Supplementary Materials

Timing of citizenship acquisition and immigrants' children educational outcomes: a family fixed-effects approach*

Marie Labussière

Final version 30 May 2023

Contents

*The code used in this study is available in the Harvard Dataverse archive `https://doi.org/10.7910/DVN/P39QGO`

1 Covariates description

Table S2 on p.4 gives the detailed coding of the full set of covariates, while I detail below the construction of three important variables: the score obtained at the Cito test (the outcome), age at naturalisation (the variable of interest) and parental level of education (an important confounder).

Cito test scores

A substantial share (23%) of students in the sample did not take the Cito standardised test at the end of primary school. This can be for different reasons: some primary schools do not offer any test (this is common in special primary education, SBO), while others have a different test provider than Cito. Unfortunately, there is very little external information about the profiles of such schools in the considered year. I have checked in the sample whether schools without a Cito test are systematically different from those with a Cito test in terms of their students' sociodemographic characteristics (results available upon request). It seems that students that are enrolled in a school without the test come from a slightly more disadvantaged social background than those who take the test. However, the differences are limited (2 percentage points at most).

Parental education level

Unfortunately, information on the educational level of first generation parents is very limited in the Dutch register data, especially in the 1990s and 2000s. To maximise the coverage, Statistics Netherlands obtains or derives information on education level from various registers and the Labor Force Survey.[1] However, although information on education level exists for an increasing share of the population, reaching 58% in 2012 [2], information is still missing for a substantial group of individuals, specifically for older cohorts and immigrants. Moreover, the level of education tends to be underestimated in administrative data because education abroad or at private institutions is difficult to take into account. As a result, an imputed educational level is available in the most recent registers. Since first-generation immigrants are particularly at risk of having their educational level underestimated, I use the imputed level of education whenever available. This measure of education is more reliable but it is also more recent, which introduces a risk of anticipatory analysis to the regression. Nevertheless, parents' education level is arguably relatively

[1] This combination of register and survey data requires the use of weights to get a representative measure of education for the entire Dutch population. However, the weights are not meant to ensure representativity for the sub-population of first generation immigrants alone; they have therefore not been used.

[2] Gans, S., Linder, F., Mooren, F.v., 2015. Gebruikershandleiding Onderwijsinformatie uit het SSB.

stable from the moment their children enter secondary education,[3] which justifies treating it as a time-constant characteristic in the absence of any more reliable information.

Age at naturalisation

Ideally, naturalisation should be measured before the outcome. However, one caveat is that students do not take the Cito test at the same age (see Table S1). In the Netherlands, entry into compulsory education is based on the age at the start of the school year, which means that each school level consists of at least two birth cohorts (including possible grade-repetition and grade-skipping). I therefore censor age at naturalisation at age 14, when the vast majority of students have already taken the test. This implies that those who naturalise at ages 11-13 may acquire Dutch citizenship shortly after taking the Cito test. Although this introduces a risk of anticipatory analysis in the regression, I do not expect a major bias to the extent the naturalisation process starts at least a year before the respondent officially becomes – and is registered as – a Dutch citizen. Yet, as a robustness check, I used an alternative specification where age at naturalisation is censored at age 10. This ensures temporal consistency but gives a conservative estimate for the age at naturalisation, as those who naturalise after 10 before the Cito test are included in the censored group. As expected, coefficients are smaller in magnitude, but the results do not substantively alter my interpretation (see p.26). I also conducted the analysis with different age groupings (including age dummies), which yield analogous conclusions about the effect of timing (see section 6 for all robustness checks).

Age	N	%
≤ 10	333	0.4
11	43,660	46.0
12	45,847	48.3
13	5,129	5.4
≥ 14	46	0.1
Total	95,015	100.00

Table S1: Distribution of age at the time of the test in the full sample of children of immigrants born in the Netherlands taking the Cito test between 2008 and 2015

[3] When looking at non-missing cases, 17.0% of mothers and 18.1% of fathers have more than one level of education recorded during the period under observation. Note that this gives an upper bound for variation in parental education level over time, because the measure relies on a very detailed classification of educational credentials.

Variable	Operationalisation	Measurement
Cito test score	Standardised by year	See note on p.2
Age at naturalisation	(1) 0 (Dutch from birth) (2) 1-2 (3) 3-4 (4) 5-6 (5) 7-10 (6) 11-13 (7) $\geq$ 14	Censored at age 14. See the robustness checks for an alternative specification with the variable censored at age 10 (p.25).
Gender	(1) Male (2) Female	Time-constant
Birth order	(1) First-born (2) Non first-born	Time-constant
Date of birth	Year dummies from 1995 to 2003	Time-constant
Number of minor children in the household	Range: 1-7	Year of Cito
Type of household	(1) Two-parent household (2) Single-parent household	Year of Cito
Father's and mother's socioeconomic status	(1) Employee or self-employed (2) Reception of benefits (3) No declared income (4) Missing information	Year of Cito
Parents' homeownership status	Dummy: at least one parent is a homeowner	Year of Cito
Family income	Terciles	Year of Cito
Highest educational degree among parents	(1) Low education (primary education) (2) Middle education (completed some secondary education) (3) High education (bachelor or higher) (4) Missing education	Time-constant See detailed note about the measurement of immigrants' level of education on p.2.
Mother's country of birth	(1) Afghanistan (2) China (3) Irak (4) Iran (5) Morocco (6) Other EU countries (7) Other non-EU countries (8) Turkey (9) Yugoslavia	Time-constant Note: 90.6% of children in the sample have parents originating from the same country of birth.

