PREFACE

The 1993 Indonesian Family Life Survey (IFLS) provides data at the individual and family level on fertility, health, education, migration, and employment. Extensive community and facility data accompany the household data. The survey was a collaborative effort of *Lembaga Demografi* of the University of Indonesia and RAND, with support from the National Institute of Child Health and Human Development, USAID, Ford Foundation, and the World Health Organization. In Indonesia, the 1993 IFLS is also referred to as SAKERTI 93 (Survai Aspek Kehidupan Rumah Tangga Indonesia).

The IFLS covers a sample of 7,224 households spread across 13 provinces on the islands of Java, Sumatra, Bali, West Nusa Tenggara, Kalimantan, and Sulawesi. Together these provinces encompass approximately 83 percent of the Indonesian population and much of its heterogeneity. The survey brings an interdisciplinary perspective to four broad topic areas:

- fertility, family planning, and contraception
- infant and child health and survival
- education, migration and employment
- the social, economic, and health status of adults, young and old

Additionally, extensive community and facility data accompany the household data. Village leaders and heads of the village women's group provided information in each of the 321 enumeration areas from which households were drawn, and data were collected from 6,385 schools and health facilities serving community residents.

This current document contains the User's Guide to the IFLS database. The User's Guide describes the IFLS data structure and presents detailed descriptions of the variety of information available, and how it can be put together. This Guide is meant to be a companion to the IFLS Codebooks and provides guidelines on how to build analysis files from the data.

The complete public use file documentation consists of the following publications:

DRU-1195/1-NICHD/AID, *The 1993 Indonesian Family Life Survey: Overview and Field Report.* This report describes the purpose, design, field work and response rates for both the household and the facility components of the IFLS.

DRU-1195/2-NICHD/AID, *The 1993 Indonesian Family Life Survey: Appendix A, Household Questionnaires and Interviewer Manual.* This document provides the English translation of the questionnaires used during the household and individual interviews, as well as the interviewers' instructions.

SUMMARY

The data from the 1993 Indonesian Family Life Survey (IFLS) provide a rich but complex database. This User's Guide describes the IFLS data structure and presents detailed descriptions of the variety of information available, and how it can be put together. This Guide is meant to be a companion to the IFLS Codebooks and provides guidelines on how to build analysis files from the data. For example, the document discusses how to identify various individuals of interest (e.g. husbands, wives, children, parents of respondents) and how to link data from different parts of a particular person's questionnaire with one another and with data from the questionnaires of related individuals. The multiple file structure of the IFLS makes linking files the major task in building analysis files.

DRU-1195/3-NICHD/AID, *The 1993 Indonesian Family Life Survey: Appendix B, Community-Facility Questionnaires and Interviewer Manual.* This document provides the the English translation of the questionnaires used during the community leader and facility interviews, as well as the interviewers' instructions.

DRU-1195/4-NICHD/AID, *The 1993 Indonesian Family Life Survey: Appendix C, Household Codebook*. The codebook provides descriptions of all variables and their and locations in the IFLS data files. The codebook also presents information on cases that are known anomalies.

DRU-1195/5-NICHD/AID, *The 1993 Indonesian Family Life Survey: Appendix D, Community-Facility Codebook*. The codebook provides descriptions of all variables and their and locations in the IFLS data files. The codebook also presents information on cases that are known anomalies.

DRU-1195/6-NICHD/AID, *The 1993 Indonesian Family Life Survey: Appendix E, Users' Guide.* This report document provides descriptions of the IFLS data format and data files. Guidelines for data use are provided, with special emphasis on working with the household, individual, and facility IDs and making links across different parts of the survey.

Subsequent announcements regarding the 1993 IFLS database can be found on the Family Life Surveys Home Page under the Labor and Population Program Home Page on RAND's external World Wide Web server: *http://www.rand.org*. Users should check this page periodically for updated information on the 1993 IFLS. The FLS Home Page also contains information on the First and Second Malaysian Family Life Surveys conducted in 1976-1977 and in 1988-1989.

Subsequent information about the 1993 IFLS will also appear in the *FLS Newsletter*. The *FLS Newsletter*, meant to serve all FLS data users and interested parties, is a free occasional publication of the RAND Center for the Study of the Family in Economic Development and the RAND Population Research Center. The newsletter disseminates information about surveys fielded by RAND in developing countries. Those who are not already subscribers may either subscribe via the FLS Home Page above or by sending electronic mail to the FLS subscription alias *fls-sub@rand.org*.

ACKNOWLEDGMENTS

The 1993 Indonesian Family Life Survey (IFLS) provides data at the individual and family level on fertility, health, education, migration, and employment. Extensive community and facility data accompany the household data. The survey was a collaborative effort of *Lembaga Demografi* (LD) of the University of Indonesia and RAND, with support from the National Institute of Child Health and Human Development, USAID, Ford Foundation, and the World Health Organization.

The IFLS questionnaire development, fieldwork, and public use file creation represent the combined efforts of social scientists in Indonesia and the United States. At RAND, Paul Gertler served as Principal Investigator, with Elizabeth Frankenberg and Lynn Karoly as co-Principal Investigators. Sulistinah Achmad was the LD Project Director, with co-Directors I.G.N. Agung and Sri Harijati Hatmadji. In addition, Susan Butler, Theo Downes-LeGuin, Christine Peterson, Polly Phipps and Paramita Sudharto were instrumental to the survey effort.

A number of other researchers at RAND contributed to the design of the questionnaire and implementation of the survey. In particular, Julie DaVanzo, Reta Hendratidewi, Jacob Klerman, Lee Lillard, Jack Molyneaux, Bob Schoeni, James P. Smith, and Duncan Thomas contributed tremendously to questionnaire development. John Adams and Dan McCaffrey developed procedures for sampling and for constructing weights. RAND staff Kim Linton, Nick Murray, Caron Murray, Joanna Nelsen, Judy Perlman, and Carl Serrato played key roles in the logistics of questionnaire production and editing, fieldwork, and data entry. We would also like to acknowledge the input of Angus Deaton (Princeton University), T. Paul Schultz (Yale University), and John Strauss (Michigan State University).

The survey could not have taken place without the efforts of the directors and staff at LD, particularly N. Haidy A. Pasay and Djuhari Wirakartakusumah, the current and former LD directors, respectively, as well as Aris Ananta and Sri Moertiningsih Adioetomo. Eko Ganiarto, Akhir Matua Harahap, Edy Priyono, Gatot Arya Putra, Sapruddin, Sutji Rochani, and M. Yusuf all participated in questionnaire design, training, and fieldwork for the household survey. Nargis, Darlis Rabai, and Muda Saputra participated in questionnaire design, training, and fieldwork for the community and facility survey. Gita Marina Sapprudin managed the technical production of the instrument, while Ni Wayan Suriastini, Linda Fitriawati, Sugiharso, and Hendratno took responsibility for data entry.

I. INTRODUCTION

The database for the 1993 Indonesian Family Life Survey (IFLS) is a rich but complex database. The User's Guide describes the IFLS multi-file structure for the household survey data (IFLS-HH) and for the community-facility survey data (IFLS-CF), and presents fairly detailed descriptions about the variety of information available, including how to put various types of data together. While the IFLS-HH Codebook[1] and the IFLS-CF Codebook[2] each explain what individual variables mean and note special cases within their respective databases, the User's Guide provides the guidelines on how to use those data more effectively. The User's Guide discusses how to identify various individuals of interest (e.g., husband, wives, children, parents of respondents) and then how to link data from different parts of the respondent's questionnaire and from the questionnaires of related individuals, plus how to link the IFLS-CF data with relevant IFLS-HH subfiles. The multiple file structure of the IFLS database makes linking files the major task in building analysis files, and great effort has been made to cover most linkages.

This User's Guide does not cover in detail such topics as sampling frames and response rates. Users seeking such information should consult the *Overview and Field Report*.[3] The User's Guide discusses data quality issues but does not provide any analysis of data quality. A descriptive report on the population, health and education data in the IFLS is available and should prove useful to users interested in those areas of the data.[4]

Section II provides a short overview of the IFLS survey, including sample definitions, survey instrument descriptions, and data entry and cleaning experiences which may be of interest to users. Section III presents the IFLS data structure and format plus addresses various IFLS data conventions. Section IV discusses how to identify various individuals of interest, while Section V provides detailed information on a variety of ways to link the different IFLS subfiles together, including IFLS-HH and IFLS-CF data. The final section deals with the interrelatedness of the IFLS database and how users might make use of the richness provided therein.

---

[1] *The 1993 Indonesian Family Life Survey: Appendix C, Household Codebook,* DRU-1195/4-NICHD/AID, RAND, 1995.

[2] *The 1993 Indonesian Family Life Survey: Appendix D, Community -Facility Codebook,* DRU-1195/5-NICHD/AID, RAND, 1995.

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

[4] *The Indonesian Family Life Survey: Overview and Descriptive Analysis of the Population, Health, and Education Data,* DRU-1191-AID, RAND, 1995.

The User's Guide has four appendices. Appendix A lists all subfiles for the IFLS-HH and Appendix B lists the IFLS-CF subfiles. Appendix C presents a general description of the organization of the health sector (public and private) in Indonesia. Appendix D presents a brief description of the educational system in Indonesia.

II. GENERAL OVERVIEW OF THE IFLS SURVEY

In this section, we briefly overview the 1993 Indonesian Family Life Survey (IFLS) and describe the survey instruments. More detailed descriptions of the survey and its instruments can be found in the *Overview and Field Report*. The information presented below is intended to serve as an easy reference when using the data and codebook materials. A short synopsis of the data entry and data cleaning process for the IFLS data is also included.

PURPOSE OF IFLS

The 1993 Indonesian Family Life Survey (IFLS) provides data at the individual and family level on fertility, health, education, migration, inter-household transfers, and employment. Extensive community and facility data accompany the household data. The survey was a collaborative effort of *Lembaga Demografi* of the University of Indonesia and RAND, with support from the National Institute of Child Health and Human Development, USAID, the World Bank, Ford Foundation, and the World Health Organization. The IFLS covers a sample of 7,224 households spread across 13 provinces on the islands of Java, Sumatra, Bali, West Nusa Tenggara, Kalimantan, and Sulawesi. Together these provinces encompass approximately 83 percent of the Indonesian population and much of its heterogeneity. The survey brings an interdisciplinary perspective to four broad topic areas:

- fertility, family planning, and contraception
- infant and child health and survival
- education, migration and employment
- the social, economic, and health status of adults, young and old

Additionally, extensive community and facility data accompany the household data. Village leaders and heads of the village women's group provided information in each of the 321 enumeration areas[5] from which households were drawn, and data were collected from 6,385 schools and health facilities serving community residents.

[5]There were 9 enumeration areas (EA) that had the same village head as another IFLS enumeration area. Those 9 EAs did not have the community questionnaires administered because the information had been collected earlier under the counterpart EA. The IFLS-CF Codebook lists the 9 EAs and their interviewed counterparts. Users can make duplicate records for each of the 9 EAs from their counterpart EAs.

SAMPLES COLLECTED IN IFLS

The *Overview and Field Report* provides detailed discussions of the sampling procedures used in the IFLS. Please refer to that document for information on how samples were selected. Below are listed the types of individuals/facilities covered in the IFLS-HH and IFLS-CF databases.

IFLS-HH (Household Survey)

In addition to collecting household-level information (household roster, household characteristics and economy), the IFLS-HH collected detailed data on the following individuals in the household:

- The household head and their spouse;
- Two randomly selected children of the head and spouse aged 0 to 14 (interviewed by proxy);
- An individual age 50 and above and their spouse, randomly selected from remaining members;
- For a randomly selected 25 percent of the households, an individual aged 15 to 49 and their spouse, randomly selected from remaining members.

Only a maximum of four adults were interviewed per household. Households in the 25 percent group above could have more than four possible adult respondents. A selection scheme was used to determine which four adults would be interviewed if the household had five or six possible adult respondents. See the *Overview and Field Report* for details.

IFLS-CF (Community-Facility Survey)

In addition to collecting community-level information from the village leaders and the head of the village women's group, the IFLS-CF collected data on the following types of facilities:

- Government health centers and auxiliary centers (*puskesmas, puskesmas pembantu*);
- Private physicians and clinics (*praktek umum/klinik*);
- Nurses, midwives, and paramedics (*perawat, bidan, mantri, paramedis*);
- Community health posts and Family Planning posts (*posyandu, PPKBD*)
- Traditional practices (*dungun, sinshe, tabib, orang pintar*)
- Primary schools (*SD*)
- Junior secondary schools (*SMP*)
- Senior secondary schools (*SMA)*

The number of facilities per EA targeted for interview varied by type of facility. The target numbers were: four for health centers; three for nurse/midwife/paramedic, primary schools, and junior secondary schools; and two for private doctors and clinics, traditional practices, and senior secondary schools. In some EAs, the targeted number of facilities were not met, so users will observer EAs with less than the maximum number of facilities. In addition, some facilities served multiple EAs. Such facilities were interviewed only once and

they appear under the EA where there interview was conducted. See Appendices C and D for brief descriptions of the health and education facilities.

SURVEY INSTRUMENTS

IFLS-HH (Household Survey)

The IFLS-HH data were collected with seven instruments. Table 2.1 provides a brief description of the instruments and the samples to which they were administered.

Table 2.1

IFLS-HH SURVEY INSTRUMENTS

Survey Instrument	Administered to:	General Contents
K: Control Book	All households where an interview was attempted	Questionnaire tracking, within-household sampling
I: Household Characteristics	All households interviewed in IFLS	Household roster and environment, consumption, health provider knowledge
II: Household Economy	All households interviewed in IFLS	Farm and non-farm business, labor and non-labor income, household assets, economic shocks, health insurance
III: Adult Information	Selected adult respondents age 15 and older	Education, work, migration, and marriage histories; health status, acute morbidity, and health care utilization; non-coresident parents, siblings and children; inter-household transfers; individual assets
IV: Ever-married Female Information	Ever-married female Book III respondents age 15-49	Pregnancy history , infant feeding, contraceptive knowledge and use, contraceptive calendar
V: Child Information	Selected children of Household Head age 0-14	Education history, acute morbidity, health care utilization
CA: Anthropometry	IFLS Book III and V respondents, children 0-5	Height, weight, demographics

IFLS-CF (Community-Facility Survey)

The IFLS-CF data were collected with eleven instruments. Table 2.2 provides a brief description of the instruments and the samples to which they were administered.

Table 2.2

IFLS-CF SURVEY INSTRUMENTS

Survey Instrument	Administered to:	General Contents
I: Community Characteristics	Village Heads of the 321 IFLS enumeration areas (9 EAs had same village head as another EA)	Transportation, electricity, water and sanitation, agriculture and industry, history and climate, migration, credit, history of schools and health facility availability
II: Community Characteristics	Taken from village records in the EA	Village statistics
PKK: Community Characteristics	Head of PKK (Association of Family Activities) or wife of village head	Prices of food stuffs, history of school and health facility availability
PUSK: Government Health Center/ Sub-Health Center	Selected health centers and sub-health centers (*Puskesmas* / *Puskesmas Pembantu*)	Facility management and history, service availability, staff and equipment, family planning, vignettes on types of care
DR: Private Physicians and Clinics	Selected physicians and clinics (*praktek umum* / *klinik*)	Facility management and history, service availability, staff and equipment, family planning, vignettes on types of care
BIDAN: Nurses, Midwives, and Paramedics	Selected practitioners (*perawat*, *bidan*, *mantri*, *paramedis*)	Facility management and history, service availability, staff and equipment, family planning, vignettes on types of care
PPKBD: Community Health Posts and Family Planning Posts	Selected Community Health Posts (*Posyandu*) / Family Planning Post (*PPKBD*)	Facility management and history, service availability, staff and equipment, family planning
PTRAD: Traditional Health Practitioners	Selected practitioners (*dungun*, *sinshe*, *tabib*, *orang pintar*)	Facility management and history, service availability, staff and equipment, family planning
SD: Primary School	Selected Primary Schools	Staff and school characteristics, classroom characteristics, student test scores, revenues
SMP: Junior Secondary School	Selected Junior Secondary Schools	Staff and school characteristics, classroom characteristics, student test scores, revenues
SMA: Senior Secondary School	Selected Senior Secondary Schools	Staff and school characteristics, classroom characteristics, student test scores, revenues

DATA ENTRY

The data was transcribed from the recording forms into the PC-based data entry system ISSA(Integrated System for Survey Analysis)[6] , by staff at *Lembaga Demografi* (LD) of the University of Indonesia. The data entry program was developed by Nick Murray of RAND with the assistance of LD staff. 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. Upon receipt of the IFLS data at RAND, the ISSA ASCII files were converted into SAS® files for use in the data cleaning process and the preparation of a public use file version of the data.

Due to the double entry and data entry program checks, data entry errors were basically nil. The source of remaining data errors was interviewer error and respondent error. Based on problems uncovered so far, there appears to be about a 1-2 percent interviewer/respondent error rate. For files that have, for example, 20,000 records, the 1-2 percent error rate suggests 200-400 records with potential problems. In more complicated sections of the questionnaire, this rate may be a bit higher.

DATA CLEANING

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 data cleaning strategy that would meet basic user needs and make the data available to the research community in a reasonable time frame. Given 100%-verification at data entry, the basic approach, then, was to concentrate on those data cleaning activities which required access to information that was privacy protected. Such cleaning activities could only be done at RAND. 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, the non-coresident sibling and children rosters, and facility identifiers. Within the IFLS-HH, efforts also focused on trying to clean the household roster data so that it could serve as the main source of basic demographic information on household members. Users could then take information from the household roster and use it throughout to provide consistency in characteristics. Additional areas where data checking efforts were made reflect those sections of interest to projects within the P01 grant that included original IFLS funding[7] and those of interest to the report prepared for AID, a

---

[6]Survey 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.

[7]The P01 is the Center for the Study of Families in Economic Development at RAND.

sponsor of the survey. Those areas included anthropometric data, income data, outpatient and inpatient utilization, education status and expenses, pregnancy histories and infant feeding, and interhousehold transfers. 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.

Data problems due to interviewer error were the types of problems on which cleaning efforts focused. Discrepancies or oddities due to respondent confusion or to different respondents referring to the same event were generally not addressed during data cleaning. After public release, subsequent data cleaning efforts sponsored by RAND projects will continue and results of those efforts will be made available to the IFLS user community.

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 and IFLS-CF 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 guess at the correct response. Many of these suggested fixes came from preliminary analyses of selected sections of the IFLS database. The IFLS-HH and IFLS-CF codebooks identify those variables which have suggested fixes available. Users are welcome to incorporate these corrections in their data if they so choose. These "fixes" files are listed in the respective IFLS codebook introductory sections.

OBSERVATIONS ON DATA QUALITY

Below we discuss some of the data problems uncovered. Many of the data problems discussed below have been corrected in the IFLS public use database, and are so noted below. In some cases, however, only suggested corrections are provided via the "fixes" files described above, and are noted accordingly. In other cases, decisions on how to handle a particular problem belong in the hands of the research analyst and in such cases, we alert users to the type of problems we have uncovered, but do not provide suggested fixes. The discussion below may also help users understand remaining interviewer and respondent errors not detected before public release.

