Nang Rong Projects

Codebook for the 1994 Moved Households

Carolina Population Center, University of North Carolina at Chapel Hill
The Institute for Population and Social Research, Mahidol University

May 26, 2004

Using this Codebook

This codebook is a product of notes that were written on copies of the 1984 questionnaire (Form A) that were brought into the fieldwork during the 1994 data collection.   If none of the members from the 1984 old household were living in the village in 1994, then the interviewers wrote on the copy of the 1984 questionnaire noting the household's current location. The information came from others in the village, possibly neighbors or relatives of the household. Below are some important things to know when working with this codebook:

Data File Pointers

There is a data file "pointer" at the top of every page. All variables described on the page are located in the data file named.

Last Modified: 05/21/03   Data File ==> MOVED94

Column x Column

Question No.	Variable			Question	Code
	Name	Type	Size		

- **Question No**. This column is not relevant since there was no questionnaire. Therefore, it has been left blank.

- **Variable** lists the...

  > **Name** which is the SAS variable name in the data file.
  > **Type** which tells whether it is a numeric variable (N) or an alphanumeric variable (A).
  > **Size** which gives the maximum length of the variable.

- **Question** provides a short description of the variable. In this same column, below the description, is either the labels for each of the codes on this variable or a pointer to go see one of the codebook appendices.

- **Code** lists either the actual values on the variable or the format of those values.

Variable Naming Convention?

There is not a strict naming convention per se, but there are some things to point out:

- We still adhere to 8 characters or less for all variable names.
- Most variables start with a 'MH' which stands for moved household.
- An identifier variable ends with the last 2 digits of the survey year from which it came.

Missing Values

- **Not Applicable**:   This code is always SAS system missing, which is a value of "." for numeric variables or a value of " " (or blank) for alphanumeric variables.

- **Missing/Don't know**: This code always has a value consisting of all 9's.   The number of 9's is determined by the length of the variable. For instance, if the variable length is 3, then the variable is assigned a '999' code.

Standard Shading

As shown below, the column headers are always shaded as well as the row that contains the actual question or description that corresponds with the variable.

Question No.	Variable			Question	Code
	Name	**Type**	**Size**		

				Address status	
	MHSTATUS	N	1	Household moved and address is known	1
				Household moved but address is unknown	2
				No information is known	4

Restricted-Use Shading

There are variables described in the codebook which are only available in the restricted-use data files. For these variables, all of the columns – question number, variable name, question and code – and all of the codes and labels have been completey shaded:

				District to which the migrant moved.	
	MHDIS	N	4	See District Appendix	112 – 6713
				Missing/Don't know	9999

Some of the restricted-use variables are recoded for the public release data files, while others are not available in any form. The variable above (MHDIS) is not available in any form in the public release data files. When a restricted-use variable is recoded for the public release data files, then a note and the public release codes will follow the variable:

				1984 Village Number	
	VILL84	A	2	Village number from the 1984 household questionnaire identifier.	01 - 51
				1984 House Number	
	HOUSE84	A	3	A 3-digit sequence number that is part of the 1984 household questionnaire identifier.	001 – 184

The VILL84 and HOUSE84 variables are not in the public use data, but they have been replaced with one variable called PHHID84. The values for this new variable were randomly generated based on the original values from VILL84 and HOUSE84:

Public Use Codes	PHHID84	N	4	Randomly generated sequence number	1 – 5860

Identifier Variables for the Community and Household Data in 1984, 1994 and 2000

Understanding the identifier variables is vital in working with the data files and merging them appropriately. Analysts will need to consistently use three different types of identifiers – village, household and individual - in working with the community and household data files. Each identifier type represents a different level of observation and detail and provides the key to the most important mechanism for matching and linking villages, households and people. Below, there is a description for each identifier type, which is then followed by the identifier variable names (both restricted- and public-use data) for each survey year:

- **Village** - A unique number was assigned, during or just prior to the data collection, to villages that existed within the boundaries of the 1984 Nang Rong District.