Table S2: Information about the coding of all variables

2 Descriptive Statistics

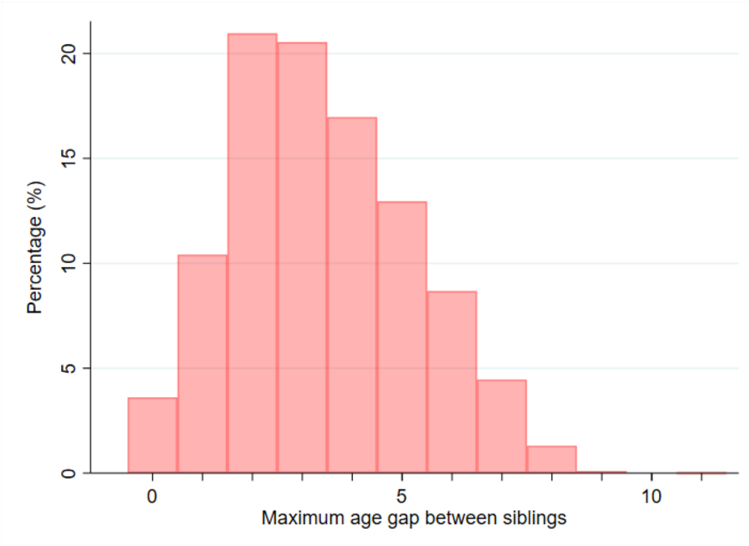

Figure S1: Age difference between siblings in the full sample (N=94,727)
Note: I take the *maximum* age difference in the families with three siblings or more.

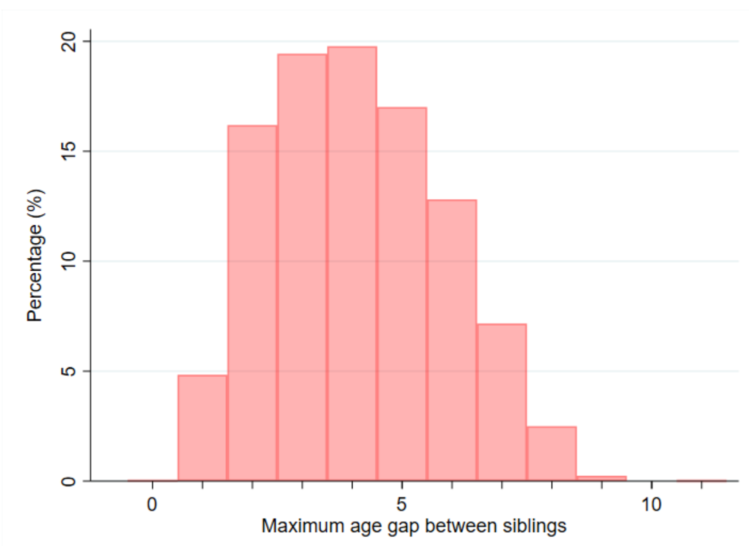

Figure S2: Age difference between siblings in the sample of exposure-discordant families (N=14,393)
Note: I take the *maximum* age difference in the families with three siblings or more.

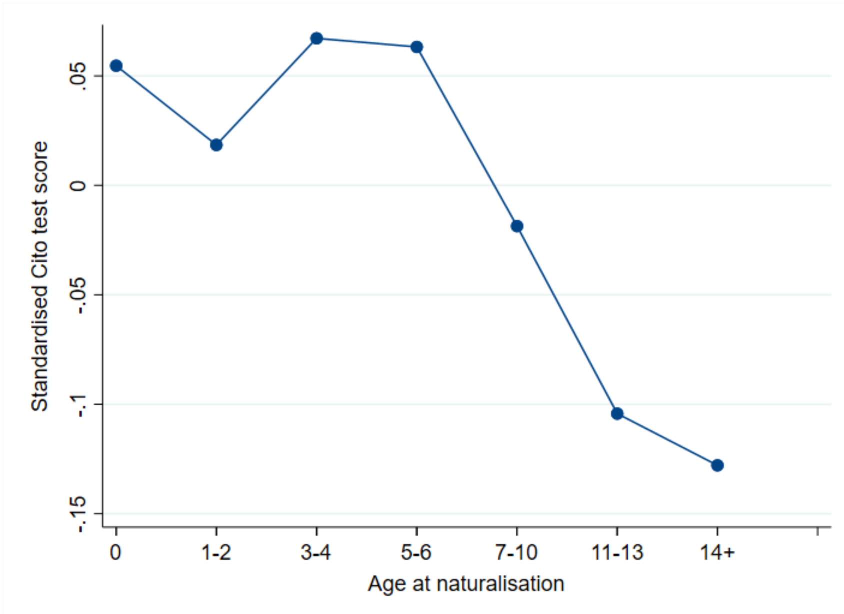

Figure S3: Average standardised Cito test scores for the children of immigrants born in the Netherlands, by age at naturalisation groups

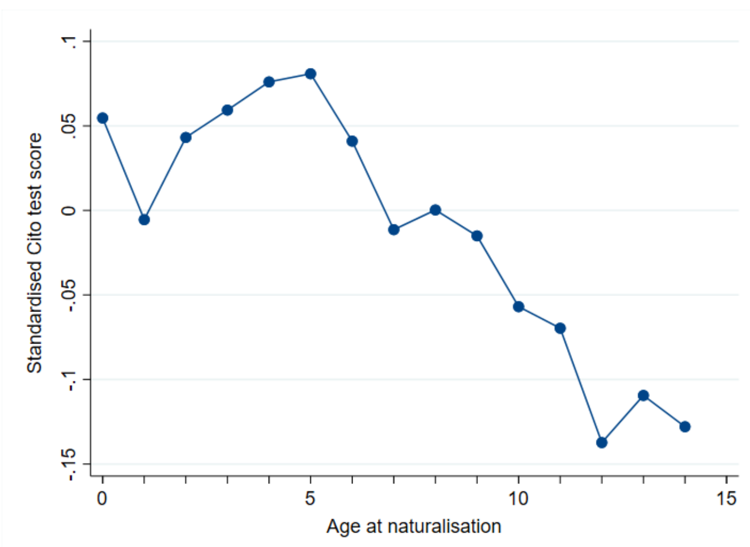

Figure S4: Average standardised Cito test scores for the children of immigrants born in the Netherlands, by age at naturalisation