Control files in Book K

With regards to the Book K control files, there were some inconsistencies between identifiers and completion codes in the FP and PS sections (files BUKKFP1 and BUKKPSK). In the FP section, sometimes the Book III, IV or V proxy respondent identifier code is listed instead of the individual to whom the book refers; a few cases have codes indicating completed, but the individual has incomplete data. The PS section (subfile BUKKPSK) lists all the household members from which the various book respondents were to be selected. In some cases, the order of household members is different from that in the household roster (section AR of Book I), and thus the person identifier codes in the PS section (PS00) may not coincide with the roster identifier codes. The respondent identifier codes in the FP section (subfile BUKKFP1) were based on the identifier code in the PS section and not that in the household roster, so discrepancies exist between the individual book covers, where person identifier codes were based on the household roster, and the respondent identifier codes in FP. Where possible, corrections for identifier codes have been made to the FP and PS subfiles (BUKKFP1 and BUKKPSK). However, not all such cases may have been corrected. We recommend that users use the respondent identifier variable on the individual book cover file (see Table 4.1 in Section IV of this document for a list of those files) to identify the book respondent and not use the FP data. Nonrespondents to IFLS-HH questionnaire books, however, are only found in the FP data.

Household and Person Identifiers

The IFLS-HH data have been checked for the quality of the two main identifier codes: household identifier (CASE) and person identifiers. Errors were found in both identifier codes found in the cover information for Books III, IV and V. Known problems have been corrected so that the Book III, IV and V respondent identifiers now link properly with the household roster member identifiers.

Errors were also found in the household member identifiers listed in the birth history (question CH27 in Book IV). Most of those errors resulted from interviewers writing in the wrong person number or writing in the “relationship to household head” code. Some of those errors, though, reflect interviewer problems in filling out the birth history wherein the interviewer wrote information about child “i” in the wrong column (i.e., not in the “i”th column). These sources of errors have largely been corrected; however, a few cases where the wrong lines were filled in may still exist.

Errors were found in the household member identifiers recorded in the AK section of Book II (health insurance) and in the PH section of Book II (income). Some were transcription errors, some arose because interviewers entered the relationship to household head code instead of the household

member number from the household roster. Known errors have been corrected in files related to those sections.

Errors were found in the household member identifier (question CA02) in the anthropometry roster. These errors were due largely to transcription errors from the household roster to the anthropometry roster and to entering the relationship to head of household code instead of the household member identifier. Known errors in question CA02 have been corrected.

Errors were found in the transcription of non-coresident roster person identifiers between the full rosters for non-coresident siblings and for non-coresident children and the rosters of selected siblings and children from those rosters on whom detailed information was collected. For non-coresident siblings, names were entered in both the full and selected rosters, so errors were detected by differences in names with the same identifier code (BA30A in the full roster, BA34 in the selected roster) between the two rosters in the same household. For non-coresident children, however, names were only entered for the full roster--the data entry program did not allow the names on the selected children roster to be entered. Detecting errors in BA74, the identifier code in the selected child roster, with respect to the full roster identifier variable, BA63A, was much more difficult. Comparing the age the child now with information on the general age of the child when he/she left home provided clues to possible identifier code problems. In addition, there was a particular selection procedure for children and if the identifiers in the selected roster did not follow that pattern, this too provided clues to possible identifier problems. Known problems have been fixed in the non-coresident sibling and children rosters.

In the household roster, errors in the "birth father/birth mother" identifier codes (AR10/AR11) were found. Among those with the head and spouse listed in the AR10/AR11 fields were people who were older than the head and his spouse, people with relationship to head codes that did not imply "child" of the head/spouse, and people who did not show up in the birth history of the female head/spouse of head. Sometimes a head/spouse was listed as his/her own parent. In some cases the father/mother codes were reversed. In households where the head of the household was a woman, some interviewers still made AR10=1 and AR11=2, the standard response for children in a male head and female spouse household, out of force of habit. In other cases it appears everyone got AR10=1 and AR11=2 regardless of who they were. A set of suggested fixes have been prepared for these cases and a program containing those suggested fixes is available (BUKKAR2.FIX).

Another point about the "birth father/birth mother" codes is the following. For children of the household head who are the product of a different wife than the current spouse, the AR02 code is "3", meaning child of head, but the AR11 mother code is often equal to "2" for such a child as opposed to using a code signifying "not in household" or "dead" for the child's actual birth mother. Crosschecks of children in the household roster who have a Book IV respondent's identifier code as the "birth mother" against the children listed in the Book IV birth history uncovered these cases. Checks of the

father's marriage history revealed earlier marriages with dates that encompassed the children in question. Thus, even though the AR10/AR11 questions specifically refer to "birth parent", responses do not always reflect that fact---sometimes responses appear to reflect who is viewed as the father/mother. Such "step" children are not flagged in the data. Users must crosscheck the roster household member identifier code for the child with household member identifier codes in the birth history and with the marriage date of the woman to her current husband to isolate such children.

Event identifiers

In the outpatient visit histories in Book III, there were errors in the visit identifier code, RJ05A, due to interviewer transcription problems from the outpatient visit summary page. Similar problems were found in the Book V outpatient data (BUK5RJ2, variable RJA05A). Known errors of this type were corrected.

In the inpatient visit histories in Book III, there were errors in the visit identifier code, RN05A, due to interviewer transcription problems from the inpatient visit summary page. Similar problems were found in the Book V inpatient data (BUK5RN2, variable RNA05A). Known errors of this type were corrected.

In the marriage history data, interviewers were supposed to collect information on the current marriage (or most recent marriage is the respondent was not currently married) and the ask about the next most recent marriage and so on down the line to the first marriage. In a few instances, the interviewer collected information on the first marriage first, then the second and so on. Corrections to the variable ENTRYORD (which runs from current to first marriage) can be found in the program files BUK3KW2.FIX and BUK4KW2.FIX.

Users should note that because polygamy exists in Indonesia, the current marriage of a male respondent (i.e., the marriage to the woman he is living with) may have occurred before subsequent marriages to other women. Thus, users will not necessarily see a chronologically descending order of marriage dates. To help users, a chronological order marriage identifier was created. The variable MARRNUM reflects the chronological order from the earliest marriage (MARRNUM=1) to the last marriage.

In the pregnancy history in Book IV, some interviewer confusion existed over what the responses to CH04 and CH05 represented. The pregnancy history has one record for each pregnancy outcome. CH04 is the pregnancy count---it increments each time the woman becomes pregnant anew--and multiple births (twins, triplets) all have the same value of CH04 since they come from one pregnancy episode. CH05 is the count of pregnancy outcomes--it increments for each outcome--and outcomes that are part of multiple births have different CH05 values. For example, if a woman has 3 pregnancies and the second pregnancy resulted in twins, the sequence for CH04 would be 1,2,2,3 and for CH05 it would be 1,2,3,4. Those interviewer errors in assigning CH04 and CH05 values that

have been uncovered were corrected. CH05 is the pregnancy outcome identifier. Additional corrections to CH04 and CH05 that are not based on checks of the hard copy questionnaires, but based on examination of the data, can be found in the program file called BUK4CH1.FIX.

In the pregnancy history in Book IV, births after July 1, 1990 were asked more detailed questions about recent infant feeding practices. Question CH42 recorded the CH05 number for the given birth. Errors were detected in CH42 due to transcription error and interviewer confusion on what identifier went into CH42. Known errors in CH42 were corrected.

Facility Identifiers

In the contraceptive use section of Book IV, respondents were asked about where they got family planning services. The information collected was the name of the facility and the village, kecamatan, kabupaten, and province where the facility was located. In the case of community health posts (*posyandus*), a village often has more than one such facility. Given multiple *posyandus* per village, the location information collected was often insufficient to accurately identify the specific *posyandu* to which the woman was referring. Therefore, it is suggested that users create EA-level averages for *posyandu* characteristics from the Book PPKBD facility data rather than using the specific *posyandu* facility data for the facility code attached to the Book IV record. This topic is discussed in more detail in the *Overview and Field Report*.

Location codes (Book K and LK sections of IFLS-CF data)

In Book K of the IFLS-HH data, a few cases had errors in the kecamatan code, SC03, for the given enumeration area in which the household resided. A few cases also had miscoded values for the urban/rural indicator SC05. These miscodings were due to transcription errors by the interviewers. All corrections for these miscodings have been incorporated into the data.

In the IFLS-CF data, a few communities had miscodes for kecamantan in the community survey portion of the IFLS-CF (files BUKILK1, BKIILK1, and BUKPLK1). These miscodings were due to transcription errors by the interviewers. Corrections for these miscodings have been incorporated into the data.

No cleaning of location codes was done in any other IFLS-HH or IFLS-CF subfiles. As noted in the IFLS-HH codebook, there are some location codes in the 5000 range. These are places where the village, kecamatan, and/or kabupaten names were inconsistent with the standard coding scheme used by the Indonesian Bureau of Statistics (BPS). In addition, some place names could not be matched to existing code lists and have nonresponse values in location code fields. Given that there were some miscoded location codes in the Book K subfile of the IFLS-HH database and in the LK1 subfiles of the IFLS-CF community data, users should expect that some miscoded location codes exist in other subfiles with location code data.

For privacy protection, the names of villages (desas) mentioned in the IFLS-HH and IFLS-CF databases have been dropped. In addition, in all but the migration history data, the identifier codes of villages when entered have been dropped as well. In the migration history data, Sections MG and SR in Book III, we have substituted a scrambled village identifier code so that users can detect moves between villages and moves to and from the same village. This scrambled code bears no relation to the standard desa codes and cannot be matched to any BPS desa code listings.

Non-identifier variables: IFLS-HH

Below we present a number of areas in the IFLS-HH data involving data not related to identifier variables. The first set of areas are problems that appear in more than one book of the IFLS-HH. These areas are birth dates and ages, gender codes, marital status, education levels and expenses, asset ownership, and rupiah fields. The remaining areas are confined to individual books: Book I household roster and consumption, anthropometry roster and measurements, Book III non-coresident rosters, and Book IV birth history and infant feeding.

Books I, III, IV, V, CA: Birth dates and ages

In the household roster, the birth date of a household member was asked first before any questions about age. If the respondent could give a year of birth for the household member, the respondent was not asked to give the current age of the person. If a year of birth was given, then the interviewer computed a current age based on the interview date and the birth date. Interviewers were not to fill in the birth year if the respondent only knew the person's current age. In some cases, the interviewers made incorrect calculations of the current age from the reported birth year. Most often the current age computation was off by 10 years due to poor subtraction (e.g., someone born in 1978 was listed as age 25 instead of 15). Users should be aware of this problem and should cross check year of birth with current age to find inconsistencies. A suggested global correction is found in the BUKKAR2.FIX program, which requires adding the interview date to the BUKKAR2 (household roster) file. Please note that ages on Book III, IV and V covers were often taken from the household roster--if age was computed wrong in roster, then it was wrong on the Book III, IV or V cover. The following program files contain corrections to respondent ages on book covers: BUK3S3A.FIX, BUK4SIV.FIX, and BUK5SE5.FIX. In addition, ages and birth dates were often copied from the household roster to the anthropometry roster. Here to, we have prepared corrections for the anthropometry roster that incorporate corrections made to the household roster (BUKCCA2.FIX).

A couple other points regarding the birth date information in the household roster. For a few children born less than one year before the survey, the “year of age” value in AR09YR was coded “96”/”98”/”99” rather than zero, as it should have been. In a few case, full birth dates were given and the interviewer did not try to compute a current age. Those individuals over age 95 were to be coded

as AR09YR=95. There were a handful of cases where the actual age was recorded (96,97,98 or 99). Since the 96-99 codes have specific meaning for missing values, those few cases should be set to 95. The suggested fixes in BUKKAR2.FIX include code to correct these cases. In addition, a few Book III respondents had ages and birth dates coded as "98" (don't know) in the household roster, but on the Book III cover the interviewer wrote down "95", mistakenly assuming the "98" code for the roster age meant "98 years old". Corrections for those few cases can be found in BUK3S3A.FIX.

Lastly, disagreements between age reported on the Book III, IV, and V covers and the household roster age still exist. Likewise some disagreements still exist between the household roster and the anthropometry roster. These differences are due to respondent error. Sometimes a respondent would give two different age responses for the same person, sometimes different respondents give different ages for the same person, and sometimes a respondent appears to have rounded an age to the nearest multiple of 5 or 10. For example, someone who is age 32 might be reported as age 30 in the roster and age 32 on the Book IV cover. Users should be aware that age heaping around these multiples of five does exist in the IFLS-HH data.

Books I, V, CA: Gender codes

Inconsistencies between the gender code in roster (AR07), gender code in the anthropometry roster (CA05), and the gender code in the Book IV birth history (CH08) were investigated. Where AR07 was obviously incorrect due to gender-specific names or to corroborating evidence from the anthropometry roster (based on name matches) or the Book IV cover respondent (names in roster and Book IV matched), corrections to AR07 were made. Sometimes the error was in the birth history, and CH08, the sex identifier code, was changed accordingly. Sometimes the error was in the anthropometry roster. Some additional corrections to gender codes in the household roster can be found in the program file BUKKAR2.FIX; some additional corrections to gender codes in the Book IV birth history data are in the program file BUK4CH1.FIX. In addition, corrections to gender codes in the anthropometry roster (which were also often taken from the household roster) that account for changes made above are found in the program file BUKCCA2.FIX.

Books I, III, IV: Marital Status and Spouse Identifier

Errors were found in marital status codes and spouse identifier codes. Some women in the household roster who were the wife of the head were coded as "not currently married" in AR13, the marital status variable. In the household roster, heads and spouses were coded as married, but the spouse identifier codes (AR14) were "96" instead of "2" for the head and "1" for the spouse. This use of "96" makes one wonder if sometimes interviewers felt it wasn't necessary to write down the marital status code (AR13) since one could tell who the spouse was from the relationship code (AR02), and thus it was "not applicable". In addition, there were children in the household roster

under age 10 who were coded as “married” (AR13=2), and some age 10-13 children as well. For problems in the marital status (AR13) and spouse identifier (AR14) codes, a program of suggested fixes, BUKKAR2.FIX, is available to use.

As with ages, marital status was usually copied from the household roster to the Book III and Book IV covers (STATUS3 and STATUS4)--if the marital status on the roster was wrong, then it was wrong on the Book III and IV covers. The following program files contain corrections to respondent marital status on Book III and IV covers: BUK3S3A.FIX and BUK4SIV.FIX.

Books I, III, V: Level of Education

Cursory checking of the level of education information in the household roster (subfile BUKKAR3) and in the Book III and Book V education sections (subfiles BUK3DL2 and BUK5DL1) revealed discrepancies between education levels reported in the household roster and those reported by the Book III/V respondents in their relevant education sections. Some of these discrepancies appear to be due to miscodings of the education level, others appear to be due to responses begin given by different respondents between the roster and the Book III/V data, and others appear to be due to differences in how those who very recently graduated a given level were handled.

A few miscodings found in the household roster have suggested fixes in the program file BUKKAR3.FIX. Differences due to different respondents are considered to be valid and tend to show the level of knowledge about others living in the same household. The difference in how recent graduates are handled has also not be resolved in the data, but the issue is discussed below.

Recent graduates of a level may be coded one of two ways. They may be coded as having graduated a given level (i.e., the grade completed variable has a code of 7), or they may be coded as in the current level but have a grade completed code of 96 (not applicable) since they have not completed the first grade of the current level. For example, someone who just graduated from primary school and is in the first year of junior secondary may be coded as either level=2 and grade=7 or as level=3 and grade=96. If users want to create highest level and grade completed, they must recode the grade=96 cases back to the previous level.

A last item related to education level is the coding in of the “other, specify” values. Virtually all of the responses to the “other, specify” field for education level were responses that fell in the range of codes originally provided to respondents. Thus, for these cases, the education level variable was recoded to the appropriate response value. This was done in the subfiles BUKKAR3 (variable AR16), BUK3DL2 (variable DL06), and BUK5DL1 (variable DLA07). An indicator variable was added to each file to let users know which cases had their education level variable recoded.

Books III, V: Education expenses

Education expenses in Book III (subfile BUK3DL3, variables DL31BR1-DL31BR5) and Book V (subfile BUK5DL2, variables DLA25R1-DLA25R5) were to be recorded in thousands of rupiah, i.e., rounded to the nearest thousand rupiah. For example, 100,000 rupiah was to be recorded as 100. In some cases, it appears that the interviewers recorded the amount as actual rupiah and not rounded to the nearest thousand rupiah. In such a case, the 100,000 rupiah above would have been recorded as 100,000, implying an expense of 1,000,000 rupiah instead of the actual 100,000 rupiah. An analysis of education expenses led to a set of suggested corrections to cases where it appeared that this interviewer error occurred. Those suggested fixes appear in the program files BUK3DL3.FIX and BUK5DL2.FIX.

Books II, III: Percent share of asset ownership by husband and wife

In the asset sections of Books II and III (BUK2HR1 and BUK3HI1), respondents were asked about the percentage of an asset owned by the household (HR07 and HI08), the percentage owned by the respondent (HR12A and HI13A), and the percentage owned by the respondent's spouse (HR12B and HI13B). That percentage represents the percentage of the total asset and not the percentage of the respondent's and respondent's spouse's share. However, there are some cases where the respondent gave the percentage split between he/she and his/her spouse instead. For example, a respondent says that his household owns 70 percent of an asset but then says he is the sole owner of the asset within the household and thus owns 100 percent of that asset and his spouse owns none of it. The correct response should be that he owns 70 percent of the asset, not 100 percent. Another problem with respondent confusion takes the form of reporting equal ownership of an asset as each owning 100 percent. For example, a household owns 70 percent of an asset and the husband and wife are the only members of the household who own the asset and the husband reports that he owns 100 percent and his wife owns 100 percent. Here the correct response should be that he owns 35 percent of the asset and his wife owns 35 percent.

The most obvious cases with such problems as those above has suggested fixes available in the program files BUK2HR1.FIX and BUK3HI1.FIX. However, these corrections do not cover all the possible problem cases. Users should doublecheck the consistency between the percentage owned by the household, who in the household owns the asset, and the percentage of the asset owned by the respondent and the respondent's spouse. After incorporating the corrections in the two "fixes" files, users should again check for cases where the sum of the respondent and the respondent spouse shares exceed the percentage owned by the household (e.g., (HR12A+HR12B)>HR07 in BUK2HR1) Such cases should then be investigated by users to determine how to handle them.

Books I, II, III, IV, V: Rupiah fields

Sometimes the "not applicable", "refused", "don't know", and "missing" codes (96/97/98/99) were not always entered correctly in question fields for large rupiah amounts. For example, a "don't know" rupiah amount entered in three parts (millions, thousands, and hundreds), should have been coded as 999/999/998. However, in some cases, it was coded 998/998/998 returning a numeric rupiah amount of 998998998 instead of the expected 999999998. Suggested fixes for those coding errors can be found in: BUK2UT2.FIX, BUK2NT2.FIX, BUK2PH1.FIX, BUK2PH2.FIX, BUK2GE1.FIX, BUK2HR1.FIX, BUK3TK1.FIX, BUK3TK5.FIX, BUK3MG3.FIX, BUK3BA2.FIX, BUK3BA6.FIX, BUK3BA7A.FIX, BUK3TF3.FIX, BUK3TF4.FIX, BUK3HI1.FIX, BUK3KW2.FIX, BUK4KW2.FIX, BUK4CX1.FIX, BUK1KR1.FIX, BUK1KS3A.FIX, BUK1KS3B.FIX, and BUK5DL2.FIX.

