GÖTEBORGS UNIVERSITET
INSTITUTIONEN FÖR PEDAGOGIK OCH DIDAKTIK

GOLD
Gothenburg Educational Longitudinal Database

IMMIGRATION, NATIONAL ORIGIN AND ADOPTION – VARIABLES IN THE GOLD DATABASE

Jan-Eric Gustafsson & Bo Nielsen

Immigration, national origin and adoption –
Variables in the GOLD-database

Background

The immigration and national origin variables are a constructed variable frequently used in studies that uses data from the GOLD-database at Gothenburg University. The GOLD – population that originates from the VALUTA-project consists of all persons living in Sweden at the age of 16 born in the period 1972-1987.

In Sweden the immigration has increased over time and in 2004 about 15% of the population has another background than Swedish. However the term 'foreign background' is defined and interpreted in several different ways. In the context of the GOLD-project it is investigated, among other things, how this variable influences school achievement and educational careers.

Statistics Sweden has published directions how to code national origin in scientific studies in Sweden. These directions are based on a person's immigration status and citizenship (SCB 2002). However, for the GOLD-project this is not sufficient because second-generation immigrants are not coded in this model. This group is very important to study, therefore the coding of the GOLD-project differs from the coding suggested by Statistics Sweden.

Statistics Sweden guidelines

In the guidelines two variables with two categories each are recommended. One concern where the person is born. Born in Sweden or born outside Sweden. The other variable also distinguishes between Swedish citizenship and other citizenship. The recommended variable is the latter with an expanded description, see table 1, (Statistics Sweden, 2002).

Table 1. The Statistics Sweden definition of foreign and Swedish background.

A Persons with foreign background	B Persons with Swedish background
Born outside Sweden with two parents born outside Sweden	Born outside-Sweden with at least one parent born in Sweden
Born in-Sweden with two parents born outside Sweden	Born in-Sweden with one parent born in Sweden and one parent born outside Sweden
	Born in-Sweden with two parents born in Sweden

These criteria take the origin of the parents into account but are only based on the biological parents.

Gold-data

Background

In the Gold-project a lot of the data on the parents are missing and using only information about biological parents does not account for the influence of the family environment on the children. Therefore, data was included about the parents the child was living with at the age of 16 (these will be referred to as 'house parents'). This makes it possible to identify adopted children from another country as well as other small groups. In some cases the information

about the parent is missing; however nearly all missingness is due to late immigration in Sweden and other foreign births.

The variable logical national background

The following coarse categories were starting points for the classification:

1. The child was born in Sweden and at least one of the house parents was born in Sweden. These are considered to be Swedish.
2. The child was born outside Sweden and both house parents were born in Sweden. This group primarily consists of adopted children from abroad and is considered to be Swedish.
3. The child was born in Sweden and both parents were born outside Sweden. This group primarily consists of second-generation immigrants.
4. The child was born outside Sweden, and at least one parent was born outside of Sweden. These generally are the first generation of immigrants.

This variable had a large amount of missingness because of missing information about the parents

To obtain a more exhaustive classification all combinations of the immigration status of the child, the biological parents and the house parents were created as a composite string variable. The national origin were coded as S = born in Sweden, F = born outside Sweden and M= missing data. From this composite variable, which comprised 156 categories, the status of the child was categorized into ten groups (see table 2 and 3).

Categories

In the database there are six different variables representing foreign/Swedish background: one with 10 groups; three in which these 10 groups have been reduced to five, four and two groups, respectively; and two in which the latter two variables have been expanded with categories showing age of immigration for immigrating persons. Table 2 shows the different variables in relation to each other. The variables are named floinv10(named floinv9), floinv5, floinv4, floinv2, floinv4t and floinv2t in the database.

The ten primary groups are discussed below:

1. CSPS. The child was born in Sweden and at least one parent was born in Sweden. If data on the biological parents are missing the data concerning the house parents are used. N = 1 444 832 and the number of categories of the combination variable (ncv) are 46. This category is considered to be persons of Swedish origin.
2. CSPI. The child was born in Sweden and both of the biological parents are immigrants and at least one of the house parents has immigrated. If information is missing it is assumed that the parent is born outside Sweden. N= 84 154 and ncv = 27. This category consists of second-generation immigrants.
3. CIPS. The child was born outside Sweden and both biological parents were born outside Sweden or data is missing. Furthermore, both house parents are born in Sweden or data is missing on one parent. N = 23 053 and ncv = 12. This category primarily consists of adopted children from outside Sweden and they are treated as persons with Swedish origin.
4. CIPI. The child was born outside Sweden and both the biological and the house parents were born outside Sweden or data is missing on one parent. N = 106 540 and nmg = 32. This category consists of first generation immigrants.

