

HOUSEHOLD SURVEY CODEBOOK FOR THE  
INDONESIA FAMILY LIFE SURVEY, WAVE 2.

DRU-2238/5-NIA/NICHD  
March 2000

We recommend the following citations for the IFLS data:

For papers using IFLS1 (1993):

Frankenberg, E. and L. Karoly. 'The 1993 Indonesian Family Life Survey: Overview and Field Report.' November, 1995. RAND. DRU-1195/1-NICHD/AID

For papers using IFLS2 (1997):

Frankenberg, E. and D. Thomas. 'The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and 2.' March, 2000. RAND. DRU-2238/1-NIA/NICHD.

## PREFACE

This codebook describes each variable in the public-use files for the household survey (HHS) portion of the Indonesia Family Life Survey, wave 2 (IFLS2). It is the fifth of seven volumes documenting the IFLS2. The Indonesia Family Life Survey is a continuing longitudinal socioeconomic and health survey. It is addressed to a sample representing about 83% of the Indonesian population living in 13 of the nation's 26 provinces. The survey collects data on individual respondents, their families, their households, and the communities in which they live. The first wave (IFLS1) was administered in 1993 to 7,224 households. IFLS2, administered in 1997, was addressed to the same sample. A follow-up survey (IFLS2+) was conducted in 1998 with 25% of the sample. The next wave will be fielded in 2000.

IFLS2 was a collaborative effort of RAND, UCLA, and the Demographic Institute of the University of Indonesia (LDUI). Funding for IFLS2 was provided by the National Institute on Aging (NIA), the National Institute for Child Health and Human Development (NICHD), U. S. Agency for International Development (USAID), The Futures Group (POLICY Project), the Hewlett Foundation, the International Food Policy Research Institute (IFPRI), John Snow International (the OMNI project), and the World Health Organization. MACRO International developed the data-entry software and had responsibility for certain parts of the data processing.

The IFLS2 public-use file documentation, whose seven volumes are listed below, will be of interest to policymakers concerned about socioeconomic and health trends in nations like Indonesia, to researchers who are considering using or are already using the IFLS data, and to those studying the design and conduct of large-scale panel household and community surveys. Updates regarding the IFLS database subsequent to publication of these volumes will appear at the IFLS Web site, <http://www.rand.org/FLS/IFLS>.

## DOCUMENTATION FOR IFLS, WAVE 2

- DRU-2238/1-NIA/NICHD: The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and 2. Purpose, design, fieldwork, and response rates for the survey, with an emphasis on wave 2; main results from both waves 1 and 2.
- DRU-2238/2-NIA/NICHD: Users Guide for the Indonesia Family Life Survey, Wave 2. Descriptions of the IFLS file structure and data formats; guidelines for data use, with emphasis on using the wave 2 and wave 1 data together.
- DRU-2238/3-NIA/NICHD: Household Survey Questionnaire for the Indonesia Family Life Survey, Wave 2. English translation of the questionnaires used for the household and individual interviews. Includes interviewer's instructions.
- DRU-2238/4-NIA/NICHD: Community-Facility Survey Questionnaire for the Indonesia Family Life Survey, Wave 2. English translation of the questionnaires used for interviews with community leaders and facility representatives. Includes interviewer's instructions.
- DRU-2238/5-NIA/NICHD: Household Survey Codebook for the Indonesia Family Life Survey, Wave 2. Descriptions of all variables from the IFLS2 Household Survey and their locations in the data files.
- DRU-2238/6-NIA/NICHD: Community-Facility Survey Codebook for the Indonesia Family Life Survey, Wave 2. Descriptions of all variables from the IFLS2 Community-Facility Survey and their locations in the data files.
- DRU-2238/7-NIA/NICHD: Crosswalk between the Survey Instruments for the Indonesia Family Life Survey, Waves 1 and 2.

## RE-RELEASE OF IFLS1 DATA

To facilitate using the IFLS1 and IFLS2 data together, a revised version of IFLS1 data has been released. Abbreviated IFLS1-RR (1999), the re-release incorporates adjustments outlined in the . files, joins subfiles having the same unit of observation, and adds identifiers that make it easier to link IFLS1 and IFLS2 data. The IFLS-RR data are available at <http://www.rand.org/FLS/IFLS> and are documented in

DRU-1195/7-NIA/NICHD: Documentation for IFLS1-RR: Revised and Restructured Indonesia Family Life Survey Data, Wave 1.

## PREVIOUS DOCUMENTATION FOR IFLS, WAVE 1

DRU-1195/1-NIA/NICHD: The 1993 Indonesian Family Life Survey: Overview and Field Report. Purpose, design, fieldwork, and response rates.

DRU-1195/2-NIA/NICHD: The 1993 Indonesian Family Life Survey: Appendix A, Household Questionnaires and Interviewer Manual. English translation of the questionnaires used for the household and individual interviews. Includes interviewer's instructions.

DRU-1195/3-NIA/NICHD: The 1993 Indonesian Family Life Survey: Appendix B, Community-Facility Questionnaires and Interviewer Manual. English translation of the questionnaires used for interviews with community leaders and facility representatives. Includes interviewer's instructions.