The problem with incorrectly coded nonresponse codes was less of a problem in the IFLS-CF data. The few corrections needed there were incorporated into the IFLS-CF data.

Users should also note that some rupiah fields collect actual rupiah amounts while others only collect rupiah amounts rounded to the nearest thousand rupiah. As seen in the comments about education expenses above, there may be some cases where instead of writing down the amount in thousands of rupiah, the interviewer wrote down the actual amount. When investigating outliers, users may discover such cases. Another item of note is that some rupiah amount questions collected monthly amounts and others collected annual amounts. Users should be sure to check the unit of time to which the given question applies.

Book I: Relationship to Household Head

The "relationship to household head" question, AR02, appears to generally have been filled in correctly. For some households, though, the interviewer put down the sequence number for the person and not the relationship code (e.g., a household of 5 people consisting of a head, spouse and 3 children, had AR02 codes of 1,2,3,4,5 instead of 1,2,3,3,3). There were also cases where the head of the household was a woman and the family members were her children and the second person is the household was listed as her spouse out of force of habit--person one is always the head and person two is his spouse. In other cases, the AR10/AR11 codes indicated that the parent of the spouse was in the household, but the AR02 code for that individual was "parent of head". Checks of the father/mother information in the Book III section on parents confirmed that the spouse's parent was in the household and not the head's parent. Sometimes, it was the other way around, with the head's parent in the household but the AR02 code said "parent in law". For these very obvious cases, AR02 has been fixed in the BUKKAR2 data.

Book I: Consumption

Duplicate food item records with zero values were found for five households in the food consumption subfile BUK1KS1. These blank duplicate records were dropped and the variable containing a count of the number of food items purchased was fixed to reflect the correct number of items.

In the consumption section, the subsections about transfers of food (subfile BUK1KS2A) and nonfood items (subfiles BUK1KS2B and BUK1KS3A) contained a question regarding whether the expenditure was for the whole household or for just some portion (variables KS05, KS07 and KS09). In examining expenditures reported in those subsections by that question, it appears that for some goods, most respondents report household expenditure. For other goods, 10 to 25 percent (only recreation and transportation are over 20 percent) of the respondents report expenses that are not for the entire household. Reporting expenses for less than the entire household appears to occur more often in higher income households (i.e., respondents from higher income households were more likely to report what appears to be just their own expenses). Users should be aware that the KS05, KS07 and KS09 variables indicate that for a given item there may an underreporting of the total household expenditure on that item. Such information has generally not been available before to researchers. Unfortunately, there is no information on just how much of the entire household expenses were reported by the respondent. Comparison of expenses reported as for the entire household to those reported as for less than the entire household suggests the gaps between the two for some items can be quite large at higher income levels. When creating household expenditure measures, users should be sure to utilize the information on underreporting provided in variables KS05, KS07 and KS09.

Book II: Health Insurance

In the AK section of Book II, health insurance questions were broken down into four groups: employer-provided insurance (subfiles BUK2AK1 and BUK2AK2), individual policies (subfiles BUK2AK3 and BUK2AK4), medical benefits provided by the company (subfiles BUK2AK5 and BUK2AK6), and company health facilities (BUK2AK7 and BUK2AK8). These groupings were meant to be mutually exclusive--a given health insurance plan/benefit/facility could fall in only one of the four groupings. A few respondents were confused by these groupings and users will find inconsistent responses. For example, users will find a few cases where government provided plans for government employees (e.g., ASKES, PHB) are listed under individual policies, medical benefits, and company clinics; in a few cases, users will find insurance plans or health facilities listed under medical benefits.

In such cases where government insurance plans are listed in subfiles BUK2AK3 to BUK2AK8, users may want to pull out those cases and pool them with the employer-provided

insurance data.[8] Users should do this pooling if they want to investigate the amount of health insurance coverage provided by government plans relative to private plans.

Respondents were asked about health insurance used by members of the household, and then were asked a few questions about each of the health insurance providers/benefits used by the household. The questions about the providers/benefits were to be just one record per provider/benefit. In a few cases, the interviewers repeated the provider/benefit information as many times as there were household members that used the provider/benefit. Such duplicates were dropped (subfiles BUK2AK2, BUK2AK6, and BUK2AK8). In a couple of cases, interviewers wrote down the names of the covered individuals in the fields for provider/benefits (again creating duplicates). These cases were fixed by inserting the proper provider/benefit and dropping the extra records.

In a few cases, the number of providers/benefits listed in the first part of a given AK subsection (subfiles BUK2AK1, BUK2AK5, and BUK2AK7) did not agree with the number of provider/benefits listed in the second part (subfiles BUK2AK2, BUK2AK6, and BUK2AK8). Such cases were not corrected since it was unclear which data was in error.

To aid users, we created insurance type character variables from the insurance plan names given in the employer-provided health insurance section (subfiles BUK2AK1, BUK2AK2). The insurance type variable groups plans into eleven categories, with eight representing forms of government provided insurance and three representing non-government supplied plans. In subfile BUK2AK1, the type variables are TYPE, SOURCEA, SOURCEB, SOURCEC, and SOURCED; in subfile BUK2AK2 the type variable is TYPE. The names of the insurance plans recorded by interviewers are still included in the data for users who wish to examine them (variables AK05, AK02A, AK02B, AK02C, and AK02D in subfile BUK2AK1, and variables AK06 in subfile BUK2AK2).

Book III: Non-coresident parents

The non-coresident status of parents recorded in Section BA of Book III was compared to the presence of parents in the household roster. Some discrepancies do exist. Parents listed in Section BA (variable BA04 in file BUK3BA1) as living in the household do not appear in the household roster, and parents living in the household roster are listed as not living in the household in Section BA. Some of these discrepancies appear to be due to different definitions being used in Section BA and in the household roster as to what constitutes the household. If interviewers did not remind the respondent about the definition of who is considered a household member, the respondent may have followed a broader interpretation than that used for the household roster--thus a person who is

[8]Government sponsored programs will have any of the following in the name reported by respondents: ASKES, PERUM HUSADA BHAKTI, ASTEK, INSURANSI KESEHATAN.

"outside the household" by the roster's definition may be "in the household" by the respondent's definition. Another source for the discrepancy may be the confusion by respondents as to who was to be listed as their father or mother in the household roster. In some cases where the spouse's parents lived in the household, the respondent also said his/her spouse's parents were his/her own parents. It is possible that some of this confusion arose in Section BA of Book III as well. While some household roster records were corrected for errors in father and mother identifier codes, no corrections on inside/outside household status have been made to the BUK3BA1 data on parents. Users should crosscheck BA04 in BUK3BA1 (and BA05--parent alive or dead) against the household roster AR10 and AR11 codes for the Book III respondent. Users must then make their own decisions as to how to handle discrepancies between the BA section and the household roster information.

Book III: Non-coresident rosters

In the non-coresident selected child roster in Book III, question BA76 asks "how long ago" did the child leave and the response was supposed to be in length of time (i.e., number of years ago). Some respondents gave a date as opposed to a length of time. In such cases, the interviewer wrote the date information into BA76. Thus, in the file BUK3BAA, there are values for BA76YR which are actually year dates (e.g., 89, 90, 91 and 92). A suggested set of fixes for such cases has been proposed and is found in the BUK3BAA.FIX program.

Another problem area in the non-coresident sibling and children sections was the counts of types of siblings and number of eligible siblings and children. For non-coresident siblings, the counts in BA29 did not agree with the full sibling roster and/or the counts in BA30H did not agree with the selected sibling roster. For non-coresident children, the "eligibility" flag in BA71 was sometimes turned on for children under age 15--only those age 15 and over were to be eligible for the selected non-coresident children roster (questions BA74-BA90). In some cases, children under age 15 were included in the selected children roster. This tended to occur when there were less than 5 total non-coresident children and the interviewer listed all non-coresident children in the "selected" children roster. Many of these included children are listed as age 14 in BA66. It may be that interviewers were uncertain about the exact age and chose to include them. A suggested set of corrections to uncovered problems in BA29 and BA30H can be found in the programs BUK3BA3.FIX and BUK3BA5.FIX; suggested fixes to BA71 are in BUK3BA8.FIX.

Book IV: Stillbirth coding

We found records in the Book IV birth history where a child who had died after birth was coded as a "stillbirth" (CH06=3) instead of a live birth (CH06=2). These children died days, weeks, months and even years after they were born. Infant feeding information exists for them as well as death dates. A check of the pregnancy summary information in section BR of Book IV (subfile

BUK4BR1) showed that the woman reported having children born alive who subsequently died and do not report any stillbirths. Such cases have not been fixed in the data and a suggested set of corrections to CH06 is available in the program called BUK4CH1.FIX.

Book IV: Birth history skips

In the birth history in Book IV, question CH12 is a skip indicator. The code of 5, which told the interviewer to skip to question CH41, was not to be circled for an existing pregnancy; it was only to be circled in the column after the last pregnancy. However, some interviewers were confused by the CH12 question and circled "5" on the record of the last pregnancy. This led them to skip the CH13-CH40 questions for the last pregnancy. In other instances, it was the editor who was confused. In such cases, the editor erased the interviewer's coding and circled the 5 code on CH12 for the last pregnancy record, even though the interviewer had asked questions CH13-CH40. Because the data entry program took the CH12=5 as a valid skip, the program skipped over questions CH13-CH40, forcing the data entry person to start with the CH41 responses.

We were able to recover the CH13-CH40 information for a handful of these cases (there were 40-50 of them) by checking the original hardcopy questionnaire, and those fixes have been incorporated into the data. For those few cases, it appears that after all the questions had been asked for the last pregnancy, someone, either the interviewer or the editor, went back and circled the "5" code in CH12 and marked out the previously correct CH12 response. The remaining CH12=5 records had no CH13-CH40 information filled in on their hardcopy questionnaires. Thus some youngest births to Book IV respondents have no CH13-CH40 data (e.g., no prenatal care, birthweight, alive/dead, or infant feeding information).

Book IV: Infant feeding data

In the Book IV birth history, some confusion arose over the use of the "96" code for the questions on when first regular food was given (CH31) and when breastfeeding was stopped (CH32). The questionnaire specifically states that "96" means "not yet/still breastfeeding". However, there are children who are too old to still be breastfeeding or to have not received regular food (e.g., 10, 12, and 20 years old) and other children who are dead, and thus cannot be still breastfeeding. Examination of the cases relating to dead children suggests that in some cases, the interviewer coded "96" if the child was still breastfeeding when it died. This was not always done for children who died while breastfeeding, many such children have an age at weaning response in CH31 and CH32. In cases where the child did not die, it is not clear what the "96" code means--it could mean "not applicable" as it does elsewhere in the survey, but it's not clear what would be "not applicable". In some of these "too old" cases, all children have "96" codes in CH31, and in others only one or two. A

few corrections to CH31 and CH32 are available in the program file BUK4CH2.FIX. However, that program file does not include any suggested fixes for the types of cases described above (i.e., kids too old to still be breastfeeding, kids who are currently dead, and so on). Users need to decide for themselves how to handle such cases.

Book CA: Anthropometry roster

The transcription of data from the household roster to the anthropometry data rosters suffered from the fact that the order of "father identifier, mother identifier" was reversed between the two rosters as was the "date of birth, age" order. In the household roster, AR10=father identifier, AR11=mother identifier; in the anthropometry roster (CA), CA03=mother identifier, CA04=father identifier; in the household roster, AR08 was birth date, AR09 was age; in the anthropometry roster, CA06 was age, CA07 was birth date. In transcribing from the household roster, sometimes the anthropometrist did not notice the changed order, especially for the mother/father identifiers. Age and birth date were to be asked again, so there appears to be less transcription error here, except where the anthropometrists did not re-ask the information. Another transcription problem between the household and anthropometry rosters was that the number and order of people on the anthropometry roster could be different from the household roster. Sometimes anthropometrists did not read the demographic information (CA03-CA07) from the correct line in the household roster. Fixes to CA02 and CA05 (sex) based on name matches have been incorporated into the data. Fixes to known errors to CA03, CA04, CA06 and CA07 due to transcription can be found in the program file BUKCCA2.FIX.

In general, we suggest that users use the household roster data for demographic information and not the anthropometry roster, except for individuals who are in the anthropometry roster but not in the household roster (this did happen in a few cases). Discrepancies still exist between the household roster and the anthropometry roster even after all the corrections in the "fixes" files have been incorporated. Many of those discrepancies are due to different reports from the same respondents and not interviewer error. Our general rule was to not correct cases of respondent-based differences since it is not always clear which response is correct. Also, since information about an individual is obtained from that individual and from various proxy respondents, it is expected that answers are not always consistent.

Book CA: Height and Weight

The scales used to measure weight could be set to either kilograms or pounds. The scales were to always be set at kilograms. However, it appears that in some cases, the scales were accidentally changed to pounds and the weights recorded in the anthropometry section were in pounds and not kilograms. Most of these cases have been detected, although

probably not all, and those detected have been corrected in the data. In addition, there were a few cases where the anthropometrist wrote weight in the height column and height in the weight column. Such transposition problems that have been uncovered have been corrected in the data.

Some anomalies in height/weight/age combinations remain in the public use data. Users should check the relationship between height, weight and age (and sex) to isolate remaining incongruities. It is up to the users to decide how to handle such cases.

Nonidentifier variables: IFLS-CF

Rupiah fields

Sometimes the “not applicable”, “refused”, “don’t know”, and “missing” codes (96/97/98/99) were not always entered correctly in question fields for large rupiah amounts. For example, a “don’t know” rupiah amount entered in three parts (millions, thousands, and hundreds), should have been coded as 999/999/998. However, in some cases, it was coded 998/998/998 returning a numeric rupiah amount of 998998998 instead of the expected 999999998. The problem with incorrectly coded nonresponse codes was less of a problem in the IFLS-CF data. The few corrections needed there were incorporated into the IFLS-CF data.

A Note about Data Collected in the Form of Lists

In a number of parts of the Community and Facility questionnaires, respondents were asked to provide data in a list format. In some cases respondents could generate a longer list than there was space for in the instrument. For example, in Sections I and J of the Village Leader (Book I) and Village Women's Group (Book PKK) questionnaires respondents were asked to list health facilities and schools used by village residents. Respondents could list up to five health centers. If the respondent could think of more than five health centers, the information was not recorded because there was no additional space. In cases such as these, the list is truncated. Users may wish to check the maximum number of responses allowed and treat cases in which the maximum is attained as lower bounds on the particular item.

General Notes about the Information in the Village Leader Questionnaires

Books I, II, and PKK of the Community Survey were designed to provide general information on village infrastructure. Examination of the data suggests that respondents were readily able to answer the questions. Patterns of responses and correlations across responses do not suggest any major problems with data quality. This finding is not terribly

surprising, since the questions are generally fairly straightforward and the respondents are used to providing similar data for various government purposes.

Book I: Public Transportation

Questions in the transportation section focused on availability of public transportation and on travel times, costs, and transportation modes to the nearest bank, market, post office, telephone office, and to the subdistrict, district, and province capitals (subfile BUKIA01A). The information on travel time and cost was only collected if the institution in question was not located in the village. Village leaders were generally able to provide answers to about distance, travel time, and travel cost that are consistent with expectations that more plentiful institutions are closer than are "rarer" institutions. For example, on average, subdistrict capitals are closer than district capitals, which in turn are closer than province capitals. On average, banks are no further away than the subdistrict capital, which is logical since most subdistrict capitals have a bank of some sort. On average markets are closer than post offices, which are close than telephone offices.

Book I: Electricity , Water and Sanitation

Questions in Sections B (electricity, subfile BUKIB01) and C (water and sanitation, subfile BUKIC01) were designed to identify sources of electricity, water for drinking and bathing, and methods of sewage and garbage disposal. The questions generally follow a similar progression. First, respondents are asked whether any households in the village receive the service in question from a specified list of sources. Next, the respondent is asked to identify the most important source of the service. For example, electricity may be available from the national government electric company (PLN), or from local governments, private cooperates, or privately owned generators. Respondents are asked to identify any problems with the service/source, such as daily or seasonal disruptions in supply.

From responses in Sections B and C it appears that multiple sources of a certain service or resource are quite common in Indonesian villages. There is also differential access to services/resources within villages, based on responses to questions about the percentage of village residents with electricity, and whether all households are connected to the sewage system. The variables in the water and sanitation section are meant to be used in combination rather than in isolation. Precursory consistency checks of distributions of various characteristics do not turn up any major anomalies. For example, fewer villages claim piped water as the main source of bathing and cleaning water than claim piped water as the main source of drinking and cooking water. We have not checked whether a respondent's answers to one question are inconsistent with his answers to another question

(e.g. does a respondent tell us that no one in the village has piped drinking water, but that the main source of drinking water is piped water).

Book I, PKK: Date of Establishment of Village Infrastructure

Respondents to the Village Leader (Book I) and Village Women's Group (Book PKK) questionnaires were asked to provide information on when certain services and types of infrastructure became available in the community (subfiles BUKIB01, BUKIC01, BUKII01, BUKIJ01, BUKPI01, and BUKPJ01). Respondents were asked to report either the year in which the change occurred, or the number of years ago that the change occurred. Missing values for questions of this type appear to be more common than on other items in the Book I questionnaire. Information on dates of major changes/events in the village was also collected in Section E of the questionnaire (subfile BUKIE02B). It may be possible to use data from Section E to fill in missing values on certain questions asked in other sections of the questionnaire.

Frequency distributions on variables measuring years elapsed since a service or infrastructure type became available appear plausible. The majority of responses suggest that major changes in infrastructure have occurred in the last 25 years. For answers of more than ten years ago there is some heaping of responses on values that mark either numbers of years ago ending in 5's and 0's (e.g. 15 years ago) or numbers of years ago ending in 3's or 8's (e.g. a change occurred in 1980 or 1975).

Book I: Agriculture and Industry

Question D29 asks about factories in or near the village (subfile BUKID04), while question D33 asks about cottage industries of village residents (subfile BUKID05). It appears from the text answers to these questions that the distinction between factories and cottage industries is somewhat blurred. Users are encouraged to draw on additional information (e.g. estimates of cottage industry revenue) in classifying industries as cottage or large-scale.

Book PKK, PPKBD: Availability and Prices of Foodstuffs

Section H of the Village Women's Group (Book PKK) and Community Health Post (Book PPKBD) questionnaires asks about prices and availability of basic foodstuffs (subfiles BUKPH01 and BKPPKH01). An effort is made to collect the data in a way so that prices of foodstuffs can be compared across villages. Accordingly, it is important to ask about quantity and about price for a known unit. In a few cases respondents did not provide adequate information from which to calculate price per volume. For example, rather than giving the

unit of measurement (grams, cc, etc.) the respondent gave an answer such as "piece" or "packet". This problem is relatively rare and so should affect only a few food items for a few communities.

