

Coding Instructions for Sociomatrices  
 Siblings, 2000 Household Interview  
 November 18, 2002  
 September 9, 2002

file: coding\_siblings.doc

This memo describes how the sociomatrices for sibling networks are constructed for the 2000 household data.

### **Relevant Questions from 2000 Household Interview:**

Form 4 asks about siblings for individuals in the household who are 18-41 years old and any resident spouse of one of these individuals, regardless of age.

Variables X4\_4\_1 to X4\_4\_12 record CEP00 codes for siblings:

For “old” households, siblings that are on Forms 1, 2, or 3.

For “new” households, siblings that are on Form 3.

Variables X4\_5R1 to X4\_5R16 record the locations of other siblings:

For “old” households, siblings that are not on Forms 1, 2, or 3.

For “new” households, siblings that are not on Form 3.

The location can be:

This village

Another village in Nang Rong

Another district in Buriram

Another province

Another country

N/A

Missing/Don’t know

Variable X3 codes the respondent’s age.

Variables X4\_5A1 to X4\_5A16 record siblings’ ages.

### **Relations and Sociomatrices**

The sibling relation records the number of individuals in a location (household, village, district, or province) who are named as siblings by members of the household, where respondent and sibling are in specified age ranges (as described below).

There are five sets of matrices coding sibling ties. These sets restrict the ages of respondents and/or named siblings in different ways. The first three sets restrict the ages of respondents but not of named siblings:

1. The first set has responses from all 18 to 41 year olds in the household and all named siblings.
2. The second set has responses from 18 to 35 year olds (the same age range as used for the 1994 sociomatrices) and all named siblings.
3. The third set has responses from 24 to 41 year olds (the cohort used for the 1994 sociomatrices) and all named siblings.

The other two sets of matrices restrict the ages of both respondents and of named siblings. We do this to investigate the mutuality of naming siblings.

4. The fourth set has responses from 18 to 35 year olds in the household and named siblings who are also 18 to 35 years old.
5. The fifth set has responses from 24 to 41 year olds and named siblings who are also 24 to 41 years old.

The number of siblings is accumulated for siblings named from the same location (household, village, district, or province). The values of the ties in the sociomatrix are counts of the number of individuals (either total or in the specified age range) from each location who are named as siblings of people in the specified age range in the household.

Sociomatrices are organized at the village level. There are 51 villages, corresponding to the 1984 villages. In all sociomatrices the rows are the households responding to the survey and the columns are the possible locations of siblings. Ties are recorded as siblings from other households in the village, other villages in Nang Rong district, other districts in Buriram province, or other provinces in Thailand (including ties abroad).

There are four kinds of sociomatrices for this relation:

$X^v$ : household by household, for siblings nominated by households in village  $v$  who are in other households in village  $v$

$Y^v$ : household by village, for siblings nominated by households in village  $v$  who are in other villages within Nang Rong district

$Z^v$ : household by district, for siblings nominated by households in village  $v$  who are in other districts in Buriram province

$W^v$ : household by province, for siblings nominated by households in village  $v$  who are in other provinces outside Buriram province, including outside Thailand

Since sociomatrices are constructed for each of five different age restrictions, there are  $5 \times 4 = 20$  sociomatrices for each village.

### **Coding Siblings in Unknown Locations or Abroad**

There are four kinds of anomalous data that arise when siblings are in unknown location or abroad. These are coded in extra columns in either the household by household sociomatrix or the household by province sociomatrix. The household by household

matrix has one extra column and the household by province matrix has three extra columns.

- 1) *Siblings within the village but the household id is unknown.* This is coded as one extra column in the household by household matrix
- 2) *Siblings living abroad.* This is coded in the 1<sup>st</sup> extra column in the household by province matrix.
- 3) *Siblings outside the village but the location is unknown.* This is coded in the 2<sup>nd</sup> extra column in the household by province matrix.
- 4) *Siblings in an unknown location that could be either inside or outside the village (it was impossible to track down the location of the sibling).* This is coded in the 3<sup>rd</sup> extra column in the household by province matrix.

### **Diagonal Entries**

Diagonal entries in the household by household sociomatrix are equal to zero.

### **Missing Data:**

For comparability with 1994 coding,

If CEP=99, then code as an unknown location.

