy respondent code. The variable PERSON on books III, IV and V is also an identifier code for the book respondent.

Proxy respondents were administered the entire questionnaire for Books I, II and V. If a proxy respondent answered a book III or book IV questionnaire in place of the selected respondent, the proxy respondent was administered only a portion of books III and IV. The following lists were finalized after fieldwork was underway, so there may be some instances where a proxy respondent answered more than the questions/sections listed below.

Proxy respondents for Book III answered only the following:

- Questions DL1-DL7
- Questions TK1-TK27 (if selected respondent worked); TK1-TK8 (if not)
- All of section AW
- Questions KW1-KW3
- Questions BR1-BR16 for female selected respondent age 50+
- All of section SR
- All of section KM
- All of section MA
- All of section RJ
- All of section RN

Proxy respondents for Book IV answered only the following:

- Questions KW1-KW3
- Questions BR1-BR16

TEXT RESPONSES

All text responses in the IFLS-HH were entered in Bahasa Indonesia. There are two basic types of text responses: "other, specify" and open-ended questions. All "other, specify" responses have English counterparts. The voluminous and non-standard nature of respondents' answers to open-ended questions made it difficult to incorporate complete translations and code assignments for those variables. At the time of this Codebook, some open-ended responses have not yet been translated into English.

In order to make the IFLS-HH data available to the public in a timely fashion with the resources available, priorities were established in translating the large number of text responses. The "other, specify" were given the highest priority as they outnumbered the open-ended questions. Among open-ended questions, efforts were directed to items viewed as most needed by the general research community. All Bahasa Indonesia text responses are in the data, so users can perform translations themselves if desired. As we get additional items either translated or categorized, those translations and/or categorizations will be made publicly available.

"Other, Specify"

Many questions in the IFLS-HH allowed respondents to give an answer other than one of those listed in the response categories. These are the "other, specify" responses. English translations for these responses are found in the subfile called OTHTRANS. All "other, specify" responses are found in this file. A record represents a given respondent's answer to a

given question. The variables CASE and PERSON define the respondent, the variable OTHQUES denotes the question, and the variable NEWRESP contains the English translation.

No additional codes were added to the response categories in the IFLS-HH database on these "other, specify" responses, with the exception of the level of schooling variables where kindergarten was added. Users must consult the OTHTRANS subfile to see what responses make up the "other, specify" category for a given variable.

Cursory examination of these responses has shown that in a number of cases, many of the "other, specify" responses actually fit within the specified categories provided in the questionnaire. Interviewers, when in doubt, seemed to go for detail rather than generality. In the case of education levels, a research project at RAND made the assignments of "other, specify" responses in to the existing education level categories. These assignments were passed on and incorporated in to the IFLS-HH public use data.

Open-ended Questions

Below we discuss several key areas where open-ended responses occurred and the current nature of translated responses for those areas.

Employment Questions

The employment text variables (where respondent works, what is made where respondent works, and what duties the respondent has) found in Book III have not been directly translated in to English. However, we have attempted to categorize these responses into 2-digit occupation categories using the International Standard Classification list of occupation codes. Respondents often gave less detailed information than needed to accurately assign them to an occupation category.

Table 1.5 lists the employment variables with Bahasa Indonesia responses and their occupation code variable counterpart.

Table 1.5
EMPLOYMENT TEXT VARIABLES AND OCCUPATION CODE VARIABLES
BY BOOK III SUBFILE

Book III Subfile	Employment variables (Bahasa Indonesia)			Occupation Variable (2-digit)
	Where work	What make	What duties	
BUK3TK1	TK10	TK11	TK12	OCC12
BUK3TK2	TK18A	TK19A	TK20A	OCC20A
BUK3T2B	TK18B	TK19B	TK20B	OCC20B
BUK3TK3	TK30	TK31	TK32	OCC32
BUK3TK4	TK40	TK41	TK42	OCC42
BUK3TK5	TK50	TK51	TK52	OCC52

No translations have yet been done on the type of product made at the respondent's workplace. For privacy protection, the place of work has been dropped. This latter variable often was not very informative since respondents gave their boss's name or a company name as opposed to a type of company or place.