5. CIBS. The child was born outside Sweden and both the biological parents and the house parents were born in Sweden. Some combinations with missing data for one parent or one parent born outside Sweden are included. N = 1 879 and ncv = 10. This category includes children to Swedes working abroad. The children are treated as Swedish.

6. CSBSIPI. The child was born in Sweden and one of the biological parents has immigrated and the other was born in Sweden. Furthermore, both house parents have immigrated or data is missing on one of the house parents. N = 8 771 and ncv = 8. This category primarily consists of second-generation immigrants.

7. CIPM. The child was born outside Sweden and data is missing on biological as well as house parents. N = 3 516 and the ncv = 1. This category consists of first generation immigrants and adopted persons to 29%.

8. CIBSMPM. The child was born outside Sweden and one of the biological parents was born in Sweden and data on the other is missing. Data on house parents are missing or includes a parent born outside Sweden. N = 503 and ncv=4. This category is considered to be first generation immigrants based on the assumption that missing data indicate a parent born outside Sweden.

9. CIBSI. The child was born outside Sweden and one of the biological parents was born in Sweden and the other was born outside Sweden. The same pattern applies to the house parents. Variations in the pattern with missing data for one of the parents, or that both house parents were born in Sweden occurs. N = 3 281 and ncv = 11. This category primarily includes children to emigrated parents or one parent that is working abroad. The group shows many similarities with the CIBS group and is treated as Swedes.

10. CIBSMPSM. The child was born outside Sweden and one parent was born in Sweden and the data is missing on the other. N = 712 and ncv = 5. This group is very similar to CIBSI and CIBS and is also treated as Swedes.

Time variable

In the four-group and two-group variables the time of immigration is included in two additional variables. The time categories have been chosen to cover the main school events. Second generation immigrants have been assigned immigration time zero and first generation immigrants has been divided in three groups: immigration below seven years of age; immigration between the age of seven and below thirteen years; and finally immigration from thirteen years of age up to the year they are sixteen. The upper limit is because the population is defined as persons being in Sweden at the age of sixteen and born between 1972 and 1987. This means that immigrants arriving after the age of sixteen are excluded from the population.

The construction of the time dependent variable has as an effect that some immigrants, (particularly adopted children and children of Swedish parents working abroad) do not have any immigration time. It is in this respect that our coding differs mostly from the guidelines of Statistics Sweden. Statistics Sweden acknowledges the latter group as existing but says that the group is so small that they do not consider treating them as anything other than immigrants. Adopted children are by Statistics Sweden treated as immigrants. However, the household is treated as Swedish as long as the child lives at home but treated as immigrants when the child moves out (Statistics Sweden, 2002).

Adoption

Data from the adoption register are available. In total 1.9 % of the total population has been adopted. The variable is coded to show those born in Sweden (about 21%) and those born outside Sweden (about 79%). Table 4 shows the proportion of adopted in the migration background groups.

References

Statistics Sweden. 2002. Reports on Statistical Co-ordination for the Official Statistics of Sweden 2002:3. Statistics on persons with foreign background. Guidelines and recommendations.

Table 2. Scheme over the six origin variables in the GOLD-project.

10 – groups	5 groups	4 groups	2 groups	4 groups and time	2 groups and time
CSPS	CSPS	CSPS	Swedish	CSPS	Swedish
CSPI	CSPI	CSPI	Immigrant	CSPI time 0[1]	Immigrant time [1]
CIPS	CIPS	CIPS	Swedish	CIPS	Swedish
CIPI	CIPI	CIPI	Immigrant	CIPI time [2]	Immigrant time [2]
CIBS	CIBS	CSPS	Swedish	CSPS	Swedish
CSBSIPI	CSPI	CSPI	Immigrant	CSPS	Immigrant time [1]
CIPM	CIPI	CIPI	Immigrant	CIPI time [2]	Immigrant time [2]
CIBSMPM	CIPI	CIPI	Immigrant	CIPI time [2]	Immigrant time [2]
CIBSI	CIBS	CSPS	Swedish	CSPS	Swedish
CIBSMPSM	CIBS	CSPS	Swedish	CSPS	Swedish

1) Born in Sweden   2) distributed over the ages: 1_6, 7_12, 13_16

Table 3. Explanation of the composite string variable.

Position	1	2	3	4	5
Explanation	Child	Biological father	Biological mother	Father the child is living with at the age of 16	Mother the child is living with at the age of 16

Table 4. Percentage adopted in migration variable groups. Cross tabulation of 10 and 5 groups.

	CSPS	CSPI	CIPS	CIPI	CIBS	Total
CSPS	0,4					
CSPI		0,4				
CIPS			99,6			
CIPI				1,2		
CIBS					2,1	
CSBSIPI		0,5				
CIPM				28,9		
CIBSMPM				1,6		
CIBSI					0,5	
CIBSMPSM					0,6	
Total	0,4	0,4	99,6	2,1	1,0	1,9