Codebook of variables

The intergenerational impact of naturalisation reforms: the citizenship status of children of immigrants in the Netherlands, 1995-2016

Marie Labussière and Maarten Vink

Naturalisation

ind_nat_t5	Naturalisation **or** change from unknown to Dutch nationality is observed during the year	Time-varying dummy
sum_nat_t5	Naturalisation **or** change from unknown to Dutch nationality is observed during the observation period	Time-constant dummy .b – Dutch from birth
date_nat_t5	Date of naturalisation (based on sum_nat_t5)	Time-constant variable 4-digit .b – Dutch from birth .c – Does not naturalise over the observation period
age_nat_t5	Age at naturalisation (based on sum_nat_t5)	Time-constant variable 2-digit .b – Dutch from birth .c – Does not naturalise over the observation period
natsup18	The respondent naturalises after the age of majority (based on age_nat_t5)	Time-constant dummy .b – Dutch from birth .c – Does not naturalise over the observation period
nat_with_whom	Family pattern of naturalisation (four clusters), identified by comparing the respondent's and their parents' dates of naturalisations.	Time-constant variable 1 – Naturalisation with the mother only 2- Naturalisation with the father only 3 – Naturalisation with both parents 4 – Independent naturalisation .b – Dutch from birth .c – Does not naturalise over the observation period
ind_nat_cr	Family pattern of naturalisation used for the competing risk model, based on *nat_with_whom*.	Time-constant variable 1 – Naturalisation with both parents 2 – Naturalisation with only one parent 3 – Independent naturalisation after majority

Eligibility

eli_child	The respondent becomes eligible for acquiring Dutch citizenship over the year, based on their parents' eligibility	Time-varying dummy 1 – Eligible, i.e. at leat one parent becomes eligible, is already Dutch or naturalises 0 – Non eligible, i.e. both parents are not eligible and/or dead. .b – Dutch from birth .z – post-naturalisation year
one_eli_child	The respondent is eligible at least once during the observed period	Time-constant dummy
year_eli_child	Year of eligibility	Time-constant variable 4-digit .b – Dutch from birth .c – never eligible
bothp_eli	Both parents are eligible during the year	Time-varying dummy .b – Dutch from birth .c – never eligible
sum_bothp_eli	Both parents are eligible at least once during the observed period	Time-constant dummy .b – Dutch from birth .c – never eligible
period_eli_child	Period during which the respondent becomes eligible (4 clusters)	Time-constant variable 1 – 1995-1996 2 – 1997-2002 3 – 2003-2006 4 – 2007 and after .b – Dutch from birth .c – never eligible
period_eli_child3	Period during which the respondent becomes eligible (two-year clusters)	Time-constant variable 1 – 1995-1996 2 – 1997-1998 3 – 1999-2000 4 – 2001-2002 5 – 2003-2004 6 – 2005-2006 7 – 2007-2008 8 – 2009-2017 .b – Dutch from birth .c – never eligible
period_eli_child6	Period during which the respondent becomes eligible (3 clusters)	Time-constant variable 1 – 1995-1997 2 – 1998-2002 3 – 2003 and after .b – Dutch from birth .c – never eligible

Parents' country of origin

Gbaherkomstgroepering_ma	Mother's country of origin (non-clustered), original CBS variable.	Time-constant variable 4-digit code (see SSB Catalogue)
Gbaherkomstgroepering_pa	Father's country of origin (non-clustered), original CBS variable.	Time-constant variable 4-digit code (see SSB Catalogue)
country_of_origin	Mother's country of origin at the year of birth (15 clusters). It is based on CBS variable *gbaherkomstgroepering*, which gives the country of birth for first-generation respondents. ⓘ If mother's country of origin is missing at the year of birth (e.g. if mother's identifier is missing), it is replaced by later values We use a time-constant variable because 1383 respondents (0.38%) have a change in their country of origin due to EU enlargement.	Time-constant variable 1 – Morocco 2 – Turkey 3 – Irak 4 – Yugoslavia 5 – Somalia 6 – Afghanistan 7 – former Soviet-Union 8 – Poland 9 – Great-Britain 10 – China 11- Germany 12 – Iran 13 – Angola 14 – Other EU countries 15 – Other non-EU countries
country_of_origin_bis	Father's country of origin at the year of birth (15 clusters). It is based on CBS variable *gbaherkomstgroepering*, which gives the country of birth for first-generation respondents. ⓘ If father's country of origin is missing at the year of birth (e.g. if father's identifier is missing), it is replaced by later values We use a time-constant variable because 1383 respondents (0.38%) have a change in their country of origin due to EU enlargement.	Time-constant variable 1 – Morocco 2 – Turkey 3 – Irak 4 – Yugoslavia 5 – Somalia 6 – Afghanistan 7 – former Soviet-Union 8 – Poland 9 – Great-Britain 10 – China 11- Germany 12 – Iran 13 – Angola 14 – Other EU countries 15 – Other non-EU countries
EU_child	At least one parent is from a EU country during the year Based on *country_of_origin* (mother's) and *country_of_origin_bis* (father's).	Time-varying dummy
EU_sum_child	At least one parent is from a EU country over the observation period	Time-constant dummy

Other sociodemographic characteristics

gbageslacht	Gender (original CBS variable)	Time-constant variable 1 – Man 2 – Woman
lft	Age (original CBS variable)	Time-varying variable
date_birth	Year of birth, based on *lft*.	Time-constant variable 4-digit
secm_pa/ma_agg1	Socioeconomic status of father/mother (aggregated based on secm_pa/ma) Categories have been clustered based on CBS variable SECM.	Time-varying variable 1 – Employee 2 – Self-employed 3 – Reception of benefits 4 – No declared income 5 – Missing SECM
secm_pa/ma_agg2	Socioeconomic status of father/mother (aggregated based on secm_pa/ma) Employees and self-employed are merged.	Time-varying variable 1 – Employee or self-employed 2 – Reception of benefits 3 – No declared income 4 – Missing SECM