Health Insurance Questions

In section AK of Book II, respondents were asked about health insurance coverage. If covered by a specific insurance plan or company, the respondent gave the name of that plan or company. To assist users, a categorical text variable was subsequently created that classifies the type of plan/company for those who have employer-provided health insurance (questions AK01-AK11) into government and non-government plans. Those with individual policies (questions AK12-AK21) are basically covered by private insurance since in 1993 the public health insurance organizations were not yet issuing individual policies. For those who get medical benefits and allowances (questions AK22-AK32), the text listing the medical benefit/allowance (AK27) has not be translated from Bahasa Indonesia. For those who get medical care at a company's clinic or a company-appointed clinic, no categorical variable denoting the type of clinic based on the

name of the clinic has been created. Users must be familiar with the health insurance options within Indonesia in order to best utilize this information beyond the basic gross categories of employer-provided, individual policy, medical benefit/allowance, and company clinic.

Interviewer Comments

At the end of each IFLS-HH questionnaire book, space was provided for interviewers to write down various comments. These have not been translated into English.

Anthropometrist notes have also not been translated.

CODES FOR NOT APPLICABLE, REFUSED, DON'T KNOW, AND MISSING RESPONSES

Rather than repeat the specific values of the codes for non-response used in the IFLS-HH throughout this Codebook, we present those specific values here in Table 1.6.

Table 1.6

CODES FOR NOT APPLICABLE, REFUSED, DON'T KNOW, AND MISSING RESPONSES

Type of Variable	Out of Range	Not Applicable	Refused	Do Not Know	Missing
Numeric Variables by Number of Digits					
1	5	6	7	8	9
2	95	96	97	98	99
3	995	996	997	998	999
4	9995	9996	9997	9998	9999
5	99995	99996	99997	99998	99999
(and so on)					
Decimal Variables (Width.# decimals)					
5.2	99.95	99.96	99.97	99.98	99.99
6.2	999.95	999.96	999.97	999.98	999.99
(and so on)					
Character Variables (any length)	V	W	X	Y	Z

RUPIAH AMOUNTS

Rupiah amounts were originally entered in one of two ways depending on how the amounts were recorded on the questionnaire: character strings using "." and "_", and a 3-part format where millions, thousands and units were

entered separately. These methods were designed to help both interviewers and data entry staff be sure they entered all the necessary digits (many rupiah amounts are in the hundreds of thousands and millions). In the public use data, we have converted these character and 3-part rupiah variables into numeric counterparts. Only the numeric counterparts are included in the IFLS-HH public use data in order to reduce file size and potential user confusion.

Users should be aware that some rupiah amount fields only called for the interviewer to record the full amount rounded up to thousands of rupiah. This procedure was used in the education expenses questions in the education sections of Book III and Book V, in the employment section and the non-coresident transfers section of Book III. An earlier analysis of the education expenses data suggested that once in a while interviewers wrote down the full amount as opposed to rounding it to thousands of rupiah. Such interviewer error may appear in the other sections using thousands of rupiah, so users should be alert to outliers.

PRIVACY PROTECTIONS

In order to protect the privacy of our IFLS-HH and IFLS-CF respondents, the public use data has been stripped of names of individuals and institutions and of any specific addresses (i.e., building names, street names, postal codes). In some cases, blank fields for such variables exist in the public use data (for example, school names in the education sections). In addition, desa information has been dropped; scrambled desa codes are used in the migration histories to facilitate migration analyses. Such fields are noted in the appropriate subfile section of the IFLS-HH Codebook.

DATA CLEANING AND SUGGESTED FIXES TO IFLS-HH SUBFILES

In order to make the IFLS-HH data available in a timely fashion, the enormous size and complexity of the IFLS-HH data encouraged us to create a priority list for data cleaning. Priorities were given to the cleaning of identifier variables--respondent identifiers, anthropometry roster identifiers, household members mentioned elsewhere in the IFLS-HH besides the household roster, and the non-coresident sibling and children rosters. Efforts also focused on cleaning the household roster data so that it could

serve as the main source of basic demographic information on household members. Users, then, could take information from the household roster and use it throughout to provide consistency in characteristics.

Examination of specific sections of the IFLS-HH questionnaire was left to research projects within the Center for the Study of Families in Economic Development which obtained original funding for the IFLS. Sections which have received more detailed examination are the household consumption and income sections, pregnancy and infant feeding sections, the education and health care sections, the non-coresident roster and transfers sections, and the anthropometry section.