Information about prices was asked in these two different questionnaires in an effort to get a broader picture of prices within the area. Because people shop at different places in a community, asking the head of the Village Women's Group and the individual Community Health Post respondent for prices provides that broader view. At the time of fieldwork, the Section H portion of Book PPKBD had to be added in manually by the interviewers as the section was not bound into the original IFLS-CF questionnaire book. This situation resulted in Section H not being administered as part of Book PPKBD in all EAs since some interviewers forgot to include it. Prices were collected in Book PKK for 311 EAs while prices were collected in Book PPKBD for 167 EAs. The breakdown between the two sources of prices is: 163 EAs had Section H filled out in Book PKK and in at least one Book PKKBD, 148 EAs had only the Book PKK report, and 4 had at least one Book PPKBD report (1 had one report, 3 had three reports). The average number of reports on price per EA is 2.5. The breakdown of number of price reports per EA is: 47 percent of the EAs had only one report, 3 percent had two reports, 4 percent had three reports, 45 percent had four reports, and 1 percent had five reports. Only six EAs did not have prices collected in either Book PKK or Book PPKBD, but these six EAs are among the nine EAs that had the same village head an another EA. Thus all six EAs have a counterpart EA where prices were collected. See the IFLS-CF *Codebook* for a discussion of those nine EAs.

Book PUSK: Head of Health Center/ Auxiliary Center

Ideally, the respondent to Section A (Head of Health Center) was the head of the health facility (subfiles BKPSA01-BKPSA03). Of the 993 health centers/ auxiliary centers in which we interviewed, the respondent to Section A was the head of the facility in 727 cases. From precursory examination of the data it appears that proxy respondents (who answered questions about the head of the facility) were able to provide information about the characteristics of the head of the facility.

Book PUSK: History of Development of Health Center / Auxiliary Center

Section B (Development of Health Center) contains a calendar (subfile BKPSB02) on which interviewers were to record beginnings, endings, and temporary interruptions in particular types of health center services and staffing patterns (e.g. the opening of a lab). If a service was available prior to 1980, the boxes representing calendar years were to be filled in with a straight line. If a service began in 1980, the 1980 box was to be marked with an X.

This protocol was established to distinguish between services that were in existence in 1980 and services that began in 1980. However, due to data entry practices, all boxes were filled with an X, so one cannot tell whether a particular service began in 1980 or earlier than 1980. The data entry staff were used to entering "X" for boxes with straight lines; unfortunately, a special data entry protocol for Section B of Book I was not implemented. If the service began after 1980, the box for 1980 is left empty, so the issue does not arise.

Book PUSK: Staff of Health Center / Auxiliary Center

In Section D (Health Center Employees), we inquired about counts of staff members of various types. More detailed questions were asked about the characteristics of doctors, nurses, and midwives. Approximately 8 spaces were provided to list doctors, nurses, and midwives (subfile BKPSD03). In some instances, facilities had more than 8 staff members about whom supplementary information was desired. For some of these facilities, information was collected on more than 8 employees, and in others, only 8 employees were listed, usually all doctors. Users should cross-reference the employee count data to ascertain whether the supplementary information is available for all eligible staff members.

Books PUSK, DR, BIDAN: Numbers of Patients in Registration Book in Week Prior to Visit

The goal of the questions in Section C of the Health Center/Auxiliary Center questionnaire and Section B of the Doctor's Practice and Nurse/Paramedic/Midwife Practice questionnaires was to ascertain the volume of business at the facility over a specified time period (subfiles BKPSC01A, BKPSC01B, BKPSC01C; BUKDB01-BUKDB03; BKBDB01-BKBDB03). A priori it was unclear that practitioners would allow us to record this information. However, large numbers of missing values do not seem to be an inordinate problem (the information is not recorded for about 10% of doctors and nurses). Users are advised to refer to the days of the week the facility is open when analyzing daily numbers of patients, since some facilities are not open on Saturdays and Sundays.

Books SD, SMP, SMA: Information on Principal

Section A of the school questionnaires collects information about the principal of the school (subfiles BUKSA01, BKSPA01, and BKSAA01). If the principal was not available, a proxy respondent was asked to provide information about the principal. Proxies responded to this section for 20% of the elementary schools (SD), 26% of the junior high schools (SMP), and 29% of the senior high schools (SMA). Proxy respondents appear to have been able to answer most of the questions, although there are more missing values on some items when proxies responded. The questions that appear to be most affected are year of principal's graduation

from highest level of school(about 6% missing when proxy responded), number of years of experience as a school principal prior to serving in current school (12-17% missing when proxy responded) and amount of time in training in last five years (7-10% missing when proxy responded).

Books SD, SMP, SMA: Shared School Building/Complex

In Section B of the school questionnaires, we asked whether other schools shared the same complex as the school in which we were interviewing (subfiles BUKSB01, BKSPB01, and BKSAB01). The answer was yes for 40% of elementary schools and 36% of junior high schools and senior high schools. It is not uncommon in Indonesia for two schools to share the same building, one using it in the morning and the other using it in the afternoon. Originally we had hoped to ascertain which schools were sharing the same physical building. However, the question tended to elicit responses relating to other schools within the same complex as opposed to the same building. Unfortunately, the distinction between schools sharing the same physical building and schools just sharing the complex of buildings cannot be made given the way the data were collected.

Books SD, SMP, SMA: Types of Allowances Provided to Teachers

In the school questionnaires, question C12 asks teachers what types of allowances they receive in addition to their salaries. Unfortunately, answers to this question were never keypunched (subfiles BUKSC01, BKSPC01, and BKSAC01).

Books SD, SMP, SMA: Teacher/Student Ratios

Teachers of Bahasa Indonesia and Mathematics (for Grades VI of Elementary School, III of Junior High School, and III of Senior High School) were interviewed as part the school questionnaires and were asked how many students were enrolled in their classes and how many attended on an average day (subfiles BUKSC01, BKSPC01, and BKSAC01). These data can be used as one source of information on teacher/student ratios. The other way to calculate teacher-student ratios is to use the information from Section E (subfiles BUKSE02, BKSPE02, and BKSAE02) on numbers of students at each level, numbers of classes at each level, and numbers of class teachers.

Books SD, SMP, SMA: EBTANAS Scores

In Indonesia students take a standardized test at the end of elementary, junior, and senior high school (after Form VI, III, and III, respectively). We asked administrators in schools in which we interviewed to show us their lists of the students' scores on the Math and

Bahasa Indonesia tests. From the lists we recorded the scores on these tests for a random sample of students taking the tests. The number of sampled students varied with the number who took the test, but the sampling scheme was designed to obtain data on approximately 25 students.

Of the 944 elementary schools interviewed as part of the facility survey, we failed to obtain EBTANAS scores from 52 of them. Of the elementary schools from which scores were obtained, the mean number of students taking the test was 35 (ranged from 1 to 209), while the mean number of scores that were recorded was 21.

Of the 895 junior high schools in which we interviewed, we failed to obtain EBTANAS scores from 95 schools. Of the junior high schools from which scores were obtained, the mean number of students taking the test was 149 (ranged from 1 to 578), while the mean number of recorded scores was 21.

Of the 581 senior high schools in which we interviewed, we failed to obtain EBTANAS scores from 70 schools (for Bahasa Indonesia) and from 192 schools (for mathematics). Among the schools from which we obtained Bahasa Indonesia scores, the mean number of students taking the test was 130, while the mean number of recorded scores was 21. Of the schools from which we obtained mathematics scores, the mean number of test takers was 119, while the mean number of recorded scores was 16. From comparing, for each school, the number of students who took the math test to the number of students who took the Bahasa Indonesia test, it is clear that in senior high school more students are tested in Bahasa Indonesia than in math. This is not true at the other levels of school.

Books SD, SMP, SMA: School revenues

Section F of the school facility questionnaires collected information on school revenues. Users should be aware that the quality of these revenue reports has not been assessed. It is not clear how accurate school officials estimates might be or how good their schools records are with regards to revenues. It is also not clear what incentives school officials might have to under- or over-report revenues. Thus, users should be cautious in working with these revenue estimates.

III. FILE STRUCTURE, FORMAT AND CONVENTIONS FOR IFLS SUBFILES

This section presents the structure and format of the IFLS public use database. In addition, several basic data conventions used in the IFLS database are discussed. The final topic of this section covers the sampling weights provided for the IFLS-HH and IFLS-CF data.

FILE STRUCTURE

The IFLS database is really two databases in one: the IFLS-HH (household) database, and the IFLS-CF (community-facility) database. The IFLS-HH contains 119 subfiles; the IFLS-CF has 141 subfiles. Appendix A lists the IFLS-HH subfiles by their associated questionnaire, and Appendix B lists the IFLS-CF subfiles. Those Appendix tables also include the number of records in each subfile and the unit of observation.

The IFLS data were split into these separate subfiles to facilitate the construction of subsequent analysis files. Users need only work with desired subfiles rather than the entire database. Issues of how to identify questionnaire respondents and how to link data from the various subfiles are discussed in subsequent sections.

DATA FORMAT

The 260 subfiles of the IFLS database are available in two formats: rectangular ASCII files (i.e., raw data) and SAS transport files. Each format is described below. **IFLS1-RR IS ONLY IN SAS TRANSPORT AND STATA FORMATS.**

Rectangular ASCII Files

ASCII-version subfiles carry the extension .**TXT**.

The IFLS-HH ASCII subfiles take up around 215 MB of space; the IFLS-CF files use around 30 MB of space. Individual files can be up to several MB in size. In the IFLS-HH data, half of the files are 1 MB or smaller, and most files are under 5 MB. File sizes are smaller in the IFLS-CF data, with all subfiles being 2 MB or less. Users should be sure to have sufficient space available if they wish to have all IFLS subfiles uncompressed at one time on their disk. Input statements are provided for each of the subfiles in the IFLS database. These statements are in SAS programming code, but can be edited for use by other languages. Table 3.1 lists the names of the files containing input statements.

The ASCII version files have the standard DOS end-of-line and end-of-file characters (i.e., ^M, ^L). If the user's system does not need these markers, the user can either convert

the ^M and ^L to what their system expects or the user can simply add 2 bytes to the LRECL of each subfile and drop the last record read in. LRECLs are provided in the input statement programs listed in Table 3.1.

While those who use the ASCII files may read the data in without the provided input statements, we suggest that users keep the variable names listed in the IFLS Codebook. Communication regarding specific variables will be much easier if common variable names are used. In addition, the SAS input statement programs do not include variable labels or length statements. Those using these programs in SAS will want to apply lengths to variables in order to reduce subfile sizes.

Table 3.1

PROGRAMS CONTAINING INPUT STATEMENTS

DATA READ	FILE NAME
IFLS-HH	
Book CA: Anthropometry	BUKCAINP.SAS
Book K: Control book	BUKUKINP.SAS
Books I and II: Household Roster, Characteristics, Economy	BUK12INP.SAS
Book III: Adult Information--Book Cover, Sections DL-SR	BUK3AINP.SAS
Book III: Adult Information--Sections KM-CP	BUK3BINP.SAS
Book IV: Ever-married Woman Age 15-49	BUKU4INP.SAS
Book V: Children Information	BUKU5INP.SAS
OTHTRANS: Other, specify translations	OTHTRINP.SAS
WEIGHTS: Household and Individual Weights	WGTINP.SAS
LOCATION: Geographic codes	GEOINP.SAS
RND2_EAS: EAs Completed Before Round 2 Retraining	RND2INP.SAS
IFLS-CF	
Book Sample: Book Completion	SAMPLINP.SAS
Book I: Community--Kepala Desa	BUKUIINP.SAS
Book II: Community--Village Records	BUKIIINP.SAS
Book PKK: Community--PKK	BKPKKINP.SAS
Book PUSK: Gov't Health Centers	BKPSKINP.SAS
Book DR: Private Physicians/Clinics	BUKDRINP.SAS
Book BIDAN: Nurses, Midwives, Paramedics	BKBIDINP.SAS
Book PPKBD: Community Health/FP Posts	BKPPKINP.SAS
Book PTRAD: Traditional Healers	BKTRDINP.SAS
Book SD: Primary Schools	BUKSDINP.SAS
Book SMP: Junior Secondary Schools	BKSMPINP.SAS
Book SMA: Senior Secondary Schools	BKSMAINP.SAS
WEIGHTS: Community and Facility Weights	SMPWTINP.SAS

NOTE: Programs contain SAS input statements and have DOS end-of-line characters

SAS Transport Files

SAS transport files carry the extension .**EXP**.

The IFLS-HH transport subfiles take up around 300 MB of space and the IFLS-CF transport files use around 40 MB of space. Individual SAS files can range from 10K to 10 MB. Users will need over 340 MB of free space if they wish to have the entire IFLS database uncompressed at one time.

The SAS transport files were created under SAS Version 6.09 for UNIX using the SAS export engine (XPORT) and PROC COPY. Below is an example of the SAS code used to create the transport files:

```
LIBNAME BUKU1 'pathname of sas file library';
LIBNAME BUK1KR1 XPORT '/pathname/buk1kr1.exp';
PROC COPY IN=BUKU1 OUT=BUK1KR1;
SELECT BUK1KR1;
```

The following code is an example of how to convert the export file back into a SAS data file under UNIX SAS:

```
LIBNAME BUK1KR1 XPORT '/pathname/buk1kr1.exp';
LIBNAME BUKU1 'pathname of sas file library';
PROC COPY IN=BUK1KR1 OUT=BUKU1 ;
```

The following code is an example of how to convert the export file back into a SAS data file under PC SAS (this may work as well for other platforms if the XPORT engine above does not):

```
LIBNAME BUK1KR1 SASV5XPT '/pathname/buk1kr1.exp';
LIBNAME BUKU1 'pathname of sas file library';
PROC COPY IN=BUK1KR1 OUT=BUKU1 ;
```

For more details on reading SAS transport files, users should consult SAS documentation transporting files between operating systems and platforms.[9] In addition, users with the software program DBMSCOPY can use that program to convert the SAS transport file into other file types such as SPSS and STATA.

---

[9]See SAS® Technical Report P-195: Transporting SAS Files Between Host Systems, Cary, NC: SAS Institute Inc., 1989.

Users working on CMS operating systems should be aware that in moving SAS transport files from CMS to another platform by using FTP, file transfer protocol, users should specify “BINARY F 80” as opposed to just “BINARY” for FTP transfer. For more information on moving SAS transport files with FTP, KERMIT, or PATHWORKS, see SAS Usage Note V6-COPY-5371. Usage notes can be found on SAS Institute’s World Wide Web Home Page at http://www.sas.com.

The SAS data subfiles use the variable names listed in the subfiles' sections in the relevant IFLS Codebook. In order to reduce file size, the SAS variables have been stored in the shortest lengths possible given their maximum values.

DATA CONVENTIONS USED IN IFLS

The IFLS database is fairly straightforward in terms of data conventions. The four items of note are the identifier variables, missing and not applicable codes, variable names, and file names.

Identifier Variables: IFLS-HH and IFLS-CF

There are two main identifier variables in the IFLS-HH data are:

CASE: Main household identifier

PERSON: Person number from household roster

The IFLS-HH codebook introduction describes these two identifier variables. The variable CASE is on all IFLS-HH subfiles. The variable PERSON is on all Book III, IV and V subfiles.

The two main identifier variables in the IFLS-CF data are:

EA: Enumeration Area

FACCODE: Facility identification code

The IFLS-CF codebook introduction describes these two identifier variables. The variable EA is on all IFLS-CF subfiles; the variable FACCODE is on all facility subfiles.

Some of the IFLS-HH subfiles contain references to health and educational facilities used by IFLS-HH households. Such subfiles contain the variable FACCODE to identify the facilities mentioned in the household portion of the survey. Not all facilities mentioned in the household survey were interviewed as part of the community-facility survey, so there are FACCODE values in the IFLS-HH data that have no match to the IFLS-CF data; likewise, there are facilities in the IFLS-CF data that were not mentioned in the household survey but whose names were obtained from interviews with village leaders. Thus, there will be FACCODE values in the IFLS-CF data that have no match to the IFLS-HH data. The IFLS-HH subfiles that contain the FACCODE variable are discussed in the IFLS-HH *Codebook* introduction.

SEE IFLS1-RR DOCUMENTATION DRU-1195/7 FOR ADDITIONAL IDENTIFIERS ADDED TO IFLS1-RR FOR LINKING TO IFLS2.

Non-Response Codes

If the interviewer was to skip a section of the questionnaire, the fields of that section were left blank in the data entry program. For such cases, blanks in the raw data and "." in the SAS data are used to denote "not applicable". In all other cases, non-response codes are present if the individual did not provide a response. Table 3.2, which appears in both IFLS codebooks, presents the values used for different types of non-response. The table is present here again to remind users of the non-response code values.

Variable names

Variable names in the IFLS-HH and IFLS-CF originated from the ISSA data entry program and generally reflect the question number associated with the variable. In the IFLS-HH, question numbers reflect the questionnaire section as well as the sequence number of the question. For example, variable BA04 is in the BA section of the IFLS-HH questionnaire, and variable CH25 is in the CH section. The introductory section of the IFLS-HH *Codebook* contains a table that shows in which book each section appears. In the IFLS-HH questionnaire, section names are unique and do not appear in more than one book (with the exception of the CP section on interviewer notes).

Table 3.2

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

In the IFLS-CF data, variable names include both the section name and an identifier for which book the variable is from. Unlike the IFLS-HH, section names in the IFLS-CF

questionnaire were not unique across books, so to identify which book a question was from in the ISSA data entry program, a prefix of one letter was added in front of the question number to create the variable name. The variable naming convention in the IFLS-CF used the following prefixes:

B: Book I
C: Book II
D: Book PKK
E: Book PUSK (Government Health Center)
F: Book DR (Physician/Clinic)
G: Book BIDAN (Nurse, Midwife, Paramedic)
H: Book PPKBD (Community Health Post/FP Post)
I: Book PTRAD (Traditional Practice)
J: Book SD (Primary School)
K: Book SMP (Junior Secondary School)
L: Book SMA (Senior Secondary School)

Thus, variable EC01 is question C01 in Book PUSK (health center) and variable JC01 is question C01 from Book SD (primary schools). Variable C01 differs from book to book because the contents of Section C can vary from book to book.

Within the set of health facilities, the sections for the Health Centers (*puskesmas*), Physician/Clinic, and Nurse/Midwife/Paramedic tend to be very similar and covering the same type of information. The sections in the Community Health Post (*posyandu*) and Traditional Practice books differ from the first three health facilities and from each other. For the set of schools, on the other hand, the sections are identical, so variable JC01 is the same as variable KC01 and variable LC01.

Subfile names and contents

The structure of the IFLS-HH and IFLS-CF subfiles is a product of the ISSA data entry program's structure. Subfile names reflect the questionnaire book from which the data originates (e.g., BUK3 for Book III in the IFLS-HH, BKSA for Book SMA (Senior Secondary) in the IFLS-CF) and the questionnaire section and subsection (e.g., BUK3BA1 for the first subfile from Section BA in Book III, and BKSAD01 for the first subfile from Section D in the senior secondary school book). Table 3.3 presents the subfile prefixes associated with each book in the IFLS-HH and IFLS-CF questionnaires.

Table 3.3

SUBFILE PREFIXES BY QUESTIONNAIRE BOOK

Survey	Book	Prefix
IFLS-HH	Book K	BUKK
	Book I	BUK1
	Book II	BUK2
	Book III	BUK3
	Book IV	BUK4
	Book V	BUK5
	Book CA	BUKC
IFLS-CF	Book I	BUKI
	Book II	BKII
	Book PKK	BUKP
	Book PUSK	BKPS
	Book DR	BUKD
	Book BIDAN	BKBD
	Book PPKBD	BKPPK
	Book PTRAD	BKPTR
	Book SD	BUKS
	Book SMP	BKSP
	Book SMA	BKSA
	Book SAMPLE	SAMPL

NOTE: The household roster subfiles have the BUKK prefix and not the BUK1 prefix because the data entry program included the roster data in the Book K portion of the program so that the roster could be accessed for control book checks.