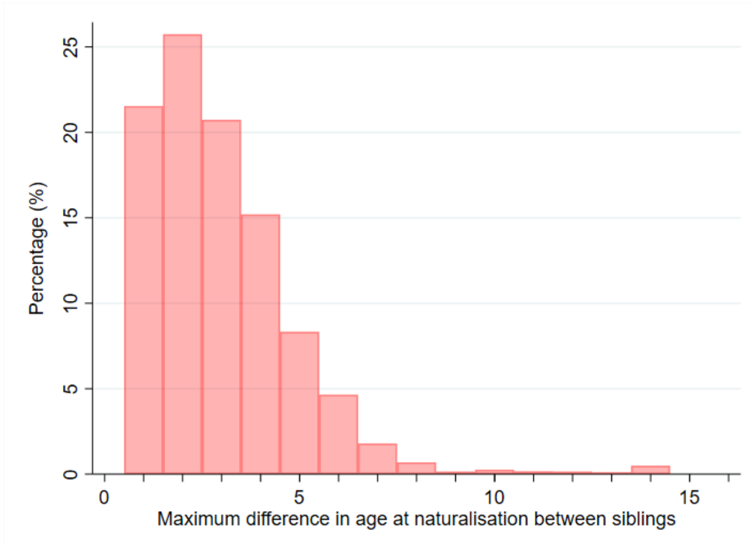

Figure S5: Difference in age at naturalisation between siblings in the sample of exposure-discordant families (N=14,393)

Note: I take the *maximum* difference in age at naturalisation in the families with three siblings or more.

Variable	N	%
Gender		
Male	47,390	50.0
Female	47,337	50.0
Birth order		
Non first-born	62,202	65.7
First-born	32,525	34.3
Date of birth		
$\leq$ 1995	5,982	6.3
1996	10,561	11.2
1997	10,819	11.4
1998	11,118	11.7
1999	11,547	12.2
2000	12,543	13.2
2001	12,489	13.2
2002	11,957	12.6
$\geq$ 2003	7,711	8.1
Number of children		
1	9,165	9.7
2	30,299	32.0
3	30,967	32.7
4	16,590	17.5
5	5,470	5.8
6	1,647	1.7
7 (and more)	589	0.6
Type of household		
Two-parent	78,115	82.5
Single-parent	16,612	17.5
SES status mother		
Employee or self-employed	36,444	38.5
Reception of benefits	30,341	32.0
No declared income	27,368	28.9
Missing information	574	0.6
SES status father		
Employee or self-employed	59,051	62.3
Reception of benefits	28,071	29.6
No declared income	4,144	4.4
Missing information	3,461	3.7

Continued

Table S3: Descriptive statistics for covariates

Variable	N	%
Parents' homeownership status		
Renters	64,312	67.9
Homeowners	30,415	32.1
Household income		
First tercile	30,786	32.5
Second tercile	31,119	32.9
Third tercile	31,871	33.7
Missing Income	951	1.0
Parents' highest educational level		
Low	42,258	44.6
Middle	26,651	28.1
High	12,359	13.1
Missing	13,459	14.2
Mother's country of birth		
Afghanistan	1,756	1.9
China	2,717	2.9
Irak	2,558	2.7
Iran	1,038	1.1
Morocco	33,729	35.6
Other EU countries	2,808	3.0
Other non-EU countries	16,830	17.8
Turkey	29,453	31.1
Yugoslavia	3,838	4.1
Total	**94,727**	**100.0**

Table S3 Continued: Descriptive statistics for covariates

Age	N	% in the full sample	% among foreign citizens at birth
0	62,338	65.8	–
1	3,555	3.8	11.0
2	3,445	3.6	10.6
3	2,815	3.0	8.7
4	2,503	2.6	7.7
5	2,153	2.3	6.7
6	1,697	1.8	5.2
7	1,293	1.4	4.0
8	1,095	1.2	3.4
9	987	1.0	3.1
10	879	0.9	2.7
11	760	0.8	2.4
12	702	0.7	2.2
13	614	0.7	1.9
14	464	0.5	1.4
15	361	0.4	1.1
16	250	0.3	0.8
17	147	0.2	0.5
18	444	0.5	1.4
19	365	0.4	1.1
20	50	0.1	0.2
21	7,810	8.2	24.1
Total	94,727	100.0	100.0

Table S4: Distribution of age at naturalisation in the full sample (N=94,727)

3 Appendix between-within linear mixed model

Model description

I analyse the within-family effect of age at naturalisation using the so-called "hybrid" or "correlated random-effects" model (Mundlak, 1978; Allison, 2009). This model separates *within-* and *between-effects* in a random-effects framework. Because of this defining feature, it has been increasingly known as the "between-within" model, following a suggestion from Sjölander et al. (2013). Contrary to standard fixed-effects modeling, the between-within model explicitly models heterogeneity bias and does not control out family-invariant covariates (Bell and Jones, 2015).

Let subscript i denotes the individual level and j denotes the family level. A standard random-effects model is:

$$\begin{cases} y_{ij} = \beta_0^{RE} + \beta_1^{RE} x_{ij} + \beta_2^{RE} c_j + (\mu_j + \epsilon_{ij}) \\ E(\mu_j | x_{ij}, c_j) = 0 \end{cases} \tag{1}$$

where y_{ij} is the educational outcome of child i in family j, x_{ij} a (serie of) covariate(s) measured at the individual level with β_1^{RE}, and c_j a (serie of) covariate(s) measured at the family level with β_2^{RE}. The idiosyncratic error ϵ_{ij} ($\sim N(0, \sigma_\epsilon^2)$) and the unmeasured cluster-level effect μ_j ($\sim N(0, \sigma_\mu^2)$) constitute the random part of the model. A main reason why fixed-effects models are usually preferred over random-effects models is that the latter assume that there is no unobserved heterogeneity at the cluster level, *i.e.*, $E(\mu_j | x_{ij}, c_j) = 0$.