DRU-1195/4-NIA/NICHD: The 1993 Indonesian Family Life Survey: Appendix C, Household Codebook. Descriptions of all variables from the Household Survey and their locations in the data files. Includes notes about cases that are known anomalies.

DRU-1195/5-NIA/NICHD: The 1993 Indonesian Family Life Survey: Appendix D, Community-Facility Codebook. Descriptions of all variables from the Community-Facility Survey and their locations in the data files. Includes notes about cases that are known anomalies.

DRU-1195/6-NIA/NICHD: The 1993 Indonesian Family Life Survey: Appendix D, Users' Guide. Descriptions of the IFLS file structure and data formats; guidelines for data use, with emphasis on working with the household, individual, and facility IDs and making links across different parts of the survey.

## ACKNOWLEDGMENTS

A survey of the magnitude of IFLS2 is a huge undertaking. It involved a large team of people from both the United States and Indonesia. We are indebted to every member of the team. We are grateful to each of our respondents, who gave up many hours of their time.

The project was directed by Elizabeth Frankenberg (RAND) and Duncan Thomas (RAND and UCLA), who were the Principal Investigators. Lynn Karoly and Paul Gertler were Principal Investigators in the early stages of the project.

Bondan Sikoki was the Project Director appointed by the Demographic Institute of the University of Indonesia (LDUI). She served as the Survey Director during the design and implementation of fieldwork. Her unswerving commitment to maintaining the integrity and quality of IFLS2, in even the most difficult circumstances, was an inspiration to us all. Prior to her appointment, the LDUI Project Director was Dr. IGN Agung.

Three LDUI staff members served as Associate Project Directors. Wayan Suriastini directed the tracking phase of the study and played a central role in the design of the Household Survey Questionnaire. Muda Saputra coordinated much of the Community-Facility Survey fieldwork and data entry. Sutji Rochani Siregar oversaw the administration of the latter phases of fieldwork and data entry.

Data-entry software and field procedures for the Computer-Assisted Field Editing (CAFE) were developed by Trevor Croft, of MACRO International, with the assistance of Hendratno of LDUI. Croft also developed the software used for the final phase of data entry/data quality checks (Look Ups). Iip Umar Ri'fai, Martin Wolfe, and Linda Fitrawati assisted with these tasks.

Eko Ganiarto coordinated the first and second pretests. Victoria Beard worked extensively on the Community-Facility Survey. Endjang Pudjani and Sheila Evans were responsible for the technical production of the Indonesian and English questionnaires. Akhir Matua Harahap coordinated the writing and production of the survey manuals. Mary Linehan managed operations in Jakarta prior to fieldwork; she developed the assessments of physical health, along with Cecep Sukria Sumantri and Merry Widayanti. Nargis, Djainal, and M. Yusuf assisted with the development of the Community-Facility Survey and the training of its staff. Donavan Bustami coordinated printing and shipping for the questionnaires.

John Adams provided critical input for the design of the follow-up protocols and guided the development of sampling weights. Christine Peterson designed the preprinted rosters, assisted with questionnaire design and processing of the pretest data, and helped calculate the sampling weights. The IFLS2 public-use data files were produced by a team based at RAND. The efforts of Paula Hamilton, Nancy Campbell, Melissa Chiu, Sue Polich, Patty St. Clair, Wayan Suriastini, and Peter Yau went well beyond the call of duty.

Many of our colleagues at RAND have contributed substantially to the survey. We are especially grateful to James P. Smith and John Strauss. We are also grateful to Kathleen Beegle, Julie DaVanzo, William Dow, Micki Fujisaki, Doug Gilbertson, Paul Gertler, Daryl Hill, Michael Hurd, Lynn Karoly, Jacob Klerman, Nancy Krantz, Donna Lee, Lee Lillard, Maria Menchaca, Eileen Miech, Jack Molyneaux, Mathew Sanders, Christine d'Arc Taylor, Jim Tebow, and Beverly Weidmer.

Much effort was put into designing IFLS2 so that it would yield information on topics of special concern in Indonesia and reflect the nation's distinctive social, economic, and policy environment. The input of a large number of scholars and policy-makers in Indonesia was key in this regard. Paramita Sudharto gave us considerable guidance on the overall survey and on its health components. Important contributions were made by Boediono, Mark Brook, Fasli Djalal, Herwindo Haribowo, Bachrul Hayat, Heryudarini, Yayah Husaini, Bambang Indrianto, Stephanus Indradjaya, Jiono, Robert Kim-Farley, Vanda Moriaga, Dr. Mujilah, Muljani Nurhadi, Ratna, Kushnadi Setjawinata, Soeharsono Soemantri, James Stein, Ace Suryadi, and Anton Wijaya.

The survey could not have taken place without the support of the LDUI directors and administrative staff, including N. Haidy Pasay, Sri Moertiningsih Adioetomo, Sri Hariati Hatmadji, Badrun, and Teguh. We are indebted to the Population Study Centers in each of the thirteen IFLS provinces, which helped us recruit the 400 field staff.

