Nang Rong Projects

Codebook for the Person Identifier Data File

Carolina Population Center, University of North Carolina at Chapel Hill

February 21, 2005

Getting Started with the Person Identifier Data File

The Data File Description
The person identifier data file is a constructed data file that contains all individuals with all of their household and individual identifiers (1984, 1994 and 2000), all of their migrant follow-up identifiers (1994 and 2000), and one person identifier that is unique to all individuals.   Each individual is listed in the data file, with their identifiers, **one time**.

The "Duplicate" Individuals
Although each individual appears in the person identifier data file one time, there are groups of individuals who appear in the household survey and migrant follow-up data files **two times**:

Code 2 Individuals
The duplicated individuals in the 1994 and 2000 household surveys are called code 2 individuals. They first appear on their origin household roster (formerly resided there) with one set of identifiers and coded 2 on the Q1.1 variable. And then, they appear on their destination household roster (currently residing there) with another set of identifiers and coded as 1 or 4 on the Q1.1 variable. As a result, the code 2 individuals have two sets of identifiers within the one survey year. For the person identifier data file, the main set of identifers correspond to the destination household where the individual currently resides. However, a second set of identifiers has been added to the person identifier data file.   This second set corresponds to the origin household, and the variable names begin with the letter "O". In 1994, there are *2,516* code 2 individuals who appear on two rosters, while in 2000, there are *2,511*.

One Individual, Two Migrant Rosters
The duplicated individuals in the 1994 and 2000 migrant follow-up surveys are individuals who appear on two migrant household rosters.   This is a rare occurrence in both surveys with only 3 individuals in 1994 and 14 individuals in 2000.   In 2000, this type of situation was anticipated and 9 of the 14 cases have a code of "2" on IVSTATUS, which is the final interview status variable. The code of "2" means the individual was found on Form 8 of another migrant questionnaire.   In the other 5 cases in 2000 and for all 3 cases in 1994, the individual was listed on both migrant rosters but did not interview in one of the two migrant households.   So, for the person identifier data file, the main set of identifiers correspond to the migrant household where the individual interviewed or was not coded 2 on IVSTATUS. However, a second set of identifiers has been added to the data file. This second set corresponds to the other migrant household, and the variable names end with the letter "B".

The Person Identifier Variable
The person identifier variable is a sequential number, in character format, that has a length of 6 digits. The first person is coded as "000001"; the second person "000002"; etc., up to and including the last person coded as "057416."   Please note from the examples that the identifier is zero-filled. The name of the person identifier variable is NRPID which stands for "Nang Rong Person Identifier."

The Data File Specifications
The person identifier data file has the following specifications:

Data Set Name	N	# of Variables	1 obs =	Sorted by
personid	57,416	26	individual	NRPID

Codebook for the Person Identifier Data File

Household and Individual Identifier Variables

Variable			Description	Code
Name	**Type**	**Size**		
			Nang Rong Person Identifier	
NRPID	A	6	A unique 6-digit sequence number that uniquely identifies an individual who participated in either the household and/or migrant follow-up phases of the Nang Rong Projects data collections in 1984, 1994 and/or 2000.	000001 – 057416

2000 Household Survey Identifier Variables

Variable			Description	Code
			Household ID from 2000 Questionnaire	
HHID00	A	9	• First 6 digits are 2000 village number • Last 3 digits are 2000 house number N/A	Village # + House # [ ]
			2000 CEP Number	
CEP00	A	2	2-digit sequence number which is the person identifier within the 2000 household N/A **For individuals coded '2' on question 1.1 only**	01 – 20 [ ]
			2000 Origin Household Identifier	
OHHID00	A	9	HHID00 of the origin household from which the code 2 individual moved since the last household survey (1984 or 1994) N/A	Village # + House # [ ]
			2000 Origin CEP Number	
OCEP00	A	2	CEP00 in the origin household from which the code 2 individual moved since the last household survey (1984 or 1994) N/A	01 – 20 [ ]

Variable			Description	Code
Name	**Type**	**Size**		

1994 Household Survey Identifier Variables

Variable			Description	Code
HHID94	N	8	Household ID from 1994 Questionnaire • First 4 digits are 1994 village number • Next 3 digits are sequence number • Last digit is type of household (1=old household, 2=new household) N/A	[.]
CEP94	A	3	1994 CEP Number Person identifier within the 1994 household N/A **For individuals coded '2' on question 1.1 only**	101 – 118, 201 – 205 401 – 494, 501 – 511 [ ]
OHHID94	N	8	1994 Origin Household Identifier HHID94 of the origin household from which the code 2 individual moved since the 1984 household survey N/A	[.]
OCEP94	A	3	1994 Origin CEP Number CEP94 in the origin household from which the code 2 individual moved since the 1984 household survey N/A	101 – 118, 201 – 205 [ ]