If there are duplicate CEP codes, use one and ignore the other.

If CEP does not match form 1, then code as an unknown location.

### **Row labels**

Row labels for households in all sociomatrices will be formed using the 2000 village number followed by the house number (variable HHID00; 9 digits).

### **Column labels**

#### Households

Same format as row labels (9 characters)

9999999 for the extra column (n + 1<sup>st</sup> column, n = # of households in the village)

for help from within the village, household id unknown

#### Villages

2000 village numbers (6 characters)

#### Districts

2000 district numbers (2 characters)

#### Provinces

2000 province numbers (2 characters)

Extra columns in the household by province matrix

col \_\_ : 98 for help from people living abroad

col \_\_: 99 for help from outside village, location unknown

col \_\_: 97 missing data, location unknown (help might be from inside or outside village)

### **Data Format**

line 1: 1984 village number (51 villages)  
 line 2: number of rows in the sociomatrix  
 line 3: number of columns in the sociomatrix  
 line(s) 4: list of row labels, separated by spaces  
 line(s) 5: list of column labels, separated by spaces  
 remaining rows, one for each household in the village: sociomatrix entries separated by spaces

### **Differences Between 1994 and 2000 Sociomatrices**

In the 1994 sociomatrices diagonal entries coded the total number of siblings in sibsets in the household. In 2000 sociomatrices diagonal entries are equal to zero.

1994 sociomatrices used respondents age 18-35. Two sets of sociomatrices were constructed: one set included all named siblings, regardless of age; the second set included only named siblings in the age range 18-35. For 2000 there are five sets of sociomatrices, with different ways of restricting ages of respondents and of named siblings (see description above).

### **Coding Instructions:**

The instructions for coding sibling ties are in two parts, those that pertain only to “old” households and those that pertain to both “old” and “new” households. Both sets of instructions are used to construct the sociomatrices for this relation.

The number of siblings is accumulated for persons in the same location (household, village, district, or province).

The value of the entry in the sociomatrix is a count of the number of individuals that respondents in the household (in the row of the sociomatrix) report from the location (in the column of the matrix). If multiple siblings are from the same location they are accumulated into the total.

### *Relevant questions*

#### *1. For “old” households*

The sibling form records locations of siblings for all eligible respondents in the household. Form 4 asks about sibling networks for individuals in the households who are 18-41 years old and any resident spouse of one of these individuals, regardless of age.

Variables X4\_4\_1 to X4\_4\_12 record CEP00 codes for siblings who are on Forms 1, 2, or 3.

Variables X4\_5R1 to X4\_5R16 record the locations of siblings who are not on Forms 1, 2, or 3.

The location can be:

- This village
- Another village in Nang Rong
- Another district in Buriram
- Another province
- Another country
- N/A
- Missing/Don't know

For code 2 people, find their location in DHHID00 (verified destination household for code 2 people).

For code 2 people, if the person is in temple, code the location as unknown but in the village.

For code 3 people, find their location in 1.9b (X9B2VILL, X9B2DIST, X9B2PROV, X9B2COUN)

Notes:

For comparability with 1994 coding,

If CEP=99, then code as an unknown location.

If there are duplicate CEP codes, use one and ignore the other.

If CEP does not match form 1, then code as an unknown location.

## 2. *For "new" households*

The sibling form records locations of siblings for all eligible respondents in the household. Form 4 asks about sibling networks for individuals in the households who are 18-41 years old and any resident spouse of one of these individuals, regardless of age.

Variables X4\_4\_1 to X4\_4\_12 record CEP00 codes for siblings that are on Form 3 (in the household)

Variables X4\_5R1 to X4\_5R16 record the locations of siblings that are not on Form 3.

The location can be:

- This village
- Another village in Nang Rong
- Another district in Buriram
- Another province
- Another country
- N/A
- Missing/Don't know

Once their location is identified, each person will be added to the count of the number of individuals from that location. The value of the tie from household  $i$  to location  $j$  should be incremented by 1 for each sibling from location  $j$ .

The location of the person is either 1) another household in the village 2) another village in Nang Rong, 3) another district in Buriram, 4) another province, or 5) abroad.

*a. Another household in the village*

If the location is another household in the village then the tie is recorded in the household by household sociomatrix.