Finally, the success of the survey is largely a reflection of the diligence, persistence and commitment to quality of the interviewers, supervisors, and field coordinators. Their names are listed in the Study Design (DRU-2238/1-NIA/NICHD), Appendix A.

```

=====
IIII  FFFFFFFF  LLL  SSSSSS  22222
II    FF  F    LL   SS   SS  22  22
II    FF       LL   SS    S   22
II    FF       LL   SSS      22
II    FFFFFF  LL   SSSSS     22
II    FF       LL   SSSSS     22
II    FF       LL       SSS    22
II    FF       LL       S   SS  22
II    FF       LL    L  SS   SS  22
IIII  FFFF     LLLLLLLL  SSSSSS  222222222
=====

```

# I N D O N E S I A   F A M I L Y   L I F E   S U R V E Y   2

A survey of individuals, households, families, communities  
and facilities conducted in 13 provinces in Indonesia  
between August 1997 and February 1998

=====

Directed by

Elizabeth Frankenberg, RAND

and

Duncan Thomas, RAND & UCLA

with Bondan Sikoki, University of Port Harcourt & RAND and  
Wayan Suriastini, Lembaga Demografi, University of Indonesia & RAND

=====

Financial support for the Indonesia Family Life Surveys has been  
provided by the National Institute on Aging, the National Institute  
for Child and Human Development, USAID, World Health Organization  
the POLICY Project (Future's Group International), The World Bank,  
OMNI (John Snow International), the Hewlett Foundation, the Ford  
Foundation, IFPRI and the United Nations Fund for Population.

=====

Website: <http://www.rand.org/FLS/IFLS>  
Send all questions to [ifls-supp@rand.org](mailto:ifls-supp@rand.org)

Copyright (1999)





```

=====
H   H   OOO   U   U   SS   EEEE   H   H   OOO   L       DDDD
H   H   O   O   U   U   S   S   E       H   H   O   O   L       D   D
H   H   O   O   U   U   S       E       H   H   O   O   L       D   D
HHHHH   O   O   U   U   S       EEEE   HHHHH   O   O   L       D   D
H   H   O   O   U   U       S       E       H   H   O   O   L       D   D
H   H   O   O   U   U       S       E       H   H   O   O   L       D   D
H   H   O   O   U   U   S   S   E       H   H   O   O   L       D   D
H   H   OOO       UUU       SS   EEEE   H   H   OOO   LLLL   DDDD

```

```

      SS   U   U   RRR   V   V   EEEE   Y   Y
      S   S   U   U   R   R   V   V   E       Y   Y
      S       U   U   R   R   V   V   E       Y   Y
      S       U   U   RRR   V   V   EEE       Y
      S       U   U   R   R       V   V   E       Y
      S   S   U   U   R   R       V       E       Y
      SS       UUU       R   R       V       EEEE       Y

```

```

      CCC   OOO   DDDD   EEEE   BBBB   OOO   OOO   K   K
C   C   O   O   D   D   E       B   B   O   O   O   O   K   K
C       O   O   D   D   E       B   B   O   O   O   O   K   K
C       O   O   D   D   EEE   BBBB   O   O   O   O   KK
C       O   O   D   D   E       B   B   O   O   O   O   KK
C       O   O   D   D   E       B   B   O   O   O   O   K   K
C   C   O   O   D   D   E       B   B   O   O   O   O   K   K
      CCC   OOO   DDDD   EEEE   BBBB   OOO   OOO   K   K

```

```

=====
B O O K :      TRACK
T O P I C :    TRACKING OF HHS & INDIVIDUALS ACROSS WAVES
=====

```

Prepared by  
 Elizabeth Frankenberg, Paula Hamilton, Suzanne Polich  
 Muda Saputra, Wayan Suriastini and Duncan Thomas  
 Created on 28 May 2000 at 16:38



BOOK : TRACK

TOPIC : TRACKING OF HHS &amp; INDIVIDUALS ACROSS WAVES

## DATA FILES

Data are available in RAW (ascii), SAS (6.12) export and STATA (6.0) format. Data and electronic documentation are available on the IFLS website. RAW (ascii) data for this book are contained in the data file:

trk.dat with logical record length 141

The survey instrument contains multiple sections. Where possible, sections have been merged together to form a single data module. The modules are organized by level of observation. The module names are listed below. They serve as the names of the SAS and STATA data files. The module names are contained in the first 8 columns of the ascii data file so that each module can be easily identified. The codebook is also organised by module. The modules are:

Module	Description	Level of observation	# of recs
htrack		Household	8116
ptrack		Individual	39785

## USER SUPPORT

These data should be used in conjunction with the Questionnaires and Users' Manuals. Check the IFLS web page for updates and answers to frequently asked questions. Please send all questions to ifls-supp@rand.org.

## USAGE NOTES

## READING THE DICTIONARY

For each variable in the dataset, the dictionary provides

Variable name : Typically the question number

Start posn : Starting position of variable in raw data file

Length : Number of characters used by that variable

Decimal : Number of decimal places to be assigned to variable

Format : A if alphanumeric and N if numeric

Label : Describes variable

Variable coding is described as follows:

Value : Meaning of values for coded variables

Record count : Frequency of each value

Note1: Values with a record count of zero do not appear in codebook. See survey instrument for complete set of allowable responses.