1984 Household Survey Identifier Variables

Variable			Description	Code
VILL84	A	2	1984 Village Number A unique 2-digit number that identifies a village. N/A	01 – 51 [ ]
HOUSE84	A	3	1984 House Number A unique 3-digit sequence number within the 1984 village. N/A	001 – 185 [ ]
CEP84	A	3	1984 CEP Number Person identifier within the 1984 household N/A	101 – 118, 201 – 209 [ ]

Migrant Follow-Up Identifier Variables

Variable			Description	Code
Name	**Type**	**Size**		

2000 – 2001 Migrant Follow-Up Survey Identifier Variables

			Migrant Follow-Up ID from 2000 Questionnaire	
MID00	A	11	• First 6 digits = 2000 Origin Village Number • Last 5 digits = Sequence Number N/A	Village # + Sequence # [ ]
			2000 Migrant Type	
MIGTYPE	A	4	Rural to urban migrant Rural to urban migrant who returned home Rural to rural migrant Rural to rural migrant who returned home N/A	R2U R2UR R2R R2RR [ ]
			2000 MCEP code	
MCEP00	A	2	2-digit sequence number which is the person identifier in the 2000 migrant household N/A **For individuals (N=14) listed on two migrant household rosters only**	01 – 10 [ ]
			Second 2000 Migrant Follow-Up Identifier	
MID00B	A	11	MID00 of a second migrant household roster on which the individual is listed N/A	Village # + Sequence # [ ]
			Second 2000 Migrant Type	
MIGTYPEB	A	4	MIGTYPE of the individual in a second migrant household roster. N/A	 [ ]
			Second 2000 MCEP code	
MCEP00B	A	2	MCEP00 in a second migrant household roster on which the individual is listed N/A	01 – 10 [ ]

3

Variable			Description	Code
Name	Type	Size		

1994 – 1995 Migrant Follow-Up Survey Identifier Variables

Variable			Description	Code
MID	N	7	Migrant Follow-Up ID from 1994 Questionnaire • First 4 digits = 1994 Origin Village Number • Last 3 digits = Sequence Number N/A	Village # + Sequence # [.]
MCEP8	A	3	1994 MCEP code 3-digit sequence number which is the person identifier in the 1994 migrant household N/A **For individuals listed (N=3) on two migrant household rosters only**	801 – 810 [ ]
MIDB	N	7	Migrant Follow-Up ID from 1994 Questionnaire MID of a second migrant household roster on which the individual is listed N/A	Village # + Sequence # [.]
MCEP8B	A	3	1994 MCEP code MCEP8 of a second migrant household roster on which the individual is listed N/A	801 – 810 [ ]

Status and Flag Variables

Variable			Description	Code
Name	Type	Size		
			Individual's 2000 Household 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 are new to the household and the village.	F2
			Individuals on Form 2 who actually link back to person in 1984 or 1994, but is NOT Code 2.	F2-OLD
			Individuals on Form 3 who are new to the household and the village.	F3
			Individuals on Form 3 who actually link back to person in 1984 or 1994, but is NOT Code 2.	F3-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	[ ]

Variable			Description	Code
Name	Type	Size		
			Individual's 1994 Household Survey Form Status	
FORM94	A	12	Individuals on Form 1 who are coded as either Q1=0, 1, 3 or 9.	F1
			Code 2 Individuals (Q1=2) who were tracked to Form 2 of another Old HH.	F1&F2&CODE2
			Code 5 Individuals (Q1=5) who were tracked to Form 2 of another Old HH.	F1&F2&CODE5
			Individuals from Combined Old HHs who were originally on Form 2 of the primary hh and Form 1 of the secondary hh.	F1&F2COMHH
			Code 2 Individuals (Q1=2) who were tracked to Form 3 of another Old HH.	F1&F3&CODE2
			Code 5 Individuals (Q1=5) who were tracked to Form 3 of another Old HH.	F1&F3&CODE5
			Individuals on Form 2 who are new to the household and the village.	F2
			Individuals on Form 2 who actually link back to person in 1984, but is NOT Code 2.	F2-OLD
			Individuals on Form 3 who are new to the household and the village.	F3
			Individuals on Form 3 who actually link back to person in 1984, but is NOT Code 2.	F3-OLD
			Code 2 Individuals (Q1=2) on Form 1 who could not be tracked to a destination hh.	CODE2-F1ONLY
			N/A	[ ]

Variable			Description	Code
Name	**Type**	**Size**		
			Identifiers flag variable that indicates which survey years the individual has household and individual identifiers.	
IDSFL	A	6	1984, 1994 and 2000 1984 and 1994 1984 and 2000 1984 only 1994 and 2000 1994 only 2000 only N/A	849400 8494__ 84__00 84____ __9400 __94__ ____00 [ ]
			Migrant identifiers flag variable that indicates which survey years the individual has migrant identifiers.	
MIGIDSFL	A	4	1994 and 2000 1994 only 2000 only N/A	9400 94__ __00 [ ]