$j$  = the position of the household in an ordered list of households in the village

$$x_{ij}^v = x_{ij}^v + 1$$

(comment: an individual from household  $i$  says they have a sibling in household  $j$  in the village, so the tie is incremented by 1 in the household by household sociomatrix)

- For the first set of matrices record all responses from all 18 to 41 year olds in the household and all named siblings.
- For the second set of matrices record responses from 18 to 35 year olds and all named siblings.
- For the third set of matrices record responses from 24 to 41 year olds and all named siblings.
- For the fourth set of matrices record responses from 18 to 35 year olds in the household and named siblings who are also 18 to 35 years old.
- For the fifth set of matrices record responses from 24 to 41 year olds and named siblings who are also 24 to 41 years old.

*b. Unknown location in the village*

If the location is unknown but it is inside the village then 1 is added to the total in the extra column in the household by household matrix

$$x_{i(n+1)}^v = x_{i(n+1)}^v + 1$$

(comment: the additional column in the household by household sociomatrix will count the number of siblings in the village but with unknown household id)

*c. Another village in Nang Rong*

If the location is another village in Nang Rong district then the tie is recorded in the household by village sociomatrix.

$j$  = the position of the village in an ordered list of villages

$$y_{ij}^v = y_{ij}^v + 1$$

(comment: an individual from household  $i$  says they have a sibling in village  $j$  so the tie from  $i$  to  $j$  is incremented by 1)

- For the first set of matrices record all responses from all 18 to 41 year olds in the household and all named siblings.
- For the second set of matrices record responses from 18 to 35 year olds and all named siblings.
- For the third set of matrices record responses from 24 to 41 year olds and all named siblings.
- For the fourth set of matrices record responses from 18 to 35 year olds in the household and named siblings who are also 18 to 35 years old.
- For the fifth set of matrices record responses from 24 to 41 year olds and named siblings who are also 24 to 41 years old.

*d. Another district in Buriram*

If the location is in another district then the tie is recorded in the household by district sociomatrix

$j$  = the position of the district in an ordered list of districts

$$z_{ij}^v = z_{ij}^v + 1$$

(comment: an individual from household  $i$  says they have a sibling in district  $j$  so the tie from  $i$  to  $j$  is incremented by 1)

- For the first set of matrices record all responses from all 18 to 41 year olds in the household and all named siblings.
- For the second set of matrices record responses from 18 to 35 year olds and all named siblings.
- For the third set of matrices record responses from 24 to 41 year olds and all named siblings.
- For the fourth set of matrices record responses from 18 to 35 year olds in the household and named siblings who are also 18 to 35 years old.
- For the fifth set of matrices record responses from 24 to 41 year olds and named siblings who are also 24 to 41 years old.

*e. Another province.*

If the location is another province then the tie is recorded in the household by province matrix

$j$  = the position of the province in an ordered list of provinces

$$w_{ij}^v = w_{ij}^v + 1$$

(comment: an individual from household  $i$  says they have a sibling in province  $j$  so the tie from  $i$  to  $j$  is incremented by 1)

- For the first set of matrices record all responses from all 18 to 41 year olds in the household and all named siblings.
- For the second set of matrices record responses from 18 to 35 year olds and all named siblings.
- For the third set of matrices record responses from 24 to 41 year olds and all named siblings.
- For the fourth set of matrices record responses from 18 to 35 year olds in the household and named siblings who are also 18 to 35 years old.
- For the fifth set of matrices record responses from 24 to 41 year olds and named siblings who are also 24 to 41 years old.

*f. Abroad*

If the location is abroad then 1 is added to the total in the 1st extra column in the household by province matrix.

$$w_{i(q+1)}^v = w_{i(q+1)}^v + 1$$

*g. Unknown location outside the village*

If the location is outside the village but the location is unknown 1 is added to the total in the 2<sup>nd</sup> extra column in the household by province matrix

$$w_{i(q+2)}^v = w_{i(q+2)}^v + 1$$

*h. Unknown location (impossible to find a location for the person)*

If the sibling is in an unknown location that could be either inside or outside the village (in other words, it was impossible to find a location for the person), then 1 is added to the total in the 3<sup>rd</sup> extra column in the household by province matrix.

$$w_{i(q+3)}^v = w_{i(q+3)}^v + 1$$