Note2: Questions with alphanumeric responses may have multiple answers. Thus, record count may exceed number observations in data file.

## SPECIAL CODES

The following values are reserved and have a special meaning:

Numeric	Alphanumeric	Meaning
5	V	Top coded/out of range
6	W	Not applicable
7	X	Refused to answer
8	Y	Don't know
9	Z	Missing

Numeric special values are preceded by as many 9s as necessary to fill the field and yield an unambiguous value. For example if a field is 4 digits wide, 9998 indicates the respondent did not know the answer.

Since special values that are embedded in continuous variables are tedious to deal with, in many cases a continuous variable, `_VAR_`, say, is accompanied by another variable, `_VAR_X` which contains the special codes. If a valid value of `_VAR_` is recorded, `_VAR_X` is equal to 1.

## DATASET VERSION

The variable VERSION identifies the release version of these data; it will be updated with each revision of the data and can be used to confirm that you are using the most recent version of the data.

## NOTES ABOUT TRACKING FILES

The TRACKING FILES, HTRACK and PTRACK, are provided to facilitate using the longitudinal dimensions of the survey. All variables included in these files are drawn from interviews conducted in IFLS1 (1993) and IFLS2 (1997/8).

## HOUSEHOLD TRACKING FILE

HTRACK contains a list of all HOUSEHOLDS that were interviewed in IFLS1; they are identified by the 1993 HOUSEHOLD identifier, HHID93. It is a 7 digit string variable. The first 3 digits are the enumeration area in which the household resided in 1993 and the next two digits are a household sequence number within that enumeration area which uniquely identifies the household. The last two digits are always '00'. (The first 5 digits of HHID93 are the same as the last 5 digits of CASE, the HH identifier variable in the original IFLS1 release.)

In IFLS2, households are identified by HHID97. If an IFLS household was found intact in 1997, it was assigned the same identifier in 1993 and 1997; in this case HHID93=HHID97. If the household had split up between 1993 and 1997, then when the first respondent from that household was re-contacted in IFLS2, that respondent's household was designated the 'original' household and the 1993 household identifier assigned to it. Each additional household that was spawned by that HHID93 was given a new HHID97. The first 5 digits of the new HHID97 are identical to the first 5 digits of HHID93 (and, therefore, all new households in 1997 that are spawned by one 1993 household share the same first 5 digits in their HHIDs). The last two digits of HHID97 are 1 (in column 6) and then a sequence number starting at 1 (in the 7th column); these digits tell us this is a split-off

household. Thus the last two digits of HHID97 are '00' for the first household found in 1997 and then '11' for the first split off, '12' for the second split off and so on.

For example, say HHID93 is 2071900. This is household 19 from enumeration area 207. The household split into 3 households between 1993 and 1997. The three households are assigned HHID97 2071900 (for the first HH relocated), 2071911 (for the first split-off HH that was relocated) and 2071912 (for the second split-off HH that was relocated).

#### Location of HH

Note that the enumeration area in 1993 (digits 1-3) in the HHID is not the location of residence of the household in 1997 (unless the household has not moved) and should not be used to determine geographic location of the household. 1993 location was built into the 1993 HH identifier, CASE. It is not built into HHID93 or HHID97.

Location information is recorded in module BK\_SC in each wave of the survey. A summary is included in HTRACK. Location in 1993 is recorded in SC01\_93 through SC05\_93; the 1993 location codes are based on the 1993 BPS codes. Some of these codes have been changed by BPS (because the community boundaries have been re-defined, for example). The 1998 BPS codes for the location of each of our respondents are recorded in the revised kabupaten code, SC02\_93R, and the revised kecamatan code, SC03\_93R. (There are no revised province codes.) The 1997 location of the respondents is recorded in SC01\_97 through SC05\_97. These locations use the 1998 BPS codes and so may be directly compared with the revised 1993 codes, SC02\_93R and SC03\_93R.

MOVER97 is intended to summarise the location of the respondent in 1997, relative to the location in 1997. It is defined only for those respondents interviewed in both waves of the survey.

#### Linking community level data

Commid is the variable that should be used to link household survey data with the community and facility data. COMMID93 identifies the community of residence of the respondent in 1993. COMMID97 is the 1997 community of residence. COMMID93 will be the same as COMMID97 if the respondent has not moved between the waves of the survey.

#### HH recontact

There were 7,730 households in the target sample for IFLS1. Of those, interviews were completed with 7,224 households. These households are included in HTRACK. For information on the 506 households that were listed but never interviewed, see Book K in IFLS1.

IFLS2 sought to re-interview all 7,224 IFLS1 households. Around 6% of the target households were not interviewed. The results of our attempts to re-interview all households are summarised in RESULT97.

In addition to the approximately 6,750 'ORIGIN' households that were interviewed in IFLS1 and IFLS2, over 850 'SPLIT OFF' households were interviewed in IFLS2. These are households in which a TARGET respondent who had moved out of an IFLS1 household

was interviewed. There are slightly over 7,600 households in IFLS2 that completed a household roster. These households, in combination with the households that were not found in IFLS2 make up the 8,116 households in HTRACK.