By contrast, the between-within approach explicitly models the correlation between μ_j and x_{ij} with the assumption that the unmeasured cluster-level effect depends on the mean values of x_{ij}: $\mu_j = \gamma \overline{x_j} + \delta_j$, with $\delta_j \sim N(0, \sigma_\delta^2)$. This allows to pick up potential correlation between the unmeasured cluster effect μ_j and the individual-level variables. Following this assumption, equation of model (1) can be reformulated by decomposing the individual-level variables into a between- ($\overline{x_j} = n_i^{-1} \sum_{t=1}^{n_i} x_{ij}$) and a within-cluster ($x_{ij} - \overline{x_j}$) component (Allison, 2009; Schunck, 2013, 66), such that

$$\begin{cases} y_{ij} = \beta_0^{BW} + \beta_1^{BW}(x_{ij} - \overline{x_j}) + \beta_2^{BW} c_j + \beta_3^{BW} \overline{x_j} + (\delta_j + \epsilon_{ij}) \\ E(\delta_j | x_{ij}, c_j) = 0 \end{cases} \tag{2}$$

Under this formulation, the within- (β_1^{BW}) and between-cluster ($\beta_3^{BW} = \beta_1^{BW} + \gamma$) effects are now clearly separated. $\hat{\beta}_1^{BW}$ is identical to the fixed-effects estimator (Mundlak, 1978) and is consistent for β_1^{BW} under rather general conditions in the linear case such as ours

(Brumback et al., 2017, 3842)[4]. By contrast, $\hat{\beta}_3^{BW}$ is not robust to unobserved heterogeneity at the family level: it is biased if $E(\mu_j|x_{ij}, c_j) \neq 0$ (Schunck and Perales, 2017, 95).

Calculating the between-cluster effect is informative although the estimator is biased. It informs about the extent to which the estimator $\hat{\beta}_1^{RE}$ is biased when unobserved heterogeneity is not modelled (Schunck and Perales 2017, 96). More specifically, testing if $\beta_1^{BW} = \beta_3^{BW}$ provides a regression-based alternative to the Hausman test: if one cannot reject the hypothesis that the within-cluster and between-cluster effects are equal, (2) collapses to the random-effects model (1) and there is no evidence against $E(\mu_j|x_{ij}, c_j) = 0$ (Snijders and Berkhof, 2008, 145).

[4]Brumback et al. (2017, 3842) show in a generalised linear mixed model framework that the estimator $\hat{\beta}_1^{BW}$ is consistent for β_1 regardless of whether $\mu_j = \gamma\overline{x_j} + \delta_j$, under the conditions that the link function is the identity link *and* that the variance $var(y_{ij}|x_{ij}, c_j, \mu_j)$ is constant.

Full between-within model

	Random-effects		Within-effects		Between-effects	
	β_1^{RE}	p	β_1^{WB}	p	β_3^{WB}	p
Age at naturalisation						
Age 0			0.105	0.123	0.094	0.000
Age 1-2			0.120	0.070	0.079	0.000
Age 3-4			0.144	0.025	0.086	0.000
Age 5-6			0.127	0.042	0.086	0.000
Age 7-10			0.101	0.068	0.053	0.008
Age 11-13			0.051	0.308	0.004	0.868
Age $\geq$ 14			ref.	ref.	ref.	ref.
Gender						
Male			ref.	ref.	ref.	ref.
Female			-0.086	0.000	-0.052	0.000
Birth order						
Non first-born			ref.	ref.	ref.	ref.
First-born			0.013	0.299	0.080	0.000
Date of birth						
$\leq$ 1995			ref.	ref.	ref.	ref.
1996			0.078	0.001	0.222	0.000
1997			0.048	0.042	0.213	0.000
1998			0.051	0.032	0.186	0.000
1999			0.011	0.663	0.181	0.000
2000			0.021	0.416	0.178	0.000
2001			0.014	0.609	0.163	0.000
2002			0.010	0.726	0.162	0.000
$\geq$ 2003			0.063	0.046	0.330	0.000
Number of children			0.016	0.157	-0.037	0.000
Household type						
Two-parent households			ref.	ref.	ref.	ref.
Single-parent household			-0.030	0.346	-0.097	0.000
Mother SES status						
Employee or self-employed			ref.	ref.	ref.	ref.
Reception of benefits			0.039	0.091	-0.033	0.001
No declared income			0.026	0.240	0.034	0.000
Missing information			-0.085	0.641	-0.041	0.321
Father SES status						
Employee or self-employed			ref.	ref.	ref.	ref.
Reception of benefits			-0.016	0.445	-0.040	0.000
No declared income			0.009	0.779	0.017	0.323
Missing information			-0.052	0.327	-0.007	0.715
Parents' homeownership status						
Renters			ref.	ref.	ref.	ref.

Continued

Table S5: Between-within model for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands

	Random-effects		Within-effects		Between-effects	
	β_1^{RE}	p	β_1^{WB}	p	β_3^{WB}	p
Homeowners			-0.053	0.161	0.110	0.000
Household income						
First income decile			ref.	ref.	ref.	ref.
Second income decile			-0.014	0.378	-0.002	0.841
Third income decile			-0.001	0.953	0.056	0.000
Missing income			-0.138	0.022	-0.086	0.015
Highest parental education level						
Low education	ref.	ref.				
Middle education	0.146	0.000				
High education	0.455	0.000				
Missing education	0.131	0.000				
Mother's country of birth						
Other EU countries	re.f	ref.				
Afghanistan	0.136	0.000				
China	0.497	0.000				
Irak	-0.041	0.140				
Iran	0.066	0.062				
Morocco	-0.200	0.000				
Other non-EU countries	-0.009	0.646				
Turkey	-0.349	0.000				
Yugoslavia	-0.073	0.003				
Constant	-0.114	0.000				
Observations	94,727					
Family clusters	69,894					

Table S5 Continued: Between-within model for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands

Note: The between-within model does not estimate separate effects for the variables of parental education and mother's country of birth, since these variables only vary between clusters. The estimated effects for these two variables are the same as those in a standard random-effects model.