Following a policy of not "over-cleaning" data, only those changes for which we had solid information on the correct value were incorporated in to the IFLS-HH data. Numerous other suggested changes are available in a set of "fixes" files which contain SAS programming statements to fix variable values that we believe are in error with our best assessment at the correct response. The IFLS-HH Codebook identifies those variables which have suggested fixes available. Users are welcome to incorporate these corrections in their data if they so choose.

Table 1.7 lists the set of "fixes" files. The file names reflect the subfile to which the fixes are to be applied. For SAS users, the simplest way of incorporating these fixes is to use the %INCLUDE statement to bring the SAS code directly into the program that reads a given subfile.

SAMPLING WEIGHTS

The IFLS-HH was designed to provide detailed information for a small sample of the Indonesian population that could be used for detailed analytical studies. While the IFLS-HH was not designed to produce national estimates of household behavior, the public use file includes a set of weights so that users can adjust for IFLS-HH sampling procedures in their analyses. The IFLS-HH database has two basic sets of weights: household weights and individual weights.

Household weights

The household weights, found in the subfile called HHLDWT, are designed to correct for the over-sampling of urban EAs and EAs in smaller provinces that was done to facilitate urban-rural and Javanese-non-Javanese

comparisons. The household weights are such that when used, the sample of IFLS-HH households will reflect the 1993

Table 1.7

PROGRAM FILES WITH DATA CORRECTIONS TO IFLS-HH SUBFILES

Program files with additional fixes:	Program files with additional fixes:
Book K files	Book IV files
bukkar2.fix	buk4ch1.fix
bukkar3.fix	buk4ch2.fix
Book I files	buk4cx1.fix
buklks3a.fix	buk4kw2.fix
buklks3b.fix	buk4siv.fix
buklkr1.fix	Book V files
buklpp1.fix	buk5dl2.fix
Book II files	buk5se5.fix
buk2ph1.fix	Book CA files
buk2ph2.fix	bukcca2.fix
buk2ut2.fix	bukcca3.fix
buk2nt2.fix	
buk2ge1.fix	
buk2hr1.fix	
Book III files	Book III files (continued)
buk3ba2.fix	buk3kw2.fix
buk3ba3.fix	buk3mg3.fix
buk3ba5.fix	buk3tf3.fix
buk3ba6.fix	buk3tf4.fix
buk3ba7a.fix	buk3tk1.fix
buk3ba8.fix	buk3tk2.fix
buk3baa.fix	buk3t2b.fix
buk3dl1.fix	buk3tk5.fix
buk3dl3.fix	buk3s3a.fix
buk3hi1.fix	

Note: Program files contain SAS® statements.

distribution of households by urban and rural status within each of and between the 13 Indonesian provinces covered by the IFLS-HH. The 1993 distribution of households by province and urban/rural status were generated from 1993 projected population counts provided by BPS and from average household sizes computed from the 1993 SUSENAS (National Socioeconomic Survey). BPS projected population counts were divided by average household sizes to get an estimate of the number of households in 1993 in each province/urban/rural strata.

Two household weights were generated: HHWT730 for the Book K sample of 7,730 households and HHWT224 for the Book I sample of 7,224 households (i.e., households in the household roster data). These weight variables can be merged on to any IFLS-HH subfile by using CASE, the household identifier variable.

Individual weights

There are three basic types of individual weights found in the subfile called INDIVWT: respondent weights, roster weights, and anthropometry weights. Each of these will be discussed in turn below.

Respondent weights

The respondent weights are designed to adjust for the intra-household sampling scheme used to select respondents. From the household roster, the number of household members eligible to be a Book III, Book IV or Book V respondent within each household was determined based on the intra-household sampling rules (see the *Overview and Field Report* for details). Sampling probabilities were then computed for individuals in each of four sampling groups:

- 1) household head/spouse of household head
- 2) non-head/spouse of head age 50 or over and spouse of such person
- 3) non-head/spouse of head age 15-49 and spouse of such person
- 4) child of head/spouse of head age 0-14 (includes fostered children)

The third group was only sought in one out of every four households, so individuals in that group had only a .25 probability of selection in addition to their probability of selection within that group. In addition, a household could only have a maximum of four Book III respondents, so in households where a group 3) respondent was sought, if there were five or six selected respondents, only four were actually interviewed. Because only 13 households had more than four selected respondents, no additional adjustment was made to the weights for such cases.