Subfile contents reflect data entry screens. If a questionnaire section required more than one data entry screen, there would be more than one subfile of data generated. Therefore, some individual sections of the IFLS-HH and IFLS-CF questionnaires were split out into separate subfiles even though they have the same unit of observation and are a continuation of same type of information. Users may wish to join such split sections back together again. Table 3.4 presents the sets of IFLS-HH subfiles and Table 3.5 presents the sets of IFLS-CF subfiles that users may want to initially link together.

NOTE: IN IFLS1-RR, THE FILES DESCRIBED IN TABLES 3.4 AND 3.5 HAVE BEEN JOINED, THUS THESE TABLES ARE IRRELEVANT FOR IFLS1-RR

Table 3.4

IFLS-HH SPLIT SECTIONS THAT CAN BE JOINED

Files That Can Be Joined		Variables to Create	Variables To Merge By
BUKCCA2	BUKCCA3		CASE, LINE_CA
BUKKAR2	BUKKAR3		CASE, AR001A
BUK1KS1	BUK1KSA		CASE, ITEM
BUK3TK2	BUK3T2B		CASE, PERSON, PERIOD
BUK3BA1	BUK3BA2		CASE, PERSON, PARENT
BUK3BA4	BUK3BA6	CREATE BA30A = BA34 IN BUK3BA6	CASE, PERSON, BA30A (ONLY KEEP MATCHES)
BUK3BA8	BUK3BAA	CREATE BA63A=BA74 IN BUK3BAA	CASE, PERSON, BA63A (ONLY KEEP MATCHES)
BUK3TF1	BUK3TF2		CASE, PERSON, RELTYPE
BUK3TF3	BUK3TF4		CASE, PERSON, WHOHELP
BUK4CH1	BUK4CH2		CASE, PERSON, PRGID
BUK4CH1	BUK4CH3	CREATE CH05=CH42 IN BUK4CH3	CASE, PERSON, CH05
BUK4KL2	BUK4KL3		CASE, PERSON, EVENT

NOTE: See Appendix A for details on individual subfiles.

Table 3.5

IFLS-CF SPLIT SECTIONS THAT CAN BE JOINED

Files That Can Be Joined		Variables To Merge By
BUKID01	BUKID02A	EA
BKIIS01	BKIIS02	EA
BKPSB01	BKPSB02	FACCODE
BKPSD01	BKPSD02	FACCODE
BKPSE01	BKPSE02	FACCODE
BUKDA01	BUKDA02	FACCODE
BUKDC01	BUKDC02	FACCODE
BUKDE01	BUKDE02	FACCODE
BKBDA01	B KBDA02	FACCODE
BKBDC01	BKBDC02	FACCODE
BKBDE01	BKBDE02	FACCODE
BKPTRB01	BKPTRB02	FACCODE
BUKSB01	BUKSB02	FACCODE
BKSPB01	BKSPB02	FACCODE
BKSAB01	BKSAB02	FACCODE

NOTE: See Appendix B for details on individual subfiles.

WEIGHTING DATA: IFLS-HH

The IFLS household Survey was designed to support a range of analyses based on a smaller, but richly detailed micro-level database covering a wide array of demographic, economic, and health outcomes. The survey was not envisioned as a database to produce national-level or provincial-level estimates of demographic or economic variables. The public use IFLS-HH database does include a series of household and individual analytic weights so that analysts can adjust, when appropriate, for the IFLS household and within-household sampling procedures. The creation of the IFLS-HH weights is discussed in the IFLS-HH *Codebook* and in the *Overview and Field Report*. All weights are inverse probabilities of selection. When using SAS, these weight variables can be used directly in the WEIGHT statement.

Household weights: HHLDWT subfile

The household weights 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 and differential sampling within urban and rural EAs. The household weights are such that when used, the sample of IFLS-HH households will reflect the 1993 distribution of households by urban and rural status within each of and between the 13 Indonesian provinces covered by the IFLS-HH.

Two household weights were created. Both represent the inverse of the probability of selection and both are normalized to sum to the number of households they represent. The household weights are:

HHWT730: for the Book K sample of 7,730 households; to be used when doing tabulations from the Book K (Control Book) data.

HHWT224: for the Book I sample of 7,224 households; to be used when doing tabulations with household-level data from all other Books in the IFLS-HH.

These weight variables can be merged on to any IFLS-HH subfile by using CASE, the household identifier variable.

Individual weights: INDIVWT subfile

There are three basic types of individual weights found in the subfile called INDIVWT: respondent weights, roster weights, and anthropometry weights. All individual weights can be merged on to the desired Book III, IV or V individual-level subfile using the identifier variables CASE and PERSON. The weights can be added onto the household roster data by setting PERSON=AR001A in the subfiles BUKKAR2 and then merging by CASE and

PERSON. Likewise, the weights can be added onto the anthropometry roster by setting PERSON=CA02 in the subfiles BUKCCA2 and merging by CASE and PERSON.

Respondent weights

The respondent weights are designed to adjust for the intra-household sampling scheme used to select Book III, IV and V respondents. There are three versions of the respondent weight provided. These three versions are provided to give the users flexibility in how they wish to use respondent weights. The three respondent weight variables, which are inverses of the probability of selection, 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 with data from Books III, IV and V. The sum of the weight variable ORIGNRWT provides the number of individuals eligible for interview across all 7,224 households with a household roster. If the weight variables, N_RESPWT or ORIGNRWT are used, the user must also interact them with the household weight variable HHWT224 in order to properly adjust for the over-sampling of urban EAs and EAs in smaller provinces and the differential sampling within urban and rural EAs ---N_RESPWT and ORIGNRWT adjust for intra-household sampling only. The sum of the variable ORIGNRWT gives the number of people eligible to e selected as an IFLS respondent across all households.

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.

ROSTERWT:	ratio of the population proportion to the IFLS-HH household roster proportion for a given strata.

The variable ROSTERWT is used when doing tabulations based on all the individuals found in the IFLS household roster data (files BUKKAR2 and BUKKAR3). Because ROSTERWT is based on population distributions, the oversampling issues related to IFLS households do not come into play. ROSTERWT can be used without interacting it with HHWT224.

Anthropometry weights

The anthropometry weights are designed to reflect the intra-household sampling scheme to select those to be measured. Like the respondent weights, there are three versions of the anthropometry weight 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 from the anthropometry data. N_CA_WT and ORIGCAWT, if used, must be interacted with HHWT224 to adjust for the household sampling procedures. The sum of the weight variable ORIGCAWT gives the number of people eligible to be measured across all interviewed IFLS households.

WEIGHTING DATA: IFLS-CF

The IFLS-CF was designed to provide extensive community and facility information to complement the IFLS-HH data. The IFLS-CF was not designed to produce nationally-representative statistics on community and facility distribution or characteristics. The weights are included so that users can adjust for sampling procedures in their analyses. The IFLS-CF database has two basic sets of weights: community weights and facility weights. The creation of the community and facility weights is discussed in the IFLS-CF *Codebook* and in the *Overview and Field Report*. All weights are inverse probabilities of selection. When using SAS, these weight variables can be used directly in the WEIGHT statement.

Community Weights: COMMWT subfile

The community weights 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. When weighted, the IFLS-CF communities reflect the number of EAs in the province/urban-rural strata in which the community lies.

COMMWT: the ratio of the number of actual EAs to the number of sampled EAs.

The COMMWT variable should be used when doing tabulations from the subfiles containing community characteristics, the Book I, Book II and Book PKK subfiles. The COMMWT variable can be merged on to any of the desired community characteristics subfiles by using the identifier variable EA.

Facility Weights: FACWGT subfile

Each facility has its own weight and all facilities are pooled together into one file. Facility weights are designed to approximate the set of facilities of a given type available to an EA. Because there is no information on the actual number of facilities of a given type, the number was estimated using a model that drew on the number of facilities mentioned in the household data and the frequency of those responses.

FACWGT: inverse probability of selection for a given facility

The facility weight variable can be merged onto any facility file by using the facility identifier code variable FACCODE. Since all facility weights are in one subfile, users will only want to keep matches between the given facility subfile data they are using and the weight data subfile. Users should use FACWGT when doing tabulations of facility characteristics.

IV. IDENTIFYING SAMPLES, HOUSEHOLDS AND INDIVIDUALS WITHIN IFLS DATA SUBFILES

This section discusses how to identify or locate the records for different households and types of individuals within the IFLS-HH database and types of facilities within the IFLS-CF and IFLS-HH databases. The type of individuals include the IFLS-HH questionnaire respondents, spouses of respondents, children and other relatives of respondents, and nonrespondents to IFLS household questionnaires. The information presented in this section should help users better understand the discussion on how to link records across files presented in Section V.

HOUSEHOLDS IN IFLS-HH

Households are identified by the variable CASE. All records with the same values of CASE belong to the same household.

IFLS-HH RESPONDENTS

The cover files for Books III, IV, and V contain one record for each respondent to the given questionnaire. In Books III, IV, and V, the variables CASE and PERSON uniquely identify the respondent to a given book. The CASE and PERSON values for the book's respondent are on all subfiles associated with that book. To get demographic characteristics of a respondent, the variable AR001A in the roster files BUKKAR2 and BUKKAR3 corresponds to the variable PERSON. Simply create PERSON=AR001A and then you can use CASE and PERSON from the desired book's cover file to link back to the household roster data. Table 4.1 lists the book cover subfiles and the respondent identifier variables.

Table 4.1

IFLS-HH SUBFILES CONTAINING BOOK COVER INFORMATION

IFLS-HH Book	Subfile Name	Respondent Identifier Variable
Book I	BUKKAR1	RESP1_1
Book II	BUK2UT2	RESP2_1
Book III	BUK3S3A	PERSON
Book IV	BUK4SIV	PERSON
Book V	BUK5SE5	PERSON

The cover files for Books I and II do not have the specific variable PERSON on them. Books 1 and 2 are household-level information files and the main identifier is the household identifier variable (i.e., CASE). The Book I and II cover files do contain variables that

identify which household member answered the questions for the given book. In the Book I cover subfile (BUKKAR1), the variable RESP1_1 contains the household member number for the intended respondent and RESP1_2 contains the identifier for the proxy respondent if the intended respondent could not be interviewed; in the Book II cover subfile (BUK2UT1), the comparable variables are RESP2_1 and RESP2_2. To link these respondents to the household roster data, just follow the same type of procedure as above. In this case, create AR001A equal to the desired RESP variable and merge to the household roster data by CASE and AR001A. To link to the Book III data, just set PERSON equal to the desired RESP variable and merge by CASE and PERSON.

Books I-V also have proxy respondent identifier variables if the individual selected as the respondent did not actually answer the questions in the given questionnaire book. In Books III and IV, proxy respondents were only given a subset of the questionnaire. The IFLS-HH *Codebook* contains a discussion of proxy respondents. Please refer to that document for details.

RESIDENT SPOUSES OF IFLS-HH RESPONDENTS

The variable AR14 (Spouse) on the respondent's BUKKAR2 subfile record gives the person number of the respondent's resident spouse. Because husbands often work outside the village, a woman can be currently married and yet her husband may not be listed in the household roster. In addition, because polygamy exists in Indonesia, if a husband has more than one wife, he may be living elsewhere with one of the other wives, and thus may not appear on the household roster. In addition, polygamy means that a given man may have more than one wife in a household. Users should not be surprised if more than one woman in the household reports the same individual to be her spouse.

The AR14 (Spouse) variable can also be used to identify spouses of other household members listed in the roster data.

RESIDENT PARENTS OF IFLS-HH RESPONDENTS

The codes in the variables AR10 (Father) and AR11 (Mother) on the BUKKAR2 subfile record of the respondent represent the respondent's parents' person numbers from BUKKAR2. The parents listed in AR10 and AR11 are supposed to be the person's biological parents. A code of 51 means the parent is alive and living outside the given household; a code of 52 means that the parent is dead.

The AR10 (Father) and AR11 (Mother) variables can also be used to identify the parents of other household members listed in the household roster data.

For Book III respondents, non-coresident parents are found in Section BA (subfiles BUK3BA1 and BUK3BA2).

RESIDENT CHILDREN OF IFLS-HH RESPONDENTS

All records in BUKKAR2 have the variables AR10 (Father) and AR11 (Mother) which contain the person numbers of the individual's parents if the parents are listed in the household roster for that case. Resident children of the IFLS respondent, then, will have the IFLS respondent's person number listed in one of those fields. Using the AR10/AR11 codes is usually the best way to identify children of an IFLS respondent.

In the case of step or adopted children, step or adoptive parents may not be listed; many times such parents are listed but sometimes not. Users interested in step-children and adopted children as well may refer to the AR02 (relationship to head of household) variable on BUKKAR2. If the IFLS respondent in question is also the household head, step children and adopted children have AR02=4. Unfortunately, if the IFLS respondent of interest *is not* the household head, users cannot so easily locate step/adopted children . Users must compare how the given IFLS respondent is related to the household head and look for relevant relationship codes. For example, if the IFLS respondent is the father of the household head in BUKKAR2, those individuals with AR02 codes of 8 (sibling) and who don't list the given IFLS respondent in the AR10/AR11 codes are most likely step and adopted children of the given IFLS respondent.

The AR10 (Father) and AR11 (Mother) variables can also be used to identify the children of other household members listed in the roster data. If the AR10 (Father) and AR11 (Mother) code equals the household member number of the individual in question, the records that match are the children of that individual.

NON-CORESIDENT PARENTS AND CHILDREN OF IFLS-HH RESPONDENTS

Information on non-coresident parents and children was collected for Book III respondents. Non-coresident parents are found in Section BA (subfiles BUK3BA1 and BUK3BA2). The variable PARENT identifies to which parent the data record belongs (PARENT=1 is father, PARENT=2 is mother).

Non-coresident children are found in Section BA (subfiles BUK3BA8 to BUK3BAB). A complete roster of non-coresident children (plus any children that died in the past year) is found in the subfile BUK3BA8. The variable BA63A is a simple ordinal identifier variable (first child has BA63A=1, second child has BA63A=2, and so on). The full roster of non-coresident children contains information on the the children's gender, ages, marital status,

education, and general information on current location of residence. For a selected set of non-coresident children, more detailed information was collected.

NONRESPONDENTS FOR IFLS-HH

The Book K subfile called BUKKPSK lists all members of the household from which those eligible for interview were selected. The variable PSAWAL provides information as to whether the given individual was eligible to be selected for interview; the variable PSAHIR provides information on who was selected to be interviewed. For the household head and spouse of household head, PSAHIR is blank; for all other selected respondents PSAHIR is greater than zero. By using CASE and AR001A, users can link an individual's household roster record in BUKKAR2 to the individuals BUKKPSK record. In BUKKPSK, the variable that is the counterpart to AR001A is PS00. Users can create AR001A by setting it equal to PS00 in subfile BUKKPSK and then merging records between BUKKAR2 and BUKKPSK using CASE and AR001A.

If the selected individual in BUKKPSK is an adult age 15 or older and does not have a record in the Book III cover file (BUK3S3A), then the individual is a nonrespondent for Book III. If the selected individual is a child age 14 and under and does not have a record in the Book V cover file (BUK5SE5) then the individual is a nonrespondent for Book V. For Book IV, all selected Book III respondents who were ever-married females aged 15-49 were supposed to be given Book IV. Such selected Book III respondents who do not have a record in the Book IV cover file (BUK4SIV) are nonrespondents for Book IV.

Problems existed in the coding of PS00, the person identifier variable in BUKKPSK. Sometimes the order of household members in BUKKPSK was not the same as that in the BUKKAR2 household roster. Efforts were made to make PS00 match the variable AR001A in the BUKKAR2 household roster by matching names between the BUKKPSK and BUKKAR2 subfiles. Some incorrect PS00 values may still remain.

INELIGIBLE RESPONDENTS FOR IFLS-HH

There are cases where an individual was interviewed (i.e., given a Book III or IV or V) but the individual was not actually eligible to be interviewed. Such examples include cases where children selected for Book V were not the offspring of the household head and whose own parents were present in the household (only the head's children age 0-14 and children age 0-14 with no biological parents in the household were eligible to be selected for Book V). To identify Book III, IV or V respondents who should not have been interviewed, users can link the INDIVWT individual weight file to the given book cover by using CASE and PERSON. Any individuals in the book cover files for whom the RESPWT variable is zero is

an ineligible respondent. Depending on a given analysis, users may wish to drop these individuals. If weighted tabulations are done, such individuals are automatically dropped due to their zero weight value.

FACILITIES IN IFLS-HH

In various sections of the IFLS-HH survey, respondents were asked questions about health facilities and schools. In Book I, respondents were asked to give the names of health facilities they tended to use or of which they were acquainted (BUK1PP1); in the household roster section (BUKKAR3), the survey collected information on schools which household members age 25 and under were currently attending. This information (along with family planning clinic information from Book IV (BUK4CX1)) was used as the basis for selecting facilities for the IFLS-CF survey. In Book III, users were asked about schools they attended (BUK3DL2) and health facilities actually used for outpatient and inpatient treatment (BUK3RJ2 and BUK3RN2). Facilities providing family planning services were asked about in Book IV (BUK4CX1, BUK4KL4B). For children in Book V, information on schools and health facilities was collected in the same fashion as in Book III (BUK5DL1, BUK5RJ2, BUK5RN2). In all these files, the variable FACCODE exists.[10]

The fourth digit of FACCODE identifies the type of facility. Below are the codes for each type:

0: Hospital (not interviewed for IFLS-CF)
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

FACCODE is only unique for facilities found in the IFLS-CF data; facilities that were not interviewed have codes that reflect the EA and the type of facility only, with the remainder of the code denoting it to be a non-interviewed facility. When a health facility was not interviewed, the last 3 digits of the facility code are in the 500's and 600's. When the last 3 digits of the health facility code are in the 700's, it means that insufficient information was given in the household survey regarding the name and location of the health facility so that we could not determine whether the facility was among those we interviewed. When the last 3 digits of the health facility code are in the 800's, the facility was misclassified and is listed under the wrong type (i.e., the 4th digit of FACCODE is not correct for the given type of

[10]In subfile BUK4KL4B, the facility identifier variables are FACCODE1, FACCODE4, and FACCODE6.

facility). When possible we have tried to assign misclassifieds the correct 4th digit, however, a few do remain in the household survey data. For schools, when the last 3 digits of FACCODE are less than 100, the school was not interviewed.

FACILITIES IN IFLS-CF

The variable FACCODE is on all IFLS-CF subfiles and it uniquely identifies a given facility. All records with the same FACCODE belong to the same facility. FACCODE can also be used to link information on facility characteristics back to records in the IFLS-HH subfiles that contain FACCODE.

V. LINKING IFLS DATA SUBFILES

Sections of the IFLS questionnaires are stored in separate subfiles for easier handling (e.g., BUK4CH1, BUK4CH2 and BUK4CH3 contain the pregnancy history and infant feeding data; BUK4CX1, BUK4CX2 and BUK4CX3 contain the contraceptive knowledge and current use data; BUK3MG1, BUK3MG2, and BUK3MG3 contain migration histories). These data subfiles can be linked or matched in a number of different ways to produce a variety of analysis files. In this section, we discuss some of the various ways to link the different IFLS data subfiles. The most basic linking involves linking records for a given person. Other linkages we discuss include husbands and wives, and children and parents, and linkages between the IFLS-HH and IFLS-CF data. Finally, we present a general strategy for linking past events, such as a birth with the concurrent events (e.g. home sanitation characteristics.)