	1984 Survey	1994 Survey	2000 Survey
Restricted	VILL84	VILL94	VILL00
Public	n/a	n/a	n/a

- **Household** - Each household received <u>two</u> unique identifiers. The first identifier was assigned by an interviewer within the household's respective village as the village list was being compiled. The second identifier was written on the household's questionnaire by the interviewer. This identifier includes the village number and the house number (or unique sequential number in 1994).

	1984 Survey	1994 Survey	2000 Survey
Restricted #1	HOUSE84	LEKTI94	HOUSE00
Restricted #2	HHID84	HHID94	HHID00
Public	PHHID84	PHHID94	PHHID00

- **Individual** - The individual identifiers (CEP codes) are found on the household roster. The CEP code was assigned by the interviewer as the household roster was being updated during the interview.

	1984 Survey	1994 Survey	2000 Survey
Restricted	CEP84	CEP94	CEP00
Public	CEP84	CEP94	CEP00

Given the three identifier types, the following table shows what identifier variables are needed to uniquely identify a village, a household and an individual in both the restricted- and public-use data:

Community and Household Survey Data

To uniquely identify a …	#	Restricted-Use Data			Public-Use Data		
		1984 Survey	1994 Survey	2000 Survey	1984 Survey	1994 Survey	2000 Survey
Village	--	VILL84	VILL94	VILL00	n/a	n/a	n/a
Household	1	VILL84 HOUSE84	VILL94 LEKTI94	VILL00 HOUSE00	n/a	n/a	n/a
	2	HHID84	HHID94	HHID00	PHHID84	PHHID94	PHHID00
Individual	1	VILL84 HOUSE84 CEP84	VILL94 LEKTI94 CEP94	VILL00 HOUSE00 CEP00	n/a	n/a	n/a
	2	HHID84 CEP84	HHID94 CEP94	HHID00 CEP00	PHHID84 CEP84	PHHID94 CEP94	PHHID00 CEP00

Identifier Variables for the Migrant Follow-Up Data in 1994 and 2000

The Migrant Follow-Up data has only two identifier types – household and individual.   Please note, however, that the household survey identifiers are part of the Migrant Follow-Up data to allow linkages between the migrant and household data.   The two migrant identifier types represent a different level of observation and detail that provides the key to the most important mechanism for matching and linking migrant households and people. Below, there is a description for each identifier type, which is then followed by the identifier variable names (both restricted- and public-use data) for each survey year – 1994 and 2000:

- **Household** – Each migrant household had an identifier written on the questionnaire by the interviewer. This identifier includes the village number from which the migrant originated and a unique sequential number.   In the 2000 data, there were two phases of the Migrant Follow-Up data collection – the urban phase and the rural phase.   Since the household identifier (MID00) could be the same across the two phases, a second variable (MIGTYPE), signifying the migrant type, is needed to uniquely identify the household.

	1994 Survey	2000 Survey
Restricted	MID	MID00, MIGTYPE
Public	PMID94	PMID00

- **Individual** - The individual identifiers (MCEP codes) are found on the migrant's household roster. The MCEP code was assigned by the interviewer as the migrant's household roster was being updated during the interview.

	1994 Survey	2000 Survey
Restricted	MCEP8	MCEP00
Public	MCEP8	MCEP00

Given these two migrant identifier types, the following table shows what identifier variables are needed to uniquely identify a migrant household and an individual in both the restricted- and public-use data:

Migrant Follow-Up Data

To uniquely Identify a …	Restricted-Use Data		Public-Use Data	
	1994 Survey	2000 Survey	1994 Survey	2000 Survey
Migrant Household	MID	MID00 MIGTYPE	PMID94	PMID00
Individual	MID MCEP8	MID00 MIGTYPE MCEP00	PMID94 MCEP8	PMID00 MCEP00