The computed sampling probability for the individual respondent was then inverted to create a respondent weight for that person. Only Book III/IV/V respondents who were *eligible* respondents as well were given a respondent weight. Book III/IV/V respondents who were incorrectly chosen by interviewers were given a respondent weight of zero. Examples of such

"ineligible" respondents are children age 0-14 who are not biological or adopted children of the household head and spouse but who have a parent in the household, and non-head/spouse age 15-49 who are not in households selected to draw respondents from group 3) above.

The respondent weight (i.e., the inverted sampling probability) was then normalized within each of the sampling groups above. This normalized weight would then sum to the number of eligible respondents within the respondent's sampling group across the 7,224 households where a Book I was done.

The normalized respondent weight was then capped at a value of 3 (99% had a weight of 3 or less) to adjust for outliers---individuals with tiny probabilities of selection and thus given very large weights who could distort weighted tabulations.

There are three versions of the respondent weight provided in the subfile called INDIVWT. These three versions are provided to give the users flexibility in how they wish to use respondent weights. The three respondent weight variables are:

RESPWT:	Capped, normalized respondent weight * HHWT224 (i.e., includes adjustment for household sampling weight)
N_RESPWT:	Capped, normalized respondent weight
ORIGNRWT:	Uncapped, normalized respondent weight

The respondent weight variable, RESPWT, is the one users will generally use when doing weighted tabulations.

Roster weights

The roster weights are designed to make the age and sex distribution of individuals in the IFLS-HH roster data reflect the 1993 population age and sex distribution by urban and rural strata within the 13 provinces covered by the survey. Five-year age groupings were used with those age 75 and older being treated as one group. The population distribution data came from the 1993 SUSENAS. Weights were assigned to all roster members. If the individual's age was missing, an age group for the individual was imputed. The imputation involved examining the age of the individual's spouse and children; if the individual was a Book III/IV/V respondent, dates and ages

provided in those questionnaires were examined. The imputed age variable and a imputation flag are included in the INDIVWT subfile.

The roster weight variable, ROSTERWT, is the ratio of the 1993 SUSENAS population proportion to the IFLS-HH household roster proportion for the given province/urban-rural/sex/age group strata into which the individual falls.

Anthropometry weights

The anthropometry weights are designed to reflect the intra-household sampling scheme to select those to be measured. All Book III/IV/V respondents and any additional children under age 6 living in the household were to be measured. Book III/IV/V respondents were given an anthropometry weight equal to their respondent weight (unnormed and uncapped), other children under age 6 were given the household weight (HHWT224 for the 7,224 household sample), and those individuals who should not have been measured were given an anthropometry weight of zero.

This anthropometry weight was then normalized to sum to the number of those across all households who were eligible to be measured. Not all of those eligible were able to be measured--i.e., there are some Book III/IV/V respondents with no anthropometric measurements, as well as some children under age 6.

As with the respondent weight, the anthropometry weight was also capped at 3 to control for those with tiny probabilities of selection.

There are three versions of the anthropometry weight provided in the subfile called INDIVWT. These three versions are provided to give the users flexibility in how they wish to use anthropometry weights. The three anthropometry weight variables are:

CA_WT:	Capped, normalized anthropometry weight * HHWT224 (includes adjustment for household sampling weight)
N_CA_WT:	Capped, normalized anthropometry weight
ORIGCAWT:	Uncapped, normalized anthropometry weight

The anthropometry weight variable, CA_WT, is the one users will generally use when doing weighted tabulations.

EAS COMPLETED BEFORE SECOND ROUND RETRAINING

Early in the second round of field work, field observation of the household survey interviewers concluded that the proficiency of questionnaire administration for the household survey was lower compared with Round 1. (See the *Overview and Field Report* for more details). Fieldwork had been done in 32 enumeration areas before second round interviewing was halted and interviewers retrained.

The IFLS-HH database includes a subfile with 32 records that contains the EA identifiers for those enumeration areas. The subfile, RND2_EAS, contains the enumeration area identifier variable, EA, and the interview dates when fieldwork was halted before retraining (variables EAMTH and EADAY). Households in these 32 EAs with interview dates earlier than the EAMTH and EADAY values in the file may have interviews with more interviewer error problems than interviews conducted after those dates. Interviews conducted after those dates in the EA took place after the retraining. Users can merge on EAMTH and EADAY to any subfile by using EA. EA can be created in any subfile by using digits 5-7 of the variable CASE. For a listing of those 32 EAs, please see Appendix B.