LINKING DIFFERENT RECORDS FOR A GIVEN PERSON

The method for linking an individual's records from different subfiles depends on who the individual is: an IFLS respondent, a child of a Book IV respondent, another household member, a non-coresident sibling, or a non-coresident child of an IFLS-HH respondent. Below we discuss how to link records from various IFLS data subfiles for these various types of individuals.

IFLS-HH Respondent

The sequence of the variables CASE and PERSON found on all Book III, IV, and V subfiles uniquely define the IFLS respondent for the given book. To link subfiles within a given Book questionnaire, e.g. Book III, you merge by CASE and PERSON. However, you can only do this if you are linking subfiles that contain just one record per each respondent or if you are linking one file with multiple records per respondent to another with just one record per respondent. You cannot link two subfiles that each have multiple records per respondent by using CASE and PERSON. Linking subfiles with multiple records per respondent is discussed under the subheading for linking past events. To link the IFLS respondent to the household roster data, the variable AR001A in the BUKKAR2 and BUKKAR3 subfiles is the same as PERSON. Simply rename AR001A to person and merge by CASE and PERSON.

As noted in the previous section, the respondents for Books I and II are identified by the variables RESP1_1 and RESP1_2 in subfile BUKKAR1 and variables RESP2_1 and

RESP2_2 in subfile BUK2UT1. To link these respondents to records they may have in Books III and IV and in the household roster, see the discussion in the previous section.

Children of Book IV Respondent

All children of the Book IV respondent, living or dead, are listed in the Book IV pregnancy history data. The subfile BUK4CH1 contains the basic birth outcome information for each child (e.g., pregnancy outcome, sex of child, date of birth, and age of mother at birth). The subfile BUK4CH2 contains information about the birth including whether the child is still alive and where the child currently resides (i.e., inside or outside of the household). To link the information in BUK4CH2 to BUK4CH1, the variables CASE, PERSON and PRGID are used. The variable that shows whether the child lives in the household is CH27A; the household member identifier for a child living in the household can be found in the BUK4CH2 variable CH27B. The values in CH27B link back to the AR001A values in BUKKAR2. To link the birth history data for a child in the household to the child's roster data (files BUKKAR2 and BUKKAR3), users must create an AR001A variable equal to CH27B and then merge by CASE and AR001A. Please note that *the variable PERSON in BUK4CHx subfiles is the household person number of the Book IV respondent* and not the identifier number of the child.

Other Household Members

Within the household roster data (BUKKAR2 and BUKKAR3), CASE and AR001A uniquely define an individual. Information on such individuals is limited and appears only in household roster data and in the Book II household economy file BUK2PH1. The BUK2PH1 subfile contains individual income information for non-Book III respondents (Book III respondent income data is found in BUK3TK2 and BUK3T2B (earned income) and BUK3HI1 and BUK3HI2 (assets and nonearned income)) . The individual identifier variable in BUK2PH1 is PH00--it is the counterpart of AR001A in BUKKAR2/BUKKAR3. To link BUK2PH1 data to BUKKAR2 data for non-Book III respondents, rename PH00 to AR001A and merge by CASE and AR001A.

LINKING HUSBANDS AND WIVES

Users will often want to link the records of the husbands and wives. For example, users may want to link father's education to each child's birth record listed by the Book IV respondent. Such a match requires linking the birth history records with the father's household roster demographic record. However, only the mother's identifiers, CASE and PERSON, appear on the birth history subfiles (BUK4CH1-BUK4CH3). One must first link

the husband's household roster data from subfiles BUKKAR2 and BUKKAR3 to that of his wife and then link that information to the birth history file.

The BUKKAR2 record contains an identifier for the individual's resident spouse (AR14). As a first step, users should merge the BUKKAR2 and BUKKAR3 subfiles to form a single household roster file (see Section III). To link husbands and wives, users can split the merged BUKKAR2/BUKKAR3 file into men and women . In the male file, recode the variable AR14 to equal the value of PERSON, and sort the file by CASE and AR14. The female file must be sorted by CASE and AR14 and is then linked to the male file by CASE and AR14.[11] Women with no husbands in the household will have no match to the male file. The husband's demographic data can now be merged onto any Book IV subfile through the IFLS respondent's identifiers CASE and PERSON, or to any Book III subfile belonging to the woman. To link the wife's data to the husband's Book III records, just reverse the process by recoding AR14 to PERSON in the female file.

As mentioned earlier, polygamy exists in Indonesia. Households with multiple wives do exist in the IFLS data. In some cases, the other wives do not live in the interviewed household but are listed in the husband's Book III marriage history (BUK3KW2). In linking BUK3KW2 marriage records of the husbands to the marriage records of wives (women age 15-49 are in BUK4KW2 and women age 50 and above have marriage records in BUK3KW2), users must be careful to see that the correct marriage records are linked when more than one male marriage record in BUK3KW2 has a marital status code showing "currently married" (KW16=2). In addition, the last marriage chronologically for a male Book III respondent is not necessarily the marriage to his current spouse in the household (if one exists). The male respondent may have taken the co-resident spouse as his first wife and then subsequently married one or more women. The records for those marriages will follow that of the marriage to the co-resident spouse. Therefore, users should exercise some caution in linking marriage records. In cases where male Book III respondents have multiple current marriages, users should match female marriage records (BUK4KW2 for women age 15-49 and BUK3KW2 for women age 50 and over) to male marriage records in BUK3KW3 by using marriage dates and ages as well.

LINKING PARENTS AND CHILDREN

The BUKKAR2 household roster data contains identifiers for the resident mother and father of each household member. The variables AR10 (father) and AR11 (mother) represent

---

[11] Please note that users must rename the variables in the male file before merging with the female file, otherwise users will overwrite the women's variable values with her husbands.

the household member numbers of the individual's parents who are listed on the household roster. The strategy for linking parents and children depends on whether the user wants to add the parents' information to the child's record or information from the child's record to the parents.

Linking parents' records to children's

In this example we will assume that the user wants to link the IFLS-HH respondent's parents' household roster demographic information onto the IFLS-HH respondent's records. Again we suggest merging the BUKKAR2 and BUKKAR3 subfiles to make one household roster file. One method for linking records is to split the merged BUKKAR2/BUKKAR3 data into males and females, and to make separate files containing the roster information for the IFLS-HH respondents. Users can link the BUK3S3A, BUK4SIV and BUK5SE5 files to BUKKAR2 by CASE and PERSON (renaming AR001A to PERSON first) to create those latter files for the Book III, IV and V respondents. Using the female merged BUKKAR2/BUKKAR3 data, recode the variable AR11 to equal the value of PERSON, and rename the relevant roster variables to be added. Sort the data by CASE and AR11 and sort the respondent file by those same three variables. Merge the two files by CASE and AR11 to add the mother's demographic data to the respondent's record (if the respondent's mother is not in the household, no data will be added). Repeat the same process with the male merged BUKKAR2/BUKKAR3 data by reassigning the value of AR10 to be equal to PERSON, then renaming and sorting the data accordingly.

To link any parent in the household roster data to their child, you follow the same approach, but instead of a respondent file, you will have all the individuals in the household roster.

Linking children's records to parents'

Users may wish to attach information about an IFLS-HH respondent's resident children to the respondent's IFLS-HH data subfiles. For example, one may wish to attach the number of adult children living with the Book III respondent to the Book III respondent's health status record. Such a match requires collapsing data about all children in the household roster into one record that can then be linked to the IFLS-HH data files. A suggested strategy here is to take the desired IFLS-HH cover file (e.g., Book III), create a RESPID variable setting it's value to that of PERSON, and then merge this RESPID variables onto the household roster data by CASE (the household identifier). If the variable AR10 or AR11 equals this RESPID, then that household member is the child of the given IFLS-HH respondent. These records can be split out to a separate file for future use or

counters can be turned on to count the number of children of desired characteristics for a given IFLS-HH respondent.

Information about non-coresident children can be linked in a similar manner. The BUK3BA8, BUK3BA9, BUK3BAA and BUK3BAB contain information on non-coresident children. In these files (except for BUK3BA9) records represent individual children of the respondent. Users can collapse the information across records with the respondent's identifying variables (CASE and PERSON), and then match that aggregated data to the IFLS-HH respondent.

Users may want to use just the household roster data to look at other individuals and their resident children. Again the variables AR10 and AR11 give the household member numbers for the parents of the given individual in the household roster. Users can count up the number of household roster records reporting the same AR10 and/or AR11 codes, output those counts to a separate file and then link back to the household roster data to locate the father's and mother's records, by finding the individuals with AR001A values equal to the AR10/AR11 codes on the children's records.

LINKING TO IFLS-HH ANTHROPOMETRY ROSTER

The variable CA02 in the anthropometry roster file BUKCCA2 corresponds to the AR001A variable in the household roster subfile BUKKAR2 (and to PERSON as well). To link household roster members to their anthropometry records, create AR001A equal to CA02 and link by CASE and AR001A. Not all household members were measured, so there will not be a one-to-one match between the BUKKAR2 and BUKCCA2 subfiles.

To obtain height and weight measurements, users must first link the BUKCCA2 and BUKCCA3 subfiles by using CASE and LINE_CA as noted earlier in Section III. Once those two files have been joined, users can then link the household roster data to that combined anthropometry file as discussed above. Likewise, users can link IFLS-HH respondents to that combined anthropometry file by setting PERSON equal to CA02 in the anthropometry file and merging by CASE and PERSON to the IFLS-HH respondent file.

LINKING IFLS-HH AND IFLS-CF

Community data can be linked to the IFLS-HH subfiles by using the variable EA. EA, while not specifically on IFLS-HH subfiles, can be easily created since the EA code is a component of the variable CASE. Users can create a 9-digit character format variable counterpart to CASE and then right justify the characters in the field. The 5th, 6th, and 7th digits of the 9-digit field are the EA. Users can also merge on the variable SC07 from

BUKKSC1 to any IFLS-HH subfile using the variable CASE for linking files. SC07 is the same as EA, so just set EA=SC07 after the merge.

As noted earlier, the variable FACCODE in the IFLS-HH subfiles that contain references to health facilities and schools is used to link to the IFLS-CF subfiles. For IFLS-CF subfiles that contain multiple records per facility, users must first collapse that data into a single record per facility before merging such IFLS-CF facility data to the desired IFLS-HH subfile.

LINKING PAST EVENTS IN THE IFLS-HH

A common form of linkage with retrospective/longitudinal data is the matching of one event with other events from the same time period. For example, users analyzing infant mortality may want to know the household's water and sanitation conditions at the time of each child's birth. Such linkages often require linking files that each contain multiple records per person, i.e., a record for each occurrence of a given event type (in this example, births and changes of residence). A simple merging of records by the respondent's identifiers CASE and PERSON will not produce a file where concurrent events are linked.

We propose here a general strategy for linking concurrent events. Users may develop their own strategies in accordance with their needs. The following suggested linkage strategy merely provides one example of how to link concurrent events. For illustrative purposes, we will use the above example (matching births to household water and sanitation conditions) in discussing the linkage strategy.

The strategy's objective is to locate the residence in which the woman lived when she gave birth to each child. The migration history, which contains the water and sanitation characteristics, must be linked to the pregnancy history using move dates and birth dates. Users must remember that because dates are not always available, ages may have to be used. Since births are the main event of interest, we suggest merging onto the birth file a record containing all the move dates and ages and the attendant migration record identifiers for a given woman. To create such a record, users must convert selected migration history information from multiple records per woman to one record per woman. The one-record-per-woman file would contain the dates and ages of each move, the sequence number for each move, and the woman's identifiers CASE and PERSON. This record is then merged on to the pregnancy history using CASE and PERSON, adding the same basic migration data to each pregnancy record. Once attached, the user then compares the birthdate of the child to the array of move dates (or age of mother at birth to an array of ages at move, depending on what is available) to locate the last move occurring before the given birth. We suggest that users convert month/year dates into standard month dates (i.e., the number of months from

Jan 1, 1900--(year*12) + month) for easier comparison. When only ages of moves or ages at birth are available, users may compare ages instead of dates, or may choose to impute an event date based on the respondent's birthdate and her/his age at the event[12]. Once a match had been made, the user then keeps the sequence number of that move and drops all the other migration information added. The pregnancy history data can then be sorted by CASE, PERSON and the migration history sequence number (MOVENUM) and linked to the migration history to add the desired household characteristics.

LINKING IFLS-CF DATA TO PAST EVENTS IN THE IFLS-HH

In the IFLS-CF questionnaire, some historical information was collected at the community level and at the facility level. Trying to use this information to attach to events occurring other than recently before the IFLS-HH interview can be difficult if not impossible in some cases. If the individual has not moved since the time of the event in question (i.e., at interview the individual is living in the same village as at the time of the event), users can use the event date information and compare it to data information in the relevant historical data in the IFLS-CF. The relationship of the event date to the IFLS-CF historical date information provides information on what community/facility characteristics may have been relevant at the time of the event. Users must also be aware that just because a facility has the same EA value as the IFLS-HH household, the facility may not be in the same area as the IFLS-HH household. Also, the IFLS-CF is not an exhaustive list of available facilities, so using the facility data to assess availability of services in the past may not give an accurate reflection of the true situation. The community data subfiles BUKII01, BUKII02, BUKIJ01, BUKIJ02, BUKPI01, BUKPI02, BUKPJ01, and BUKPJ02 provide historical information on schools and health facilities in the community. The information in these files should provide better data on the availability of services in the EA over time.

[12]For example, a woman born in August 1965 reports a move at age 20. The move, then, occurred between August 1985 and July 1986. If one assumes the move occurred mid-way in that interval, an imputed move date of February 1986 could be used.

VI. CONNECTIONS AMONG THE IFLS SUBFILES

A feature of the IFLS database is the interrelatedness of the various data subfiles. The same information may appear in more than one subfile, and events or individuals mentioned in one file may appear with more detail in another subfile. In addition, identical information collected for one set of individuals in one subfile may appear in another subfile for another set of individuals (e.g., ever-married women age 15-49 have their marital history data in Book IV instead of Book III). In such cases, users may want to pool information from those different sources. Users should be aware of these relationships for several reasons. First, users can avoid double-counting items which may appear in several parts of a questionnaire. Second, users can augment information in one file with data from another. Third, responses that may seem odd or confusing in one file can often be clarified by information from other files: the response may be substantiated or refuted by the corroborating data. In this section, we discuss some of the major connections between subfiles; however, we cannot present all possible connections. Thus, we highly recommend that users thoroughly review the questionnaires and interviewer instructions. Below we first discuss the interrelatedness of IFLS-HH subfiles, followed by IFLS-CF subfiles, and finally touch on the interrelatedness of some IFLS-HH and IFLS-CF subfiles.

IFLS-HH: INFORMATION USERS MAY WANT TO POOL

The discussion below lists subfiles where the same information is collected on different sets of individuals and is thus found in different subfiles. Depending on the type of analysis, users may wish to pool records from groups of subfiles covering the same type of information. Users must remember to rename variables to some common set since variable names differ across subfiles even if the type of information does not.

Current Earned Income Data

Information on current earned income is found in several subfiles. For non-Book III respondents age 10 and over, current earned income data is found in BUK2PH1. For Book III respondents, current earned income information is found in the BUK3TK2 (current primary job) and BUK3T2B (current secondary job) subfiles.

Assets and Non-earned Income Data

Household assets are found in the subfile BUK2HR1 and non-earned income for the household is found in BUK2HR2. In Book III, all respondents were asked about their assets

and non-earned income sources. Those data are in BUK3HI1 (assets) and BUK3HI2 (other income). If the Book III respondent was also the respondent to Book II (i.e., he/she answered the household assets and non-earned income), section HI of Book III was not administered. Thus the Book III respondent who was also the Book II respondent has no data in the subfiles BUK3HI1 and BUK3HI2. For Book III respondents who were not the Book II respondent, some of the information included in BUK3HI1 and BUK3HI2 may also appear in BUK2HR1 and BUK2HR2. In essence, the Book II subfiles should include the entire household; the Book III data should give more specifics on who owns the assets/receives non-earned income in the household. Users should be aware of possible double counting.

If users want to examine asset ownership and sources of other income for all Book III respondents, users must pool the BUK2HR1 and BUK2HR2 data for those Book III respondents who were also the Book II respondent with the BUK3HI1 and BUK3HI2 data of the remaining Book III respondents.

Education Data

Detailed information on respondents' education is found in several subfiles. The education information collected in Book III and Book V are very similar. Users may wish to pool the Book III (adults age 15 and above) and Book V (children age 6-14) data to have a file with education-related information covering individuals age 6 and above. The Book III education files are BUK3DL1, BUK3DL2, and BUK3DL3; the Book V files are BUK5DL1 and BUK5DL2.

Acute Morbidity and Self-Treatment

The acute morbidity and self-treatment information collected in Book III and Book V are the same, with Book V including a few childhood ailments not typically found among adults. Users may wish to pool the Book III and Book V data to have a morbidity file that individuals age 6 and above, and likewise for self-treatment. The Book III acute morbidity file is BUK3MA1 and the Book V file is BUK5MAA; the Book III self-treatment file is BUK3PS1 and the Book V counterpart is BUK5PSA.

Outpatient and Inpatient Utilization

The outpatient and inpatient information collected in Book III and Book V are identical. Users may wish to pool the Book III and Book V data to have a file with utilization information covering individuals age 6 and above. The Book III outpatient care files are BUK3RJ1 and BUK3RJ2, the inpatient files are BUK3RN1 and BUK3RN2; the comparable Book V files are BUK5RJ1, BUK5RJ2, BUK5RN1, and BUK5RN2.

Pregnancy Summaries

Book III collected pregnancy summary information in BUK3BR1 for female respondents *except* for those ever-married women age 15 to 49. Those women were asked the pregnancy summary questions in Book IV and their data are found in BUK4BR1. Users may wish to pool the pregnancy summary files to get a file that covers all women age 15 and over.

Marital Histories and Desire for More Children

Book III collected marital history information in BUK3KW1 and BUK3KW2 for all respondents *except* for ever-married women age 15 to 49. Those women were asked about their marriages in Book IV as prelude to the detailed pregnancy history about to be collected. The marriage history data for ever-married women age 15-49 are in BUK4KW1 and BUK4KW2.

Additionally, married men were asked in Book III about their desires for more children. Those questions are in BUK3KW3. The counterpart for their wives is in BUK4KW3 for wives age 15-49 and in BUK3KW3 for wives age 50 and over.

IFLS-HH: INFORMATION THAT CAN BE COMPARED ACROSS SUBFILES

The discussion below lists subfiles where similar information is collected across sections of the IFLS-HH questionnaire. Depending on the type of analysis, users may want to compare the information reported in these different sections.

Educational Status

In the household roster, information was collected on each household member's educational status and level of schooling (BUKKAR3). For those currently attending school and age 25 and under, information on the school attended was collected. In Book III and in Book V, the IFLS-HH respondents were asked again about their educational status and level of schooling, and about the schools they attended (BUK3DL2 and BUK5DL1). Users may want to compare the education information supplied in the household roster by the Book I respondents about the IFLS-HH respondents to the education information collected directly from the respondents themselves.