In 1997, we discovered 10 of the IFLS1 households had combined with another IFLS1 household. The original household members were interviewed in the new household.

In a small number of households, it was determined that all the members of the household had died by 1997. These were typically one or two member households in 1993 and the members in the 1993 household were typically relatively old. There were, however, a small number of households in which 1993 household members were still alive in 1997 but the household was treated as if all members had died. These cases arose because the TARGET individuals in the household had died by 1997 and the interviewers mistakenly thought they did not need to track the remaining members who had moved away.

#### PERSON TRACKING FILE

PTRACK is a person-level file that tracks all IFLS respondents across waves of the survey. PID93 is a two digit sequence number identifying each individual within a household. The combination of HHID93 and PID93 uniquely identifies every respondent in IFLS1. It may be used to link records within IFLS1. If an IFLS1 respondent was found in the original household, HHID97=HHID93 and PID97=PID93. All new respondents in IFLS2 are assigned PID97 that begins after the highest PID93 for that household. In split-off households, PID97 was assigned starting at 01 for the household head. The combination of HHID97 and PID97 uniquely identifies every respondent in IFLS2. It may be used to link records within IFLS2.

HHID93 and PID93 or HHID97 and PID97 should NOT be used to link respondents across waves of IFLS.

#### Linking respondents across waves: PIDLINK

Several individuals have moved across households between IFLS1 and IFLS2. In order to link records for a particular individual across waves of the IFLS, use PIDLINK. It is a unique person-level identifier which is the same in IFLS1 and IFLS2 for a particular individual. PIDLINK is a string variable comprising 9 digits. For a respondent in IFLS1 and IFLS2, PIDLINK is made up of HHID93 followed by 00 (denoting an original household member) and then PID93, the person identifier in IFLS1. If the respondent has moved from his or her original household, PIDLINK will retain the information necessary to identify the original household. For a new respondent in IFLS2, PIDLINK is made up of HHID97 and PID97.

PTRACK contains one record for every respondent. Some respondents were interviewed in both IFLS1 and IFLS2, some were interviewed only in IFLS1 and some were interviewed only in IFLS2. Note that in BK\_AR1, a respondent may appear more than once in a roster since all 1993 household members are listed in the 'ORIGIN' roster. A respondent who has moved out will be designated thus (AR01A\_97=3). If that respondent has been found in a new household, then AR01A\_97 will equal 4. Since this respondent is found in 2 different households in 1993 and 1997,

HHID93 and HHID97 are different. PID93 and PID97 will also be different in general. PIDLINK, however, remains constant.

Continuing the example of HHID93=2071900, there were 5 members in the household in 1993. In 1997, persons 1, 2 and 4 were still there but persons 3 and 5 had moved out. Person 3 was found in HHID97=2790911 and person 5 was found in HHID97=2790912.

The PTRACK records for this household are as follows:

PIDLINK	HHID93	PID93	HHID97	PID97	
207190001	2071900	1	2071900	1	) Original HH
207190002	2071900	2	2071900	2	) Persons 3 and
207190003	2071900	3	2071911	3	) 5 have split
207190004	2071900	4	2071900	4	) off and are found
207190005	2071900	5	2071912	2	) elsewhere.
207191101			2071911	1	) First split off
207191102			2071911	2	) (207190003 is in
207191104			2071911	4	) this HH.)
207191201			2071912	1	) Second split off
					(207190005 is in
					this HH.)

#### Age and sex

Information on age and sex is asked and used in many places throughout the IFLS questionnaire. PTRACK contains an estimate of the age of the respondent as of 1993 (AGE\_93) and as of 1997 (AGE\_97). These are estimates that have been constructed using all the information reported in the surveys and involve our own judgements about which of those reports is most likely to be right. They are provided only as a service to users and should not be construed as 'correct'. They are nothing more than our best estimates of each respondent's age. There is one respondent in IFLS1 whose sex is unknown; that respondent was not found in IFLS2.

PTRACK summarizes information on which respondents completed individual assessments (Books 3, 4 and 5) as well as who completed the health and cognitive assessments.

#### WEIGHTS

Household level weights are provided in HTRACK; person level weights are in PTRACK. For a full description of the construction and interpretation of the weights, see volume 1 of this documentation, 'The IFLS Study Design and Results'. That discussion contains a table summarizing the weights included with the data files.

#### Household level weights

HWT93 is the IFLS1 household level weight and is the inverse of the sample selection probability for each household interviewed in IFLS1. When it is applied to the sample, the resulting weighted distribution will reflect the 1993 distribution of households in rural and urban areas within each of the 13 provinces covered by the IFLS. This weight is based on the 7,224 HHs interviewed in 1993. HWT93SMP is based on the 7,730 HHs

included in the listing of the target sample for IFLS1.