Comparison with a standard OLS model

	OLS		**Random-effects**		**Between-within**			
Age	β_1^{OLS}	p	β_1^{RE}	p	β_1^{WB}	p	β_3^{WB}	p
0	0.102	(0.000)	0.096	(0.000)	0.105	(0.125)	0.095	(0.000)
1-2	0.089	(0.000)	0.087	(0.000)	0.119	(0.070)	0.079	(0.000)
3-4	0.100	(0.000)	0.101	(0.000)	0.144	(0.025)	0.086	(0.000)
5-6	0.096	(0.000)	0.092	(0.000)	0.127	(0.043)	0.087	(0.000)
7-10	0.060	(0.000)	0.059	(0.001)	0.101	(0.069)	0.054	(0.007)
11-13	0.009	(0.700)	0.012	(0.600)	0.051	(0.308)	0.005	(0.859)
$\geq$ 14	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Observations	94,727		94,727		94,727			
Family clusters	–		69,894		69,894			
Covariates	All*		All*		All*			
R^2	0.099		0.098		–			

Table S6: OLS, random-effects and between-within models for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, by age at naturalisation groups

*I include all the covariates described in Table S2: gender, birth order, date of birth, household type, mother's and father's SES status, homeownership status, household income, highest educational level among parents and mother's country of origin.

4 Appendix random-effects models

Interactions with parental education

➢ With naturalisation dummy

	β	p
Citizenship status		
(ref. not naturalised at age 14)		
Naturalised before 14	0.124	0.000
Parental education		
Low education	ref.	(.)
Middle education	0.150	0.000
High education	0.601	0.000
Missing education	0.269	0.000
Interactions citizenship x parental education		
Low education × Naturalised before 14	ref.	(.)
Middle education × Naturalised before 14	0.000	0.993
High education × Naturalised before 14	-0.147	0.000
Missing education × Naturalised before 14	-0.149	0.000
Observations	94,727	
Family clusters	69,894	
Covariates	All*	
R^2	0.099	

Table S7: Random-effects model for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, with interactions between naturalisation dummy and parental education

*All the other covariates described in Table S2 are included: gender, birth order, date of birth, household type, mother's and father's SES status, homeownership status, household income, mother's country of birth.

➢ With age at naturalisation

	β	p
Age at naturalisation		
Age 0	0.131	0.000
Age 1-2	0.109	0.000
Age 3-4	0.127	0.000
Age 5-6	0.101	0.000
Age 7-10	0.115	0.000
Age 11-13	0.078	0.010
Age $\geq$ 14	ref.	ref.
Parents' homeownership status		
Renters	ref.	ref.
Homeowners	0.108	0.000
Highest parental education level		
Low education	ref.	ref.
Middle education	0.149	0.000
High education	0.601	0.000
Missing education	0.268	0.000
Highest parental education level $\times$ Age at naturalisation		
Low education $\times$ Age 0	ref.	ref.
Low education $\times$ Age 1-2	ref.	ref.
Low education $\times$ Age 3-4	ref.	ref.
Low education $\times$ Age 5-6	ref.	ref.
Low education $\times$ Age 7-10	ref.	ref.
Low education $\times$ Age 11-13	ref.	ref.
Low education $\times$ Age $\geq$ 14	ref.	ref.
Middle education $\times$ Age 0	-0.002	0.927
Middle education $\times$ Age 1-2	0.014	0.707
Middle education $\times$ Age 3-4	0.014	0.712
Middle education $\times$ Age 5-6	0.054	0.210
Middle education $\times$ Age 7-10	-0.025	0.544
Middle education $\times$ Age 11-13	-0.048	0.372
Middle education $\times$ Age $\geq$ 14	ref.	ref.
High education $\times$ Age 0	-0.148	0.000
High education $\times$ Age 1-2	-0.125	0.013
High education $\times$ Age 3-4	-0.125	0.018
High education $\times$ Age 5-6	-0.107	0.062
High education $\times$ Age 7-10	-0.222	0.000
High education $\times$ Age 11-13	-0.223	0.005
High education $\times$ Age $\geq$ 14	ref.	ref.
Missing education $\times$ Age 0	-0.147	0.000
Missing education $\times$ Age 1-2	-0.116	0.009
Missing education $\times$ Age 3-4	-0.149	0.002
Missing education $\times$ Age 5-6	-0.120	0.033

Continued

Table S8: Random-effects model for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, with interactions between age at naturalisation and parental education

	β	p
Missing education × Age 7-10	-0.194	0.000
Missing education × Age 11-13	-0.284	0.000
Missing education × Age $\geq$ 14	ref.	ref.
Observations	94,727	
Family clusters	69,894	
Covariates	All*	
R^2	0.099	

Table S8 Continued: Random-effects model for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, with interactions between age at naturalisation and parental education

*I include all the covariates described in Table S2.

Interactions with parental education and homeownership status

➢ With naturalisation dummy

	β	p
Citizenship status		
(ref. not naturalised at age 14)		
Naturalised before 14	0.140	0.000
Parental education		
Low education	ref.	(.)
Middle education	0.140	0.000
High education	0.575	0.000
Missing education	0.248	0.000
Parental homeownership status		
Renters	ref.	(.)
Homeowners	0.197	0.000
Interactions citizenship x parental education		
Low education × Naturalised before 14	ref.	(.)
Middle education × Naturalised before 14	0.011	0.689
High education × Naturalised before 14	-0.119	0.002
Missing education × Naturalised before 14	-0.126	0.000
Interactions citizenship x parental homeownership status		
Renters × Naturalised before 14	ref.	(.)
Homeowners × Naturalised before 14	-0.095	0.000
Observations	94,727	
Family clusters	69,894	
Covariates	All*	
R^2	0.099	

Table S9: Random-effects model for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, with interactions between naturalisation dummy and parental education and homeownership status

*All the other covariates described in Table S2 are included: gender, birth order, date of birth, household type, mother's and father's SES status, household income, mother's country of birth.