Educational Expenses

Book I collected information on household expenses on education during the past year (BUK1KS3B). In the Book III and Book V education sections (BUK3DL3 and BUK5DL2), information on the respondent's education expenses was collected if the respondent attended school in the past year. While the sum of Book III and Book V education expenses should not

be expected to match those in Book I, one might expect that the sum in Book I should not be less than the sum of the Book III and V amounts. Some assessment of the quality of reporting education expenses may be able to be done by comparing the Book I, Book III and Book V responses.

Users should note that education expenses information in Book I is in actual rupiah amounts while the education expenses in Book III and Book V are rounded to the nearest thousand rupiah. in addition, users should note which amounts are reported as monthly and which are annual in Books I, III and V before comparing amounts.

Transfers Between Book III Respondents and Parents, Siblings and Children

Book III respondents were asked many questions about non-coresident parents, siblings and children. The following Book III BA section subfiles contain information on transfers between the respondent and these non-coresident relatives: BUK3BA2, BUK3BA6, BUK3BA7A, BUK3BAA, and BUK3BAB. Some Book III respondents were not administered Section BA (non-coresident family roster and transfers): those who were not the head of household, the spouse of the head, the selected senior householder, or the spouse of the senior householder. For Book III respondents who did not answer Section BA, the same type of transfer information was collected in the first half of Section TF of Book III. The subfiles BUK3TF1 and BUK3TF2 relate to transfers to and from non-coresident parents, siblings and children. Users may wish to pool the transfer information found across all the subfiles above to construct transfer files that contains all Book III respondents.

Children of Book IV Respondent

Information on the children of the Book IV respondent can be found in several places. In Book IV, the pregnancy history subfiles (BUK4CH1-BUK4CH3) provide information on birth dates, sex of child, alive/dead status, and living inside/outside the household. The Book IV respondent's answers to the non-coresident children section of Book III refer to the same children listed in the pregnancy history (and possibly more if there are adopted or step children). Those Book III subfiles are BUK3BA8, BUK3BA9, BUK3BAA, and BUK3BAB. The household roster has information on the Book IV respondent's children living in the household with her (BUKKAR2 and BUKKAR3). In the BUK4CH2 subfile, the variable CH25 says whether the child is still alive, the variable CH27A says if the child is inside or outside the current household, and the variable CH27B is the household member number of the child if the child lives in the current household. There is no direct link between the children listed in the Book IV pregnancy history and the non-coresident children covered in Book III. Users must use sex and birth date information to try and link specific children.

Contraceptive Calendar and Fertility-Related Events

The contraceptive calendar in Book IV marriage, pregnancy, and contraceptive use can cover at most the 10 years prior to interview (for some women the period is less). Dates for marriages and pregnancies, and types of contraception used can be compared to the marriage and pregnancy histories and to the contraceptive knowledge and current use information in Book IV. The calendar data is found in two subfiles, BUK4KL2 and BUK4KL3, where the first subfile contains the period January 1984 to December 1988 and the second has January 1989 to February 1994. Each record in those files represents an event (marriage, pregnancy, etc.) and the position of a letter (e.g., A, B, X, H) in the 60-character alphabetic string gives the date and type of each occurrence of the event during that 5-year period. Depending on the desired analysis, users may want to compare the event dates implied in the BUK4KL2 and BUK4KL3 files with those in the BUK4KW2 (marriage history), BUK4CH1-BUK4CH3 (pregnancy history), and the BUK4CX1-BUK4CX3 (contraceptive knowledge and use) subfiles.

Related Information About Husbands and Wives Across Subfiles

A great strength of the IFLS database is the ability to evaluate the responses of husbands and wives against each other. Here are but a couple of examples. Users can check the wife's marital history against her husband's, not only to check agreement of marriage dates, but also to check on whether the husband has other existing marriages. These other marriages may affect the time the husband spends with the respondent. Similarly, users can compare the migration histories to check dates of moves occurring after the couple's marriage. Husbands may be seen moving to a new district and at the same time the wife moves back to her home district, or husbands move and wives follow later. The earlier discussion of linking husbands and wives and linking past events provide insights as how to make such comparisons.

IFLS-CF: INFORMATION USERS MAY WANT TO POOL

The discussion below lists IFLS-CF subfiles where identical information is collected among questionnaires for different facility types. Depending on the type of analysis, users may wish to pool records from the different facility subfiles. Because the variable names differ across facility types even though the information is the of the same type, users must remember to rename variables to some common set of names before pooling.

Availability of Food Stuffs

The respondent to the questionnaire administered to the head of the PKK group (Association of Family Activities) was asked about the availability and prices of numerous food items (BUKPH01). Those same foodstuff questions were asked at the Community Health Posts and FP Posts (*posyandu*/PPKBD) facilities interviewed in the same EA (BKPPKH01). The questions about prices were asked in both books in order to get a broader picture of the prices faced by residents of those areas. In 163 EAs, reports on prices were collected from both sources, with 87 percent of those EAs having four to five individual reports on prices. When developing price indices or examining prices within and across EAs, users can pool information from both sources to get a complete picture of prices, especially since some prices may not be available across all sources. Users should check the reports given for each item across all reports (i.e., those from Book PKK and all those from Book PPKBD). Wide disparity in reported unit prices for a given item may indicate errors in reporting by respondents.

Health Facility Vignettes

In the questionnaires administered at government health centers (*puskesmas*), private physician offices, private clinics, and offices of nurses, midwives, and paramedics, a set of vignettes were included. Respondents at each facility were asked about how they would handle several types of patient problems. The questions are the same across all facility types. In pooling these files, users should create a categorical variable that identifies the type of health facility. The family planning vignettes are found in the subfiles BKPSH01, BKPSH02, BUKDG01, BUKDG02, BKBDG01, and BKBDG02. The pregnancy exam vignettes are in subfiles BKPSI01, BUKDH01, and BKBDH01. The cough and fever vignettes are in subfiles BKPSJ01, BUKDI01, and BKBDI01. The last vignettes, vomitting and diarrhea, are in subfiles BKPSK01, BUKDJ01, and BKBDJ01.

School Characteristics

The questionnaires administered to primary schools, junior secondary schools, and senior secondary schools were identical. Users may pool all subfiles of a given type to form one subfile containing all school types. In doing so, users must remember to create a common set of variable names and to create an categorical variable that identifies each type of school.

IFLS-CF: INFORMATION THAT CAN BE COMPARED ACROSS SUBFILES

The discussion below lists IFLS-CF subfiles where similar information is collected across sections of the questionnaire. Depending on the type of analysis, users may compare the information reported in these different sections.

History of Availability of Schools and Health Facilities

The village head was asked questions about the history of the availability of schools and health facilities within the community. In addition, the questionnaire administered to the head of the PKK group (Association of Family Activities) was asked the same questions. The village heads' responses are found in the subfiles BUKII01 and BUKII02 for schools and BUKIJ01 and BUKIJ02 for health facilities. The comparable subfiles from the PKK questionnaire are BUKPII01, BUKPII02, BUKPJ01, and BUKPJ02. Users may wish to compare and pool the information collected from the two different sources.

Village Housing

In Book II, village leaders were asked questions about housing characteristics in the village (BKIIS02). Interviewers were also asked to fill out a section on "Direct Observations" about the village (BKIIOL1). The overlap in the questions asked in these two sections could be used for checks on consistency.

INFORMATION THAT CAN BE COMPARED BETWEEN IFLS-HH AND IFLS-CF SUBFILES

The possibilities for comparison are more limited between the IFLS-HH and IFLS-CF than within the IFLS-HH itself. The most obvious areas of possible comparison are prices of food stuffs and school characteristics. Book I of the IFLS-HH collected household consumption information, one of which was food items (BUK1KS1). Book PKK and Book PPKBD both have information on food availability and prices in the community (BUKPH01 and BKPPKH01). Users may find the community-level information useful in examining odd prices implied in the household consumption data. The Book III and Book V education sections collected some information on school characteristics (BUK3DL2 and BUK5DL1). If the school mentioned in Book III or Book V is among those interviewed in the IFLS-CF, users could compare responses to questions that appear in both surveys (e.g., type of flooring). However, the overlap of questions is pretty limited.

In Section KR of Book I (BUK1KR1) of the household data, respondents provide information about their housing characteristics. Comparisons can be made between housing types of IFLS-HH respondents and housing types as described by village leaders (BKIIS02).

Another area for possible comparison is the section in the IFLS-CF on major events in the village (BUKIE02B) and questions in the IFLS-HH about economic shocks to the household (BUK2GE1) and reasons for moving (BUK3MG2). However, the usefulness of such a link may be limited here as well.

Appendix A

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 Evalualtion (CPI1-CPI3)	household		7216

[a] Used in addition to CASE to uniquely define a record.

Appendix A (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 (CPII1-CPII3)	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.

Appendix A (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)	mthly/yearly	PERIOD	19524
BUK3T2B	Current Job (Secondary) (TK18-TK27)	mthly/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.

Appendix A (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.

Appendix A (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.

Appendix A (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				
HHLDWT	Household Sampling Weights	household	CASE	7730
INDIVWT	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

Appendix B

IFLS-CF SUBFILES AND SAMPLE SIZES

File Name	File Description (variables in file)	Unit of Observation	Record Id[a]	# of Records
SAMPLE				
SAMPLFP1	IFLS-CF Books Completed	EA		321
BOOK I				
BUKILK1	Control Sheet for Book I	EA		312
BUKIA01A	Transportation to nearest item (BA0-BA5)	location	PLACE	2496
BUKIA01B	Transportation (BA6-BA10)	EA		312
BUKIB01	Availability of Electricity by Provider (BB1-BB7)	provider	ELECTYPE	1747
BUKIC01	Water Sources and Sanitation (BC1-BC21)	EA		311
BUKID01	Agriculture and Industry (BD1-BD10)	EA		312
BUKID02A	Agriculture and Industry (BD11-BD13)	EA		312
BUKID02B	Agriculture and Industry: programs (BD14-BD18)	program	BD14	199
BUKID03	Agriculture and Industry: wages (BD19-BD27)	EA		312
BUKID04	Agriculture and Industry: factories (BD28-BD31)	factory	BD29	440
BUKID05	Agriculture and Industry: cottage industry (BD28-BD37)	cottage industry	BD33	640
BUKIE01	Record of Village Name Changes (BE1-BE6)	name change	BE2	326
BUKIE02A	Climate (BE7-BE9)	EA		312
BUKIE02B	Important Events in the Village (BE10-BE11)	event	BE10	931
BUKIF01	Migration (BF1-BF8)	EA		312
BUKIG01	Credit (BG1-BG7)	institution	BG2	580
BUKII01	History of School Existence (BI1-BI10)	school	BI1	2459
BUKII02	History of School Types (BI11-BI16)	school type	BSCHTYPE	936
BUKIJ01	History of Health Facilities (BJ1-BJ8)	facility	BFACTYPE	5616
BUKIJ02	History of Health Facilities (BJ9-BJ25)	EA		312
BUKIK01	Identity of Respondents (BK1-BK8)	respondent		742
BOOK II				
BKIILK1	Control Sheet for Book II	EA		312
BKIIS01	Land Use (CS1-CS9)	EA		312
BKIIS02	Employment and Housing Conditions (CS10-CS18)	EA		312
BKIIOL1	Direct Observation: Cleanliness,welfare (OL1-OL25)	EA		312

[a] Used in addition to EA to define a record type. Number of records vary due to the unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Appendix B (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id[b]	# of Records
BOOK PKK				
BUKPLK1	Control Sheet for Book PKK	EA		312
BUKPI01	History of School Existence (DI1-DI10)	school	DI1	2336
BUKPI02	History of School Types (DI11-DI16)	school type	DSCHTYPE	936
BUKPJ01	History of Health Facilities (DJ1-DJ8)	facility	DFACTYPE	2527
BUKPJ02	History of Health Facilities (DJ9-DJ25)	EA		312
BUKPH01	Availability of Food Stuffs (DH1-DH3)	food item	DHTYPE	5616
PUSKESMAS				
BKPSLK1	Control Sheet (JLK01-JLK14)			993
BKPSA01	Head of Puskesmas--Part 1 (EA1-EA21)	facility		993
BKPSA02	Head of Puskesmas--Part 2 (Medicines) (EA23-EA24)	medicine	EMEDTYPE	16881
BKPSA03	Head of Puskesmas--Part 3 (EA25-EA26)	facility		993
BKPSB01	Development of Puskesmas--Part 1 (EB1-EB4)	facility		992
BKPSB02	Development of Puskesmas--Part 2 (EB6-EB20)	facility		992
BKPSC01A	Activities of Puskesmas--Part 1A (EC2-EC4)	facility		993
BKPSC01B	Activities of Puskesmas--Part 1B (EC6-EC9)	type of service	ESERVICE	33762
BKPSC01C	Activities of Puskesmas--Part 1C (EC11-EC13)	facility		993
BKPSC02	Activites of Puskesmas--Part 2 (EC35-EC38)	facility		993
BKPSD01	Puskesmas Employees--Government Employees (ED1-ED5)	facility		993
BKPSD02	Puskesmas Employees--Renumerated Employees (ED7-ED8)	facility		993
BKPSD03	Puskesmas Employees--Employee List (ED9-ED14)	employee	EMPID	5887
BKPSE01	Health Instruments--Part 1 (EE1-EE4)	facility		993
BKPSE02	Health Instruments--Part 2 (EE6-EE7)	facility		993
BKPSF01A	Direct Observation--Part 1A (Rooms) (EF1-EF23)	facility		993
BKPSF01B	Direct Observation--Part 1B (Pharmacy) (EF25-EF29)	medicine	EMEDTYPE	18867
BKPSF02	Direct Observation--Part 2 (EF30-EF45)	facility		993
BKPSG01A	KB Service Unit--Part A (EG1-EG2)	facility		993
BKPSG01B	KB Service Unit--Part B (EG3-EG10)	employee	EMPID	2922
BKPSG01C	KB Service Unit--Part C (EG12-EG15)	facility		993
BKPSG02	Presence of Hypothetical Cases (EGKH1-EGKK)	facility		993
BKPSH01	KB--IUD Case (EH1-EH7)	facility		993
BKPSH02	KB--Pill Case (EH8-EH14)	facility		993
BKPSI01	Pregnancy Examination (EI1-EI16)	facility		993
BKPSJ01	Cough and Fever Cases (EJ1-EJ5)	facility		993
BKPSK01	Vomitting and Diarrhea Case	facility		993

[b] For PKK: Used in addition to EA to define a record type. For PUSKESMAS: Used in addition to EA and FACCODE to define a record type. Number of records vary due to the unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Appendix B (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id[b]	# of Records
DR/CLINIC				
BUKDLK1	Control Sheet (FLK1-FLK19)	facility		549
BUKDA01	General--Part 1 (FA1-FA16)	facility		549
BUKDA02	General--Part 2 (FA17-FA22)	facility		549
BUKDB01	Practice Activities--Hours of Operation (FB2-FB3)	session	FSESSION	1070
BUKDB02	Practice Activities--Services (FB5-FB8)	type of service	FSERVICE	20313
BUKDB03	Practice Activities (FB10-FB20)	facility		549
BUKDC01	Health Instruments--Part 1 (FC1-FC3)	facility		549
BUKDC02	Health Instruments--Part 2 (FC5-FC6)	facility		549
BUKDD01A	Stocks of Medicines (FD1-FD9)	medicine type	FMEDTYPE	7125
BUKDD01B	Stocks of Medicines (FD10-FD11)	facility		549
BUKDE01	Direct Observation--Part 1 (FE1-FE18)	facility		549
BUKDE02	Direct Observation--Part 2 (FE19-FE21)	facility		549
BUKDF01	Special FP Services (FF1-FF9)	facility		549
BUKDF02	Case illustrations done	facility		549
BUKDG01	FP--IUD Case 9FG1-FG7)	facility		549
BUKDG02	FP-Pill Case (FG8-FG14)	facility		549
BUKDH01	Pregnancy Examination (FH1-FH16)	facility		549
BUKDI01	Cough and Fever Cases (FI1-FI7)	facility		549
BUKDJ01	Vomitting and Diarrhea Case (FJ1-FJ8)	facility		549

[b] Used in addition to EA and FACCODE to define a record type. Number of records vary due to the unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Appendix B (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id[b]	# of Records
NURSE/ MIDWIFE/ PARAMEDIC				
BKBDLK1	Control Sheet (GLK1-GLK19)	facility		892
BKBDA01	General--Part 1 (GA1-GA19)	facility		892
BKBDA02	General--Part 2 (GA20-GA25)	facility		892
BKBDB01	Practice Activities--Hours of Operation (GB2-GB3)	session	GSESSION	1766
BKBDB02	Practice Activities--Services (GB5-GB8)	type of service	GSERVICE	33004
BKBDB03	Practice Activities (GB10-GB20)	facility		892
BKBDC01	Health Instruments--Part 1 (GC1-GC3)	facility		892
BKBDC02	Health Instruments--Part 2 (GC5-GC6)	facility		892
BKBDD01A	Stocks of Medicines (GD1-GD9)	type of medicine	GMEDTYPE	14348
BKBDD01B	Stocks of Medicines (GD10-GD11)	facility		892
BKBDE01	Direct Observation--Part 1 (GE1-GE18)	facility		892
BKBDE02	Direct Observation--Vaccines (GE19-GE21)	facility		892
BKBDF01	Special FP Services (GF1-GF9)	facility		892
BKBDF02	Case illustrations done	facility		892
BKBDG01	FP--IUD Case (GG1-GG7)	facility		892
BKBDG02	FP--Pill Case (GG8-GG14)	facility		892
BKBDH01	Pregnancy Examination (GH1-GH16)	facility		892
BKBDI01	Cough and Fever Cases (GI1-GI7)	facility		892
BKBDJ01	Vomitting and Diarrhea Case (GJ1-GJ8)	facility		892
POSYANDU				
BKPPKLK1	Control Sheet	facility		899
BKPPKA01	General (IA1-IA13)	facility		899
BKPPKB01	Services (IB1-IB10)	facility		899
BKPPKC01	Manpower--Part 1	employee	EMPID	4313
BKPPKC02	Manpower--Part 2 (IC5-IC8)	facility		899
BKPPKD01	Health Instruments (ID1-ID5)	instrument	IDITEM	10788
BKPPKH01	Availability of Food Stuffs (IH1-IH3)	food stuff	IHTYPE	16182

[b] Used in addition to EA and FACCODE to define a record type. Number of records vary due to the unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Appendix B (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id[b]	# of Records
TRAD HEALER				
BKPTRLK1	Control Sheet (HLK1-HLK18)	facility		624
BKPTRA01	General (HA1-HA15)	facility		624
BKPTRB01	Practice Activities--Part 1 (HB2-HB5)	facility		624
BKPTRB02	Practice Activities--Part 2 (HB6-HB24)	facility		624
BKPTRC01	Traditional Midwife (HC1-HC14)	facility		624
PRIMARY				
BUKSLK1	Control Sheet (JLK01-JLK14)	school		944
BUKSA01	School Principal 9JA1-JA15, J01-J04)	school		944
BUKSB01	School--Part 1 (JB1-JB14)	school		944
BUKSB02	School--Part 2 (JB15-JB24)	school		944
BUKSC01	Teacher (JC1-JC26, JOB1-JOB4)	teacher type	TEACHTYP	1888
BUKSD01	Direct Observation in Grade VI Classroom (JD1-JD14)	school		944
BUKSE01A	School Statistics 1992-1993--Part 1A (number of classes) (JE1)	school		944
BUKSE01B	School Statistics 1992-1993--Part 1B (number of pupils/teachers) (JE2)	sex group	SEX	1888
BUKSE02	EBTANAS Scores (JE1-JE5)	test type	TESTTYPE	1888
BUKSF01	School Revenue (JF2)	school		944
JRH				
BKSPLK1	Control Sheet (KLK01-KLK14)	school		900
BKSPA01	School Principal (KA1-KA15, K01-K04)	school		900
BKSPB01	School--Part 1 (KB1-KB14)	school		900
BKSPB02	School--Part 2 (KB15-KB24)	school		900
BKSPC01	Teacher (KC1-KC26, KOB1-KOB4)	teacher type	TEACHTYP	1800
BKSPD01	Direct Observation in Grade III Classroom (KD1-KD14)	school		900
BKSPE01A	School Statistics 1992-1993--Part 1A (number of classes) (KE1)	school		900
BKSPE01B	School Statistics 1992-1993--Part 1B (number of pupils/teachers) (KE2)	sex group	SEX	1800
BKSPE02	EBTANAS Scores (KE1-KE5)	test type	TESTTYPE	1800
BKSPF01	School Revenue (KF2)	school		900

[b] Used in addition to EA and FACCODE to define a record type. Number of records vary due to the unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Appendix B (continued)

File Name	File Description (variables in file)	Unit of Observation	Record Id[b]	# of Records
SRH				
BKSALK1	Control Sheet (LLK01-LLK14)	school		584
BKSAA01	School Principal (LA1-LA15, L01-L04)	school		584
BKSAB01	School--Part 1 (LB1-LB14)	school		584
BKSAB02	School--Part 2 (LB15-LB24)	school		584
BKSAC01	Teacher (LC1-LC26, LOB1-LOB4)	teacher type	TEACHTYP	1168
BKSAD01	Direct Observation in Grade III Classroom (LD1-LD14)	school		584
BKSAE01A	School Statistics 1992-1993--Part 1A (number of classes) (LE1)	school		584
BKSAE01B	School Statistics 1992-1993--Part 1B (number of pupils/teachers) (LE2)	sex group	SEX	1168
BKSAE02	EBTANAS Scores (LE1-LE5)	test type	TESTTYPE	1168
BKSAF01	School Revenue (LF2)	school		584
WEIGHTS				
COMMWT	Community Weights	EA	EA	321
FACWT	Facility Weights	facility	FACCODE	6385

[b] Used in addition to EA and FACCODE to define a record type. Number of records vary due to the unit of observation, skip patterns, and non-response to subsections of the questionnaires.