Table 1.8
IFLS-HH SUBFILES AND SAMPLE SIZES

File Name	File Description (Question numbers found in file)	Unit of Observation	Record Id^a	# of Records
BOOK CA				
BUKCCA1	Cover Page- Anthropometry	household		7164
BUKCCA2	Anthropometry Roster (CA01-CA09)	person	LINE_CA	24515
BUKCCA3	Anthropometry Measurements (CA10-CA15)	person	LINE_CA	24515
BOOK K				
BUKKSC1	Visit Record (IK1-IK2, SC01-SC20)	household		7730
BUKKPSK	Respondent Selection	household		33157
BUKKFP1	Survey Disposition (FP1-FP9)	household		7730
BOOK I				
BUKKAR1	Book I Cover Page	household		7224
BUKKAR2	HH Roster (AR01-AR14)	household member	AR001A	33081
BUKKAR3	HH Roster, education (AR15-AR22)	household member	AR001A	33081
BUK1KR1	Household Characteristics (KR02-KR22)	household		7216
BUK1KS1	Food Item Consumption (KS01-KS02)	food item	ITEM	99900
BUK1KSA	Food Item Own Production (KS03)	food item	ITEM	8305
BUK1KS2A	Transfer of Food Items (KS04-KS05)	household		7216
BUK1KS2B	Non-food Items in Past Month (KS06-KS07)	nonfood item	NF_ITEM1	50512
BUK1KS3A	Non-food Items in Past Year (KS08-KS09)	nonfood item	NF_ITEM2	43296
BUK1KS3B	Education Expenditures (KS10-KS12)	household		7216
BUK1KS4	Purchases in Last Month (KS13-KS15)	type of purchase	ITEM_M	108240
BUK1PP1	Health Provider Knowledge (PP1-PP7)	facility	FACTYPE	50512
BUK1CP1	Interview Evaluation (CPI1-CPI3)	household		7216

^a Used in addition to CASE to uniquely define a record.

Table 1.8 (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id ^a	# of Records
BOOK II				
BUK2UT1	Book II Cover, Farm Business (UT01-UT09)	household		7185
BUK2UT2	Farm Business Assets (UT10-UT14)	asset	FARMASST	25272
BUK2NT1	Non-Farm Business (NT01-NT09)	household		7185
BUK2NT2	Non-Farm Business Assets (NT10-NT14)	asset	NFRMASST	22005
BUK2PH1	Individual Income (PH01-PH07)	household member	PH00	10211
BUK2PH2	Other Income Sources (PH08-PH11)	income source	INCSOURC	43110
BUK2HR1	Household Assets (HR01-HR12)	asset	ASSETTYP	79035
BUK2HR2	Household Data (HR13-HR15)	household		7185
BUK2GE1	Economic Hardships (GE01-GE04)	type of hardship	TYPEHARD	43110
BUK2AK1	Employer-provided Health Insurance: Who is covered (AK01-AK05)	person covered	AK04	9072
BUK2AK2	Employer-provided Health Insurance: Companies (AK06-AK11)	company	AK06	1000
BUK2AK3	Individual Health Insurance: Who is covered (AK12-AK16)	person covered	AK15	7256
BUK2AK4	Individual Health Insurance: Companies (AK17-AK21)	company	AK17	77
BUK2AK5	Employer-provided Medical Benefits: Who is covered (AK22-AK26)	person covered	AK25	7863
BUK2AK6	Employer-provided Medical Benefits: Type of Benefit (AK27-AK32)	benefit	AK27	524
BUK2AK7	Care at Company Clinic: Who is covered (AK33-AK37)	person covered	AK36	7991
BUK2AK8	Care at Company Clinic: Clinics (AK38-AK43)	clinic	AK38	524
BUK2CP2	Interview Evaluation (CP111-CP113)	household		7184