➢ With age at naturalisation

	β	p
Age at naturalisation		
Age 0	0.150	0.000
Age 1-2	0.123	0.000
Age 3-4	0.132	0.000
Age 5-6	0.115	0.000
Age 7-10	0.124	0.000
Age 11-13	0.095	0.002
Age $\geq$ 14	ref.	ref.
Parents' homeownership status		
Renters	ref.	ref.
Homeowners	0.199	0.000
Highest parental education level		
Low education	ref.	ref.
Middle education	0.139	0.000
High education	0.574	0.000
Missing education	0.246	0.000
Highest parental education level $\times$ Age at naturalisation		
Low education $\times$ Age 0	ref.	ref.
Low education $\times$ Age 1-2	ref.	ref.
Low education $\times$ Age 3-4	ref.	ref.
Low education $\times$ Age 5-6	ref.	ref.
Low education $\times$ Age 7-10	ref.	ref.
Low education $\times$ Age 11-13	ref.	ref.
Low education $\times$ Age $\geq$ 14	ref.	ref.
Middle education $\times$ Age 0	0.010	0.704
Middle education $\times$ Age 1-2	0.023	0.536
Middle education $\times$ Age 3-4	0.017	0.670
Middle education $\times$ Age 5-6	0.063	0.143
Middle education $\times$ Age 7-10	-0.019	0.655
Middle education $\times$ Age 11-13	-0.036	0.497
Middle education $\times$ Age $\geq$ 14	ref.	ref.
High education $\times$ Age 0	-0.115	0.003
High education $\times$ Age 1-2	-0.099	0.051
High education $\times$ Age 3-4	-0.110	0.039
High education $\times$ Age 5-6	-0.081	0.163
High education $\times$ Age 7-10	-0.201	0.000
High education $\times$ Age 11-13	-0.193	0.015
High education $\times$ Age $\geq$ 14	ref.	ref.
Missing education $\times$ Age 0	-0.121	0.000
Missing education $\times$ Age 1-2	-0.096	0.034
Missing education $\times$ Age 3-4	-0.138	0.005
Missing education $\times$ Age 5-6	-0.098	0.083
		Continued

Table S10: Random-effects model for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, with interactions between age at naturalisation and parental education and homeownership status

	β	p
Missing education × Age 7-10	-0.177	0.001
Missing education × Age 11-13	-0.259	0.000
Missing education × Age $\geq$ 14	ref.	ref.
Parents' homeownership status × Age at naturalisation		
Renters × Age 0	ref.	ref.
Renters × Age 1-2	ref.	ref.
Renters × Age 3-4	ref.	ref.
Renters × Age 5-6	ref.	ref.
Renters × Age 7-10	ref.	ref.
Renters × Age 11-13	ref.	ref.
Renters × Age $\geq$ 14	ref.	ref.
Homeowners × Age 0	-0.109	0.000
Homeowners × Age 1-2	-0.083	0.015
Homeowners × Age 3-4	-0.026	0.498
Homeowners × Age 5-6	-0.088	0.039
Homeowners × Age 7-10	-0.052	0.221
Homeowners × Age 11-13	-0.112	0.052
Homeowners × Age $\geq$ 14	ref.	ref.
Observations	94,727	
Family clusters	69,894	
Covariates	All*	
R^2	0.099	

Table S10 Continued: Random-effects model for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, with interactions between age at naturalisation and parental education and homeownership status

*I include all the covariates described in Table S2.

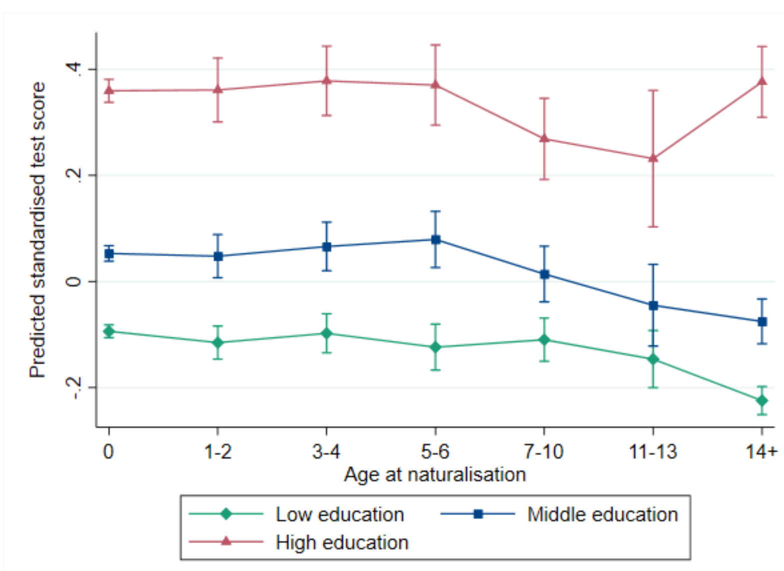

Figure S6: Predicted average standardised test score obtained at the end of primary school by children of immigrants born in the Netherlands, conditionally on age at naturalisation and parents' highest education level.