Appendix C

ORGANIZATION OF THE INDONESIAN HEALTH SECTOR

In this Appendix, a brief description of the Indonesian health care system is presented, covering the health facilities found in the IFLS-CF and IFLS-HH data. The intent is to merely to provide a general outline for those users not familiar with the structure of the Indonesian health sector.

GOVERNMENT HEALTH SECTOR

The organization of the health sector in Indonesia is complex. Four agencies are responsible for health and family planning service delivery: the Ministry of Health (MOH), the National Family Planning Coordinating Board (BKKBN), the Ministry of Home Affairs, and the Ministry for Population and the Environment. At the central level BKKBN and the MOH formulate most of the family planning and health policies, respectively.[1] The Ministry for Population and the Environment coordinates the policies at the central level, but is not active at lower levels. The Ministry of Home Affairs has little role at the central level but staff at lower levels administer programs at the provincial, district, subdistrict, and village levels.

The Ministry of Health and BKKBN are active at all administrative levels and BKKBN has its own infrastructure at each level. At the provincial level health and family planning offices are responsible to Ministry of Home Affairs officials at that level (e.g.. the head of the provincial health office reports to governor of the province). One BKKBN office coordinates family planning activities for the province. Two offices coordinate health activities at the provincial level. The provincial health office is a unit of the Ministry of Home Affairs and is responsible for the delivery of all health and family planning services. The province office of the Ministry of Health, which represents the central Ministry of Health, supervises and coordinates the activities of the provincial health office but has no operational authority. (unless there is an epidemic or natural disaster).

Organization at the district level parallels that at the provincial level. The district health office of the Department of Home Affairs is responsible for implementing health and family planning services. The MOH district office coordinates health activities and the BKKBN district office coordinates family planning activities.

The health center (*puskesmas*) is the implementing unit of the Ministry of Health within the subdistrict, but its staff are employees of the Ministry of Home Affairs. The *puskesmas* budget

---

[1] Broad health policies to the year 2000 are set forth in a 1982 document called The National Health System. Within these guidelines, specific policies and numerical targets are set every five years (as part of the five year development plan).

comes from a number of sources and administrative levels: the Ministries of Home Affairs and Health, BKKBN, and the presidential program INPRES. BKKBN employs staff at the subdistrict level. The staff sometimes have their own offices and sometimes work out of the *puskesmas.*

Hospitals

The Ministry of Health operates two basic types of fixed facilities: hospitals and health centers. Hospitals are stratified into four classes, A through D, Class A hospitals are teaching hospitals. Class B hospitals offer ten specialties, while Class C hospitals offer specialties in general medicine, surgery, pediatrics, and obstetrics and gynecology. Class D hospitals have no specialists, only general practitioners, nor do they have x-ray equipment, operating theaters, or laboratory equipment.[2]

Other ministries operate hospitals as well (Brotowasisto et al., 1988). In 1986, of the 63,155 government hospital beds (1986), 13% were operated by the Ministry of Health, 55% were operated by provincial and local government, 18% were operated by the Ministry of Defense, and 14% are operated by other ministries (USAID, 1988). Generally hospitals are not available below the district level (Berman, 1989).

Over half of Indonesia's hospitals are privately owned, many by religious organizations or industries (USAID, 1988). Private hospitals account for about one-third of the hospital beds in the country and are oriented almost exclusively toward curative care. Private hospitals are far too expensive for most Indonesians (Brotowasisto et al., 1988).

Health Centers

Health centers (*puskesmas*) are the most peripheral MOH clinic facilities. Each subdistrict has at least one health center. Some urban subdistricts have considerably more than one clinic because of their high population density (there are 5600 health centers for approximately 3200 subdistricts).

The *puskesmas* has a number of functions, including all medical care, family planning, communicable disease control, environmental health and sanitation, health education and nutrition, dental health, laboratory services, and recording and reporting. In recent years the outreach activities of the *puskesmas* have expanded considerably (see below).

The health center is headed by a medical doctor, usually a recent graduate serving his or her required two to five year term after medical school. The doctor is responsible for supervising

[2]As of 1986 there were 4 Class A hospitals, 16 Class B hospitals, 79 Class C hospitals, and 219 Class D hospitals (USAID, 1988). Expenditures on hospitals accounted for 75% of the MOH's routine budget in 1982/83, after deducting the central government's administrative expenses (USAID, 1987).

all the activities of the health center and rarely sees patients during health center hours.[3] Not all health centers claim a doctor in residence. One doctor may serve several centers in remote areas where it is hard to lure physicians. On the other hand, some urban health centers have several doctors. The ratio of doctors to health centers ranges from .35 in Irian Jaya to 1.78 in Jakarta (The World Bank, 1989).

The nurse-midwife (*bidan*) is second in authority at the health center after the doctor. The nurse-midwife oversees maternal and child health (MCH) services offered at the *puskesmas*, outreach to the community, and training for *dukun bayi* (traditional midwives). Additionally, since the nurse-midwife usually serves at a health center for longer than the doctor, she provides stability of service delivery within the community.

A number of other staff positions exist within the *puskesmas* as well. The extent to which they are filled varies from center to center. Some of these positions are dentist, nurse, sanitarian, vaccinator, nutritionist, pharmacist, and laboratory worker. Often other non-technical employees are on the payroll, such as appointment secretaries, drivers, janitors, and night watchmen. The facilities of the *puskesmas* include a reception area, an administrative office for the doctor, a dispensary, a bathroom, and several examining rooms (one of which usually contains a gynecological table).

In addition to health centers, some subdistricts also have auxiliary health centers (*puskesmas pembantu*). Subcenters are provided with the goal of increasing access to public facilities in areas where either population density is high or geographical features of the subdistrict render one health center inadequate. Staffing patterns of subcenters vary. Some subcenters are opened one to three times a week by paramedical staff from the health center. Other subcenters maintain the same hours of the health center but have a nurse or nurse-midwife as the head rather than a doctor.[4]

As the numbers of health centers has grown, so has the proportion of centers staffed by doctors. In 1974 doctors headed only 34% of health centers. By 1980 about half of all centers were headed by doctors (Hugo et al., 1987). In 1987 MOH estimates put the proportion of centers staffed by doctors at 90% and the average number of nurses per center at seven (GOI, 1988). The average area served per health center varies greatly from province to province. In Yogyakarta the average area served is 31 square kilometers, and households are generally within three

---

[3] Almost all health center doctors operate private practices after *puskesmas* hours, when they see a considerable number of patients.

[4] The distinction between health centers and subcenters seems to have changed in recent years. In the past subcenters were provided wherever one center was inadequate. Recently, some subcenters in high density areas (e.g.. Jakarta) seem to have been converted to full health centers through the assignment of doctors to serve in them. Today subcenters are more common in remote, low density areas, while high density areas often have more than one full health center.

kilometers of a health center. In Central and East Kalimantan one center serves an area of 1500 square kilometers, and households are within 25 kilometers of a center (Government of Indonesia, 1990; The World Bank, 1989).

Posyandu- Integrated Service Posts

The *posyandu* is a community-sponsored subvillage health service post that provides basic elements of maternal and child health care to neighborhood groups. In its ideal form, the *posyandu* occurs once per month (usually at the home of a villager) and serves no more than 100 children under five years old.[5] The post provides five services: family planning, maternal and child health care, nutrition, immunization, and diarrheal disease control. The *posyandu* is staffed by volunteers (*kaders*) from the village who have received a three to six day training course from the health center staff. Activities of the volunteers are supervised by health center staff and family planning field workers, who provide immunizations and keep the *posyandu* supplied with Oral Rehydration Solution (ORS) packets, growth-monitoring cards, and family-planning supplies.

PRIVATE SECTOR SERVICE DELIVERY

Apart from government sources of care at hospitals, health centers, and village-level *posyandus*, a variety of sources of private health care exist. The existence of private practitioners in Indonesia is well-established, as is the government's recognition of the private sector's potential role in providing services.

Modern Sector

Services operating within the biomedical paradigm are available from hospitals, clinics, doctors' and dentists' practices, and paramedical practices. A wide range of modern medicine is available at urban pharmacies. In rural areas small shops carry aspirins, antibiotics, and anti-malarial drugs. Widespread availability of drugs from small stalls, village shops, and pharmacies makes self-treatment a popular (and relatively inexpensive) option.

Hospitals are concentrated in urban areas. Over half the nation's 14,000 hospitals (and 30% of the hospital beds) are privately owned (USAID, 1988). Private hospitals service 20-25% of the total hospital patient load (GOI, 1988). Most private hospitals are located in a few of the larger cities and focus primarily on curative care. Some companies run hospitals and clinics for their

---

[5] The posyandu is an activity rather than a location. It takes place for a few hours one morning a month and then vanishes. Since the kaders are village residents they are physically available to provide advice to mothers, but there is no fixed service point open daily.

employees' use. Plantations often have a clinic for employees. The national oil company Pertamina, for example, runs a fully-staffed hospital in Jakarta with nurses, general practitioners, and specialists.

The private sector directly employs fifteen percent of all health workers (Brotowasisto, et al., 1988). Additionally, the majority of public sector health workers also contribute to the private sector health industry, either by running their own private practices at night or by working after hours in private clinics and hospital. Villagers typically have access to the private practices of at least one doctor, one nurse-midwife, and potentially several paramedics associated with the *puskesmas* . The distinction between public and private sector services is clearer from the perspective of a consumer, who makes a conscious choice between public and private services, than from the perspective of service providers, who typically play both roles. Household payments accounted for 64% of all health expenditures in 1982/83. Over half of these payments were for private service fees, while over one-third were for drug purchases. Most household payments for health services are out-of-pocket cash payments.

One private sector service that has expanded in availability in recent years is private maternity homes/clinics run by nurse-midwives (*bidan*). The nurse-midwives are graduates of vocational or technical schools or the nursing academy. These women offer various packages of antenatal care and delivery assistance and sometimes open a delivery clinic with two to four beds in their own home In the early 1990s, the MOH began a program called "Midwife in the Village," in which recently trained midwives are stationed directly in villages to provide quasi-private midwifery services.

Traditional Services

The types of traditional care available to Indonesians include buying or making herbal medicines (*jamu*) or Chinese remedies or using the services of various traditional practitioners (*dukuns*). The choices are not mutually exclusive. A *dukun's* cure often involves *jamu*. Rituals, such as incantations, and massage are also common elements of traditional cures.

A number of *dukuns* offer health services. The traditional practitioners with the most potential to impact maternal and child health are the *dukun bayi* (traditional midwives). *Dukun bayi* attend 80 to 90 percent of all births in Indonesia but provide little prenatal care. Increasingly women are consulting nurse-midwives for prenatal advice but using a *dukun bayi* for delivery. Over 70% of the *dukun bayi* in the country have been reached by a UNICEF training program (MOH, 1990a). The success of these efforts is unclear. In Aceh babies delivered by trained *dukun bayi* had almost the same tetanus mortality rate as babies delivered by untrained dukun bayi (Solter, Hasibuan, and Yusuf; 1988).

REFERENCES AND OTHER SOURCES OF INTEREST

USAID. 1988. "Strategic Plan. 1989-1994. Office of Population and Health. USAID/Indonesia.

World Health Organization. 1987. "Review of the program for the control of diarrhoeal diseases and the expanded program on immunization, Indonesia." Southeast Asia Region.

World Bank. 1989. "Indonesia. Poverty Assessment and Strategy Report."

World Bank. 1990a. "Health Sector Priorities Review. Overview." Population, Health, and Nutrition Division. Population and Human Resources Department. Washington, D.C.

Ministry of Health. 1990. Primary Health Care in Indonesia. MOH, Jakarta, Indonesia.

Hugo, G., T. Hull, V. Hull, and G. Jones. 1987. The Demographic Dimension in Indonesian Development. Singapore: Oxford University Press.

Hull, V. 1979. "Women, doctors, and family health care: some lessons from rural Java." Studies in Family Planning. 10(11-12):315-325.

Government of Indonesia. 1988. "Situation Analysis of Women and Children in Indonesia." GOI-UNICEF. Jakarta, Indonesia.

Berman, P. and S. Sakai. 1986. "The productivity of rural health manpower in Java." Department of International Health. The Johns Hopkins University. Baltimore, Maryland.

Berman, P., B. Ormond, and A. Gani. 1987. "Treatment use and expenditure on curative care in rural Indonesia." Health Policy and Planning. 2(4):289-300.

Berman, P. 1989. "Cost efficiency in primary health care: studies of health facilities in Indonesia." Health Policy and Planning 4(4):316-322.

Berman, P., D. Sisler, and J. Habicht. 1989. "Equity in public-sector primary health care: the role of service organization in Indonesia." Economic Development and Cultural Change. 37(4):777-803.

Brotowasisto, O. Gish, R. Malik, and P. Sudharto. 1988. "Health care financing in Indonesia." Health Policy and Planning. 3(2):131-140.

Soh-Sanu, R. 1989. "The midwife in private practice in Indonesia." Midwives Chronicle & Nursing Notes. April. 119, 122-123+.

Solter, S., A. Hasibuan, and B. Yusuf. 1986. "An epidemiological approach to health planning and problem-solving in Indonesia." Health Policy and Planning. 1(2):99-108.

Streatfield, K. and M. Singarimbun. 1988. "Social factors affecting the use of immunization in Indonesia." Social Science and Medicine. 27(11):1237-1244.

Simpson-Herbert, M., P. Piotrow, L. Christie, and J. Streich. 1980. "Traditional midwives and family planning." Population Reports. Series J, Number 22. Baltimore, Maryland: The Johns Hopkins University.

Appendix D

INDONESIAN EDUCATION SYSTEM

In this Appendix, a brief description of the Indonesian education system at the primary and secondary levels is presented since these are the levels covered by the IFLS-CF. The intent is to merely to provide a general outline for those users not familiar with the structure of the Indonesian education system.

The primary and secondary education system consists of three levels or cycles:

Primary (grades 1-6): Sekolah Dasar (SD)

Junior Secondary (grades 7-9): Sekolah Menengah Pertama (SMP)[1]

Senior Secondary (grades 10-12): Sekolah Menengah Atas (SMA)

Education is compulsory through the primary level. At the Junior Secondary and Senior Secondary levels, there are two sublevels: academic and vocational. While the public sector provides most of the education services, there is a strong private education sector as well. In addition there are both government religious schools and private religious schools rounding out the picture. All schools, public or private, religious or secular, must comply with the government-mandated curriculum. Schools may add extra courses beyond the mandated set.

The curriculum used in primary and secondary schools evolved in three phases. In the period 1968-1975, a basic national curriculum was designed and implemented but it was criticized for not providing sufficient detail in content requirements and in guidance to teachers. In 1976, the curriculum was revised and was highlighted by modifications of time allotted to key subjects and by the introduction of a system of differentiated credits for particular subjects. In 1985, the curriculum was revised again with increased emphasis on patriotism, the affective domain, and the spirit of independence.

At the end of each cycle, external examinations, called EBTANAS, are administered. Although the content of the EBTANAS is standardized across the nation, some local variation exists in test administration, and scoring is done at the provincial level. Therefore, comparison of scores across areas should be interpreted cautiously. The results of these examinations are combined with the students' grades in order to determine an overall score for performance certification for that level of education. This score in conjunction with other factors and individual preferences plays a role in determining what schools a student may enter at the next level. Because the preference for academic programs over vocational or

[1]Another acronym for the junior or lower secondary level in Indonesia is SLTP.

technical specializations, academic programs generally require higher scores than vocational or technical programs. Access to public schools also requires higher scores than most private schools, except the best ones.

The Ministry of the Interior has responsibility for primary education, the Ministry of Education and Culture oversees secondary and tertiary education, and the Ministry of Religion is responsible for religious schools.

REFERENCES AND OTHER SOURCES OF INTEREST

Kemmerer, Frances, Dean Nielsen and Patrick Lynch. *A Review of Teacher Education Issues in Indonesia. Jakarta, Indonesia: Center for Informatics*, Office of Educational and Cultural Research and Development, Ministry of Education and Culture, 1990.

Suryadi, Ace. *Improving the Educational Quality of Primary Schools*. Jakarta, Indonesia: Center for Informatics, Office of Educational and Cultural Research and Development, Ministry of Education and Culture, 1990.

Theisen, Gary, James Hughes and Paul Spector. *An Analysis of the Status of Curriculum Reform and Textbook Production in Indonesia*. Jakarta, Indonesia: Centere for Informatics, Office of Educational and Cultural Research and Development, Ministry of Education and Culture, 1990.

World Bank. "Indonesia: Basic Education Study". Wasington, D.C. 1990.

World Bank. "Indonesia Public Expenditures, Prices and the Poor". Washington, DC. 1993.