Two classes of weights have been calculated for use with IFLS2 respondents: longitudinal analysis weights and cross-section analysis weights. IFLS2 longitudinal analysis weights are intended to update the IFLS1 weights because of attrition so that the IFLS2 panel sample is representative of the Indonesian population living in the 13 IFLS provinces in 1993. The weights are constructed using a propensity score method to calculate an adjustment for attrition which is multiplied by a factor to take account of design effects.

It might be argued that the sample of IFLS1 respondents who were interviewed in 1997 in combination with the new entrants in IFLS2 is sufficiently similar to the Indonesian population living in Indonesia in 1997 that one could use the sample to describe that population. Since the IFLS1 sample design included over-sampling in urban areas and off Java, users will need to re-weight the sample to take these design effects into account. The IFLS2 cross-section analysis weights are intended to do that.

HWT97L is the IFLS2 household level longitudinal analysis weight. HWT97X is the IFLS2 household level cross-section analysis weight.

#### Person level weights

Person weights in IFLS1 are complicated by the within-household sampling scheme adopted in the first round of the survey. There are three individual-level weights.

PWT93 is assigned to every individual listed in the IFLS1 household roster. The weights are designed so that the weighted age and sex distribution of individuals in IFLS1 reflect the 1993 population age and sex distribution by urban and rural strata within the 13 provinces covered by the survey.

PWT97L is the IFLS2 person-level weight for use in longitudinal analysis. PWT97L is zero for all respondents who were new entrants in 1997 (i.e. they were not resident in an IFLS household in 1993). PWT97X is the IFLS2 person-level weight assigned to every individual resident in an IFLS2 household in 1997; those weights are intended to reproduce the age and sex distribution of individuals in the 13 IFLS provinces in Indonesia in 1997.

PWT93IN are the IFLS1 respondent weights which take into account the within-household sampling scheme used to select respondents for individual interviews in IFLS1. We refer to these respondents as the IFLS1 MAIN respondents. They include the head, spouse, a sample of their children and possibly other adults in the household.

PWT97INL are the longitudinal analysis weights that should be used with data collected from these MAIN respondents in IFLS2. There is no cross-section weight for these respondents. In IFLS2, we sought to interview all household members and so users should use all respondents with PWT97X if the goal is to calculate estimates that are representative of the 1997 population.



## IMPORTING DATA INTO SAS AND STATA

All data files are organised in ZIP files. Use unix zip, winZIP or pkzip to unzip the files. The STATA data sets should be readable on any unix or windows platform. The SAS files are in EXPORT format. They will need to be imported to be compatible with your platform.

The following SAS code should import the data files from the export file that contains the control book (Book K) datasets.

```
libname bk xport "/directory of data/hh97bk.xpt";
libname mydata "/your data directory";

proc copy in=bk out=mydata;
```

## VALUE LABELS

Those variables for which they are appropriate, VALUE LABELS have been attached in the SAS and STATA data files.

In STATA, the value labels can be suppressed using the , NOLABEL option.

SAS is a bit more complicated. SAS maintains VALUE LABELS (or what SAS calls "formats") in a library. The format libraries are included in the public use data files as export datasets. Those files should be imported prior to being accessed from SAS and made accessible to you as a format library. This may be accomplished with the following code.

```
libname fmt xport "/directory of data/hh97fmt.xpt";
libname myformat "/your format library directory";

proc copy in=fmt out=myformat;
proc format cntlin=myformat.hhfmts;
```

If you wish to suppress the formats when printing output, do not point to the format library and disable format error checking with the global options statement:

```
OPTIONS NOFMterr;
```

at the top of your program. If you wish to remove all formats from the variables use the 'FORMAT \_ALL\_' statement.

Variable name	Start posn	Len gth	Deci mal	For mat	Variable label Value	Record count
MODULE	1	8	0	A	Module name	
HHID97	9	7	0	A	Household ID (1997 wave)	
HHID93	16	7	0	A	Household ID (1993 wave)	
CASE	23	9	0	N	IFLS HH number (used in IFLS1) # obs	7224
COMMID93	32	4	0	A	Community ID (1993 wave)	
COMMID97	36	4	0	A	Community ID (1997 wave)	
RESULT93	40	1	0	N	93 Result of HH ivw 1. Done Total	7224 7224
RES93BK	41	1	0	N	93 Result Bk K 1. Done Total	7224 7224
RES93B1	42	1	0	N	93 Result Bk 1 1. Done Total	7224 7224
RES93B2	43	1	0	N	93 Result Bk 2 1. Done 3. Not done Total	7184 40 7224
RESULT97	44	2	0	N	97 Result of HH ivw 01. Interview conducted 02. Joined other IFLS HH 03. All target mems dead 04. Refused 11. Moved (w/in prov) 12. Moved (oth IFLS prov) 13. Moved (non IFLS prov) 14. Moved (diff country) 15. Moved (DK location) 21. Unable to contact 22. Building vacant 23. Building demolished 24. Building not found Total	7620 10 69 88 87 69 47 6 79 7 3 3 28 8116
RES97BK	46	1	0	N	97 Result Bk K 1. Done 3. Not done Total	7699 417 8116
RES97B1	47	1	0	N	97 Result Bk 1 1. Done 3. Not done Total	7566 54 7620