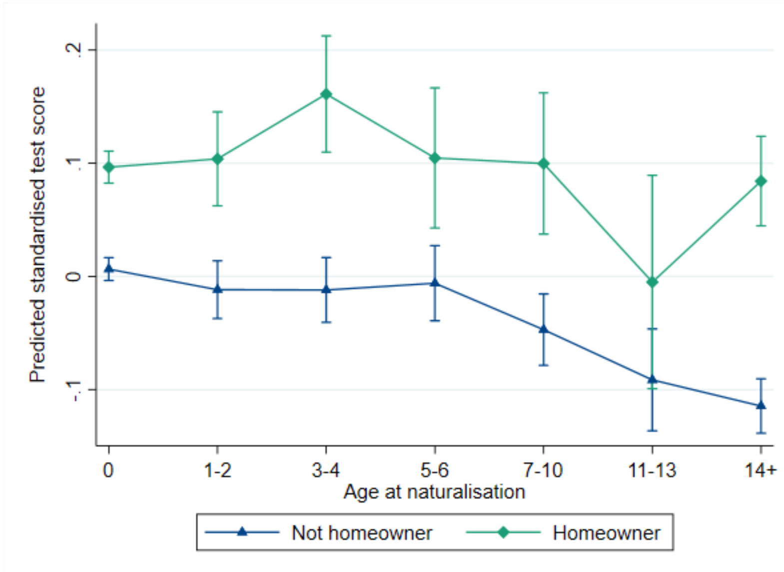

Figure S7: Predicted average standardised test score obtained at the end of primary school by children of immigrants born in the Netherlands, conditionally on age at naturalisation and parents' homeownership status.

5 Appendix bounding estimator

As a robustness check, I assess the sensitivity of the results to unobserved heterogeneity using a bounding technique. This approach has first been proposed by Altonji et al. (2005), and was further expanded by Oster (2019), who proposed a formal bounding argument linking selection bias to coefficient and R-squared stability. This method is adequate when researchers are able to include a relatively broad set of relevant regression controls, which are informative about the relationship between the regressor and unobservable confounding factors (Altonji et al., 2005, 171).

Oster developed a consistent estimator for the bias-adjusted treatment effect $\beta^{\star}$, which can be approximated with:

$$\beta^{\star} \approx \tilde{\beta} - \delta \left[\dot{\beta} - \tilde{\beta}\right] \frac{R_{max} - \tilde{R}}{\tilde{R} - \dot{R}} \tag{3}$$

where $(\tilde{\beta}, \tilde{R})$ and $(\dot{\beta}, \dot{R})$ are the coefficients for naturalisation and the (non-adjusted) R-squared resulting from the controlled and uncontrolled regression, respectively. δ is the selection ratio, which can be interpreted as "the degree of selection on unobservables relative to observables that would be necessary to explain away the result" (Oster, 2019, 195). For example, a value of $\delta = 1$ reflects the situation where the observables (*e.g.*, parental education, income) are as important as the unobservables (*e.g.*, parents' language proficiency and orientation towards the host society). Following Altonji et al. (2005) and as shown by Oster (2019, 197), it is a reasonable cutoff when the most important controls are included in the model.

Besides δ, another key input into the estimator of $\beta^{\star}$ is R_{max}. This R-squared results from the hypothetical regression where one would control for the full set of unobservables. While Altonji et al. (2005) argue that including all the unobservables would produce an R-squared of 1, Oster (2019) shows that this assumption is too conservative in most empirical settings. Likewise, in the present case, there are reasons to think that educational attainment cannot be fully explained, even when controlling for the full set of characteristics. This is so because one can expect residual idiosyncratic variation, as well as measurement error in school outcomes. As a result, I follow Oster's recommendation to set $R_{max} = 1.3\tilde{R}$ in the main specification.

To my knowledge, Oster's bounding technique has not been developed for non-binary categorical treatment variables. I therefore use a dummy variable for the effect of naturalisation, capturing whether the respondent has acquired Dutch citizenship before the age of 14. Table S11 displays the results for the OLS, random-effects and between-within specifications when this dummy is used instead of age at naturalisation groups.

	OLS	Random-effects	Between-within		Test
	β_1^{OLS}	β_1^{RE}	β_1^{WB}	β_3^{WB}	$\beta_1^{WB} = \beta_3^{WB}$
Naturalisation before 14	0.095	0.089	0.073	0.089	–
	(0.000)	(0.000)	(0.126)	(0.000)	(0.739)
Observations	94,727	94,727	94,727		
Family clusters	–	69,894	69,894		
Covariates	All*	All*	All*		
R^2	0.098	0.098	–		

Table S11: OLS, random-effects and between-within models for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, by naturalisation dummy

*I include all the covariates described in Table S2: gender, birth order, date of birth, household type, mother's and father's SES status, homeownership status, household income, highest educational level among parents and mother's country of birth.

I estimate $\beta^{\star}$ and assess the extent to which the conclusions drawn from the controlled regression are still valid when considering selection on unobservables. The standard-errors for $\beta^{\star}$ are bootstrapped using 1,000 replications.

To further explore the sensitivity of the results to omitted variable bias, I estimate $\beta^{\star}$ assuming a range of different values for R_{max} in Table S12. Note that as an alternative definition for R_{max}, Oster suggests to use sibling correlation as a benchmark (Oster, 2019, 198). In the present case, sibling correlation in test scores is equal to $\rho = 0.52$. However, setting the R_{max} to this value would far exceed the usual explanatory power of social science models for individual behaviour, which is known to be relatively low (Altonji et al., 2005, 171). As Table S12 shows below, the results for the bounding analysis are not robust to such a high R_{max}.

	Uncontrolled OLS			Controlled OLS	
	$\dot{\beta}$	$\dot{R}$		$\tilde{\beta}$	$\tilde{R}$
	0.164	0.016		0.097	0.098
	(0.000)			(0.000)	
R_{max}	0.13	0.15	0.17	0.19	0.2
Bias-adjusted treatment effect $\beta^{\star}$	0.071	0.052	0.033	0.014	0.003
	(0.000)	(0.000)	(0.003)	(0.230)	(0.805)

Table S12: Bounding the effect of naturalisation on standardised test scores obtained at the end of primary education by children of immigrants born in the Netherlands

Note: p-values are in parentheses (standard-errors are bootstrapped using 1,000 replications). I set $\delta = 1$, which reflects the situation where the observables are as important as the unobservables. The controlled regression includes all the covariates described in Table S2. The uncontrolled regression includes gender, birth order and date of birth, which are considered to be unrelated to the selection process.