^a Used in addition to CASE to uniquely define a record. Number of records vary due to unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Table 1.8 (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id ^b	# of Records
BOOK III				
BUK3S3A	Book III Cover Page	respondent		14418
BUK3DL1	Education Summary (DL01-DL07)	respondent		14418
BUK3DL2	Levels of Schooling (DL08-DL15)	level of school	DL08	13355
BUK3DL3	Last School Attended (DL17-DL31)	respondent		7778
BUK3TK1	Work Experience (TK01-TK17)	respondent		14418
BUK3TK2	Current Job (Primary) (TK18-TK27)	monthly/yearly	PERIOD	19524
BUK3T2B	Current Job (Secondary) (TK18-TK27)	monthly/yearly	PERIOD	
BUK3TK3	Primary Jobs During Past 5 Years (TK28-TK37)	year (92-88)	YEAR	52934
BUK3TK4	Secondary Jobs During Past 5 Years (TK38-TK46)	year (92-88)	YEAR	52934
BUK3TK5	Employment History: 1st, '83, '73 job (TK47-TK59)	job time point	JOBREC	19551
BUK3AW1	Time Allocation (AW1-AW3)	respondent		14418
BUK3KW1	Marriage Summary (KW01-KW08)	respondent		14418
BUK3KW2	Marital History (KW09-KW22)	marriage	MARRNUM, ENTRYORD	10516
BUK3KW3	Husband's Desire For More Kids (KW23-KW27)	married male		7572
BUK3BR1	Pregnancy Summary (BR01-BR16)	non-Book 4 female		2664
BUK3MG1	Migration History: Birth, Age 12, 1st Marr (MG01-MG21)	respondent		12996
BUK3MG2	Migration History (MG21-MG40)	move	MOVENUM	15376
BUK3MG3	Last Move to Current Residence (MG41-MG52)	respondent		6945
BUK3SR1	Circular Migration: Trigger Questions (SR01-SR02)	respondent		14418
BUK3SR2	Circular Migration History (SR03-SR18)	circular move	TRIPNUM	1788
BUK3KM1	Smoking Habits (KM01-KM09)	respondent		14406
BUK3KK1	Current Health & Activities of Daily Living (KK01-KK04)	respondent		12985
BUK3KK2	Health Conditions During Past 4 Weeks (KK05-KK07)	respondent		12985
BUK3MA1	Acute Morbidity (MA01-MA03)	symptom	SYMPTOM	172872
BUK3PS1	Self-Treatment (PS01-PS02)	respondent		12989
BUK3RJ1	Outpatient Medical Facilities (RJ01-RJ04)	respondent		14406
BUK3RJ2	Outpatient Visits (RJ05-RJ24)	visit	RJ05	4685

^b Used in addition to CASE and PERSON to uniquely define a record. Number of records vary due to unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Table 1.8 (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id ^b	# of Records
BUK3RN1	Inpatient Medical Facilities (RN01-RN04)	respondent		14406
BUK3RN2	Inpatient Visits (RN05-RN20)	visit	RN05	410
BUK3BA1	NonHouseholders: Parents (BA01-BA07)	parent	PARENT	25970
BUK3BA2	Parent Characteristics (BA08-BA26)	parent	PARENT	26169
BUK3BA3	Non-Householders: Siblings (BA28-BA29)	respondent		12459
BUK3BA4	Sibling Roster [All Possible] (BA30A-BA30G, BA33B)	sibling	BA30A	44995
BUK3BA5	Sibling Summary Counts (BA30H, BA31-BA32)	respondent		11241
BUK3BA6	Sibling Characteristics (BA34-BA51)	selected sibling	BA34	37339
BUK3BA7A	Help from/to Siblings Not Mentioned (BA53-BA57)	respondent		12459
BUK3BA7B	Trigger Questions: Non-coresident children BA58-BA63)	respondent		12459
BUK3BA8	Non-Householders: Children Roster (BA63A-BA71)	child	BA63A	10100
BUK3BA9	Non-Resid Kids Age 15+-Counts (BA72-BA73)	respondent		3776
BUK3BAA	Non-Coresident Children Characteristics (BA74-BA90)	selected child	BA74	8246
BUK3BAB	Help From/To Non-cores. Kids Not Selected (BA92-BA96)	respondent		3432
BUK3TF1	Transfers To: Parents,Sibs,Children (TF01-TF03)	type of receiver	RELTYPE	38955
BUK3TF2	Transfers From: Parents,Sibs,Children (TF04-TF05)	type of giver	RELTYPE	2007
BUK3TF3	Transfers To: Rel.,Friends,Employer,Org. (TF06-TF07)	type of receiver	WHOHELP	51940
BUK3TF4	Transfers From: Rel.,Friends,Employer,Org. (TF08-TF09)	type of giver	WHOHELP	51940
BUK3HI1	Individual Assets (HI01-HI13)	asset type	ASSETTYP	73993
BUK3HI2	Other Income In Past Year (HI14)	type of income	INCOMTYP	30570
BUK3CP3	Interview Evaluation (CP1-CP3,CPNOTE)			14405