Codebook for the 1994 Moved Households

Question No.	Variable			Question	Code
	Name	Type	Size		
	VILL84	A	2	**1984 Village Number** Village number from the 1984 household questionnaire identifier.	01 - 51
	HOUSE84	A	3	**1984 House Number** A 3-digit sequence number that is part of the 1984 household questionnaire identifier.	001 – 184

The VILL84 and HOUSE84 variables are not in the public use data, but they have been replaced with one variable called PHHID84. The values for this new variable were randomly generated based on the original values from VILL84 and HOUSE84:

Public Use Codes	PHHID84	N	4	Randomly generated sequence number	1 – 5860

	HHTYPE94	N	1	Household Type Moved Old Household N/A – individual is "unaccounted for"	 4 [.]

	MIGVILL	N	1	Sample Migrant Village Not a sample migrant village Sample migrant village	 0 1

Question No.	Variable			Question	Code
	Name	Type	Size		

Identifiers from the 2000 Household Survey

The following identifier variables are from the 2000 Household Survey.   These variables can be used to link an individual to the 2000 Survey data files. Please note that not all individuals can be linked to the 2000 data files. These individuals are coded as "N/A" on the identifiers.

Identifiers from the 2000 Survey

				2000 Household Identifier	
	HHID00	A	9	Household ID from Questionnaire in 2000 N/A	Village # + House # [ ]

The HHID00 variable is not in the public use data, but it has been replaced with a variable called PHHID00. The values for this new variable were randomly generated based on the original values from HHID00:

Public Use Codes	PHHID00	N	4	Randomly generated sequence number N/A	1 – 9415 [.]
				2000 CEP Number	
	CEP00	A	2	Person Identifier within the 2000 household N/A	01 – 20 [ ]

Note: By definition, all individuals in the MOVED94 data file do not have links to the 1994 Survey data. These are individuals who did not participate in the 1994 data collection, since their entire household moved out of the village.   Therefore, this data file does not have the 1994 identifier set.

Question No.	Variable			Question	Code
	Name	Type	Size		
	CEP84	A	3	1984 CEP Number (Person identifier within the 1984 household) Current household member in 1984 Temporarily absent household member in 1984	101 – 116 201 – 209
	MHTYPE	N	1	Type of Moved Household Individual An individual who was part of an entire 1984 household that was no longer living in the village during the 1994 data collection. An individual who "not accounted for" during the 1994 data collection even though the old 1984 household was interviewed.	1 2
	MHSTATUS	N	1	Address status Household moved and the address is known Household moved but the address is unknown No information is known	1 2 4
	MHDIS	A	4	Destination information. District to which the migrant moved. See District Appendix Missing/Don't know	9999
	MHPRO	A	2	Province to which the migrant moved. See (2-digit) Province Appendix Missing/Don't know	99
	MHADD	N	1	Clearly specified home or work address Yes No Missing/Don't know	1 2 9
	MHTEL	N	1	Clearly specified home or work phone number Yes No Missing/Don't know	1 2 9

Flag and Status Variables, created at CPC, to assist with Data Analysis

Question No.	Variable			Question	Code
	Name	Type	Size		

For Individuals who link to the 2000 Survey Data Only

				Individual's 2000 Survey Form Status	
	FORM00	A	12	Individuals on Form 1 who are coded as either X1=0, 1, 3 or 9.	F1
				Code 2 Individuals (X1=2) who were tracked to Form 2 of another Old HH.	F1&F2&CODE2
				Code 2 Individuals (X1=2) who were tracked to Form 3 of another Old HH.	F1&F3&CODE2
				Individuals on Form 2 who actually link back to person in 1984 or 1994, but is NOT Code 2.	F2-OLD
				Code 2 Individuals (X1=2) on Form 1 who could not be tracked to a destination hh.	CODE2-F1ONLY
				Individuals on Form HM whose entire household moved out of the village.	FHM
				N/A	[ ]