Variable name	Start posn	Length	Decimal	Format	Variable label Value	Record count
RES97B2	48	1	0	N	97 Result Bk 2	
					1. Done	7600
					3. Not done	20
					Total	7620
HWT93	49	8	6	N	93: HH weight =HHWT224 in IFLS1	
					Mean	1.00
					Min	0.24
					Max	1.91
					# obs	7224
HWT93SMP	57	8	6	N	93: HH wt for target sample=HHWT730	
					Mean	1.00
					Min	0.24
					Max	1.92
					# obs	7224
HWT97L	65	8	6	N	97: Longitudinal HH weight	
					Mean	0.94
					Min	0.03
					Max	3.46
					# obs	7688
HWT97X	73	8	6	N	97 Xsection HH weight (raked->SUSENAS97)	
					Mean	1.00
					Min	0.29
					Max	1.70
					# obs	7629
SC01_93	81	2	0	N	93: Province of residence (93 BPS code)	
					12. North Sumatra	608
					13. West Sumatra	401
					16. South Sumatra	405
					18. Lampung	312
					31. Jakarta	800
					32. West Java	1254
					33. Central Java	1014
					34. Yogyakarta	528
					35. East Java	1164
					51. Bali	383
					52. W. Nusa Tenggara	461
					63. South Kalimantan	375
					73. South Sulawesi	411
					Total	8116
SC02_93	83	2	0	N	93: Kabupaten of residence (93 BPS code)	
					Mean	27.48
					Min	1.00
					Max	78.00
					# obs	8116
SC03_93	85	3	0	N	93: Kecamatan of residence (93 BPS code)	
					Mean	90.72
					Min	10.00
					Max	730.00
					# obs	8116

Variable name	Start posn	Len gth	Deci mal	For mat	Variable label Value	Record count
SC05_93	88	1	0	N	93: Urban/rural residence	
					1. Urban	3827
					2. Rural	4289
					Total	8116
SC02_93R	89	2	0	N	93: Kabupaten resid(revise->98 BPS code)	
					Mean	28.14
					Min	1.00
					Max	78.00
					# obs	8116
SC03_93R	91	3	0	N	93: Kecamatan resid(revise->98 BPS code)	
					Mean	83.75
					Min	10.00
					Max	730.00
					# obs	8116
SC01_97	94	2	0	N	97: Province of residence (98 BPS code)	
					12. North Sumatra	541
					13. West Sumatra	375
					16. South Sumatra	372
					18. Lampung	291
					31. Jakarta	651
					32. West Java	1258
					33. Central Java	999
					34. Yogyakarta	486
					35. East Java	1119
					51. Bali	359
					52. W. Nusa Tenggara	456
					63. South Kalimantan	338
					73. South Sulawesi	392
					Total	7637
SC02_97	96	2	0	N	97: Kabupaten of residence (98 BPS code)	
					Mean	26.85
					Min	1.00
					Max	78.00
					# obs	7637
SC03_97	98	3	0	N	97: Kecamatan of residence (98 BPS code)	
					Mean	85.06
					Min	10.00
					Max	730.00
					# obs	7637
SC05_97	101	1	0	N	97: Urban/rural residence	
					1. Urban	3505
					2. Rural	4129
					Total	7634

Tracker files: Household and individual tracking

Module: htrack

11

Variable name	Start posn	Len gth	Deci mal	For mat	Variable label Value	Record count
MOVER97	102	1	0	N	Moved between 1993 & 1997?	
					0. Did not move	6131
					1. w/in same desa	565
					2. w/in same kec	203
					3. w/in same kab	254
					4. w/in same prov	324
					5. Other IFLS prov	160
					Total	7637
VERSION	103	17	0	A	Dataset version	

Variable name	Start posn	Length	Decimal	Format	Variable label Value	Record count
MODULE	1	8	0	A	Module name	
HHID97	9	7	0	A	97: HH identifier	
PID97	16	2	0	N	97: Person identifier	
					# obs	38250
PIDLINK	18	9	0	A	Person ID (constant across IFLS waves)	
HHID93	27	7	0	A	93: HH identifier	
PID93	34	2	0	N	93: Person identifier	
					# obs	33081
CASE	36	9	0	N	1993 HHID (numeric)	
					# obs	33081
MEMBER93	45	1	0	N	Current resident of HH (93)	
					Mean	0.83
					Min	0.00
					Max	1.00
					# obs	39785
MEMBER97	46	1	0	N	Current resident of HH (97)	
					Mean	0.85
					Min	0.00
					Max	1.00
					# obs	39785
SEX	47	1	0	N	Sex of Respondent	
					1. M	19321
					3. F	20462
					9. Ms	1
					Total	39784
AGE_93	48	3	0	N	Age of Resp in 1993 (constructed)	
					Mean	26.25
					Min	0.00
					Max	111.00
					# obs	33081
AGE_97	51	3	0	N	Age of Resp in 1997 (constructed)	
					Mean	28.26
					Min	0.00
					Max	998.00
					# obs	38248
BTH_YR	54	4	0	N	Birth YEAR - (Constructed) (97)	
					Mean	2398.48
					Min	1897.00
					Max	9998.00
					# obs	39484
BTH_MTH	58	2	0	N	Birth MONTH - (Constructed) (97)	
					Mean	32.38
					Min	1.00
					Max	98.00