^b Used in addition to CASE and PERSON to uniquely define a record. Number of records vary due to unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Table 1.8 (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id ^b	# of Records
BOOK IV				
BUK4SIV	Book IV Cover	respondent		4981
BUK4KW1	Marriage Summary (KW01-KW03)	respondent		4981
BUK4KW2	Marital History (KW04-KW17)	marriage	MARRNUM, ENTRYORD	5776
BUK4KW3	Menstrual History/Desire for Kids (KW18-KW25)	respondent		4889
BUK4BR1	Pregnancy Summary (BR1-BR16)	respondent		4980
BUK4CH1	Pregnancy Roster (CH01-CH11)	pregnancy	PRGID	17962
BUK4CH2	Information on Pregnancies (CH12-CH40)	pregnancy	PRGID	17770
BUK4CH3	Children After 7/90 (CH42-CH54)	kid after 7/90	CH42	1947
BUK4CX1	Contraceptive Knowledge/Use: Efficient (CX1-CX11)	type of method	METHOD	39120
BUK4CX2	Contraceptive Knowledge/Use: Less efficient (CX12-CX13)	type of method	METHOD	24450
BUK4CX3	Contraceptive Use (CX14-CX29)	person		4890
BUK4KL1	Cont. Calender: Marriage/calendar start/end dates	marriage		4966
BUK4KL2	Cont. Calendar (Jan 84-Dec 88) (KAL)	type of event	EVENT	29340
BUK4KL3	Cont. Calendar (Jan 89-Feb 94) (KAL1)	type of event	EVENT	44010
BUK4KL4A	Cont. Facility Visit Cost from last 30 months of calendar--Sep 93-Feb 94 (KL4,COSTRP)	month	EVNTMTH	16353
BUK4KL4B	Cont. Facility Info (KL1-KL6)	respondent		4890
BUK4CP4	Interview Evaluation (CP1-CP3, CPNOTE4)	respondent		4981
BOOK V				
BUK5SE5	Book V Cover Page	respondent		7751
BUK5DL1	Levels of Schooling (DLA01-DLA22)	level attended	DLA05	8562
BUK5DL2	School Expenses (DLA24-DLA25)	respondent		4668
BUK5MA1	Acute Morbidity (MAA01-MAA03)	symptom	SYMPTOM	116265
BUK5PSA	Self-Treatment (PSA01-PSA02)	respondent		7751
BUK5RJ1	Outpatient Medical Facilities (RJA01-RJA04)	respondent		7751
BUK5RJ2	Outpatient Visits (RJA05-RJA24)	visit	RJA05	2259
BUK5RN1	Inpatient Medical Facilities (RNA01-RNA04)	respondent		7751
BUK5RN2	Inpatient Visits (RNA05-RNA20)	visit	RNA05	100
BUK5CP5	Interview Evaluation (CP5A-CP5C,CPNOTE5)	respondent		7751

^b Used in addition to CASE and PERSON to uniquely define a record. Number of records vary due to unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Table 1.8 (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id	# of Records
OTHTRANS				
OTHTRANS	English Translations of IFLS-HH "Other, specify"	respondent /question number	CASE, PERSON, OTHQUES	90603
WEIGHTS				
HHLWT	Household Sampling Weights	household	CASE	7730
INDVWT	Individual Sampling Weights	household member	CASE, PERSON	33081
LOCATION				
PROV	Province Codes and Names	province	KODE	26
KAB	Kabupaten Codes and Names (Regency/municipality)	kabupaten	KODE	297
KEC	Kecamatan Codes and Names (Subdistrict)	kecamatan	KODE	3778
RND2_EAS				
RND2_EAS	EAs Done Before 2nd Round Retraining	EA	EA	32