Variable name	Start posn	Len gth	Deci mal	For mat	Variable label Value	Record count
						# obs 38885
BTH_DAY	60	2	0	N	Birth DAY - (Constructed) (97)	
						Mean 64.76
						Min 1.00
						Max 98.00
						# obs 36814
AR02_93	62	2	0	N	93: Relation to HH Head	
						01. Head 7226
						02. Spouse 5737
						03. Child 14869
						04. Adopted child 317
						05. S/D-in-law 657
						06. Parent 427
						07. F/M-in-law 421
						08. Sibling 361
						09. B/S-in-law 345
						10. Grandchild 1634
						11. Grandparent 46
						12. Uncle/aunt 29
						13. Nephew/niece 390
						14. Cousin 29
						15. Servant 232
						16. Relative 140
						17. Non-relative 212
						96. N/A 1
						99. Miss 8
						Total 33081
AR01A_97	64	1	0	N	97: Is person member of HH?	
						0. Dead 857
						1. In HH 25769
						3. Not in HH 3456
						4. In new HH 1464
						5. New HH mem 6704
						Total 38250
AR01B_97	65	1	0	N	97: Track respond? (Preprinted in IFLS2)	
						1. Yes 23948
						3. No 9133
						Total 33081

Variable name	Start posn	Len gth	Deci mal	For mat	Variable label Value	Record count
AR02_97	66	2	0	N	97: Relation to HH head (=ar02b in AR1)	
					01. Head	7620
					02. Spouse	6015
					03. Child	16608
					04. Adopted child	325
					05. S/D-in-law	980
					06. Parent	482
					07. F/M-in-law	520
					08. Sibling	405
					09. B/S-in-law	475
					10. Grandchild	2251
					11. Grandparent	44
					12. Uncle/aunt	53
					13. Nephew/niece	642
					14. Cousin	51
					15. Servant	353
					16. Relative	193
					17. Non-relative	202
					18. Tenant	65
					19. Friend	88
					21. Ex spouse	12
					22. Fam of ex spse	1
					Total	37385
RES93B3	68	1	0	N	Result 93: Book 3 done by Resp?	
					1. Done	12992
					2. Done by proxy	1426
					3. Not done	231
					Total	14649
RES93B4	69	1	0	N	Result 93: Book 4 done by Resp?	
					1. Done	4888
					2. Done by proxy	93
					3. Not done	166
					Total	5147
RES93B5	70	1	0	N	Result 93: Book 5 done by Resp?	
					1. Done by child	770
					2. Done by adult	6981
					3. Not done	121
					Total	7872
RES93US	71	1	0	N	Result 93: (US=CA)Health assess done?	
					1. Done	24479
					3. Not done	8602
					Total	33081
RES97B3A	72	1	0	N	Result 97: Book 3A done by Resp?	
					1. Done	19910
					2. Done by proxy	1652
					3. Not done	1633
					Total	23195



Variable name	Start posn	Len gth	Deci mal	For mat	Variable label Value	Record count
RES97B3B	73	1	0	N	Result 97: Book 3B done by Resp?	
					1. Done	19892
					2. Done by proxy	1653
					3. Not done	1650
					Total	23195
RES97B3P	74	1	0	N	Result 97: Book 3P done by Proxy?	
					1. Done	1653
					3. Not done	1651
					Total	3304
RES97B4	75	1	0	N	Result 97: Book 4 done by Resp?	
					1. Done	6160
					2. Done by proxy	154
					3. Not done	369
					Total	6683
RES97B5	76	1	0	N	Result 97: Book 5 done by Resp?	
					1. Done by child	2893
					2. Done by adult	7492
					3. Not done	374
					Total	10759
PWT93	77	8	6	N	93 person weight (=ROSTERWT in IFLS1)	
					Mean	1.00
					Min	0.09
					Max	3.10
					# obs	33081
PWT97L	85	8	6	N	97 Longitudinal analysis person weight	
					Mean	1.15
					Min	0.25
					Max	6.47
					# obs	27233
PWT97X	93	8	6	N	97 Xsection person weight	
					Mean	1.00
					Min	0.25
					Max	2.52
					# obs	33937
PWT93IN	101	8	6	N	93 MAIN respondent wt (=RESPWT in IFLS1)	
					Mean	1.19
					Min	0.25
					Max	5.72
					# obs	22019
PWT97INL	109	8	6	N	97 Longitudinal person wt--93 MAIN resp	
					Mean	1.08
					Min	0.24
					Max	2.93
					# obs	18886
PWT93US	117	8	6	N	93 psn wt for health msrs (=CA_WT IFLS1)	
					Mean	0.66
					Min	0.00
					Max	5.72

Variable name	Start posn	Len gth	Deci mal	For mat	Variable label Value	Record count
						# obs
						33078
VERSION	125	17	0	A	Dataset version	