6 Robustness checks

Family definition based on both parents' identifiers

In the main analysis, I define families based on mothers' personal identifiers only. This is because there are about ten times as many children whose father's identifier is missing at least one year during the observation period as there are children whose mother's identifier is similarly missing. This suggests that children of immigrants are more likely to have a father than a mother who is not continuously registered as a Dutch resident. As continuous residence is a requirement for naturalisation, priority was given to mothers' identifiers.

However, without fathers' identifiers, it is difficult to correctly identify complex family patterns where, for example, siblings are born of the same mother but not of the same father. To ensure that the results are robust to alternative definitions of the family clusters, I use both mothers' and fathers' identifiers in a robustness check. This leads to 70,002 family clusters, versus 69,894 previously when only mothers' identifiers were considered. As Table S13 shows, this leads to very comparable results.

	Random-effects		Within-effects		Between-effects		$\beta_1^{WB}=\beta_3^{WB}$
	β_1^{RE}	p	β_1^{WB}	p	β_3^{WB}	p	p
Age at naturalisation							
Age 0	0.098	0.000	0.101	0.143	0.097	0.000	0.946
Age 1-2	0.087	0.000	0.113	0.092	0.080	0.000	0.633
Age 3-4	0.098	0.000	0.138	0.034	0.082	0.000	0.412
Age 5-6	0.090	0.000	0.117	0.065	0.087	0.000	0.653
Age 7-10	0.059	0.001	0.093	0.097	0.055	0.006	0.528
Age 11-13	0.011	0.618	0.046	0.364	0.004	0.889	0.461
Age $\geq$ 14	ref.	ref.	ref.	ref.	ref.	ref.	
Observations	94,727		94,727				
Family clusters	70,002		70,002				
Covariates	All*		All*				

Regression coefficients, p-values in parentheses

Table S13: Random-effects and between-within models for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, by age at naturalisation
*I include all the covariates described in Table S2.

Age at naturalisation censored at age 10

Ideally, naturalisation should be measured before the outcome. However, one caveat is that students do not take the Cito test at the same age (see Table S1). In the Netherlands, entry into compulsory education is based on the age at the start of the school year, which means that each school level consists of at least two birth cohorts (including possible grade-repetition and grade-skipping). I therefore censor age at naturalisation at age 14, when the vast majority of students have already taken the test. This implies that those who naturalise at ages 11-13 may acquire Dutch citizenship shortly after taking the Cito test. Although this introduces a risk of anticipatory analysis in the regression, I do not expect a major bias to the extent the naturalisation process starts at least a year before the respondent officially becomes – and is registered as – a Dutch citizen. Yet, as a robustness check, I used an alternative specification where age at naturalisation is censored at age 10. This ensures temporal consistency but gives a conservative estimate for the age at naturalisation, as those who naturalise after 10 before the Cito test are included in the censored group. As expected, coefficients are smaller in magnitude, but the results do not substantively alter my interpretation.

	Random-effects		Within-effects		Between-effects		$\beta_1^{WB}=\beta_3^{WB}$
	β_1^{RE}	p	β_1^{WB}	p	β_3^{WB}	p	p
Age at naturalisation							
Age 0	0.093	0.000	0.066	0.241	0.093	0.000	0.639
Age 1-2	0.084	0.000	0.081	0.133	0.078	0.000	0.953
Age 3-5	0.098	0.000	0.106	0.042	0.085	0.000	0.699
Age 6-10	0.090	0.000	0.089	0.076	0.086	0.000	0.950
Age 11-13	0.056	0.001	0.063	0.124	0.053	0.007	0.822
Age $\geq$ 14	ref.	ref.	ref.	ref.	ref.	ref.	
Observations	94,727		94,727				
Family clusters	69,894		69,894				
Covariates	All*		All*				

Regression coefficients, p-values in parentheses

Table S14: Random-effects and between-within models for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, by age at naturalisation censored at age 10

*I include all the covariates described in Table S2.

Different age groupings for age at naturalisation

➢ Age dummies

	Random-effects		Within-effects		Between-effects		β_1^{WB}=β_3^{WB}
	β_1^{RE}	p	β_1^{WB}	p	β_3^{WB}	p	p
Age at naturalisation							
Age 0	0.096	0.000	0.093	0.187	0.094	0.000	0.991
Age 1	0.078	0.000	0.096	0.172	0.072	0.001	0.744
Age 2	0.095	0.000	0.122	0.079	0.085	0.000	0.609
Age 3	0.099	0.000	0.123	0.077	0.085	0.000	0.610
Age 4	0.102	0.000	0.147	0.032	0.085	0.001	0.395
Age 5	0.094	0.000	0.116	0.093	0.093	0.000	0.757
Age 6	0.090	0.000	0.129	0.059	0.078	0.008	0.493
Age 7	0.061	0.025	0.070	0.305	0.077	0.022	0.929
Age 8	0.073	0.012	0.096	0.164	0.078	0.029	0.814
Age 9	0.066	0.029	0.155	0.021	0.030	0.425	0.105
Age 10	0.030	0.344	0.080	0.232	0.016	0.688	0.408
Age 11	0.030	0.380	0.071	0.282	0.024	0.578	0.550
Age 12	0.019	0.588	0.049	0.468	0.041	0.343	0.263
Age 13	0.024	0.521	0.034	0.622	0.035	0.453	0.996
Age $\geq$ 14	ref.	ref.	ref.	ref.	ref.	ref.	
Observations	94,727		94,727				
Family clusters	69,894		69,894				
Covariates	All*		All*				

Regression coefficients, p-values in parentheses

Table S15: Random-effects and between-within models for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, by age at naturalisation (age dummies)
*I include all the covariates described in Table S2.

➢ 6-cluster variable

	Random-effects		Within-effects		Between-effects		β_1^{WB}=β_3^{WB}
	β_1^{RE}	p	β_1^{WB}	p	β_3^{WB}	p	p
Age at naturalisation							
Age 0	0.095	0.000	0.091	0.171	0.094	0.000	0.968
Age 1-2	0.086	0.000	0.107	0.098	0.079	0.000	0.674
Age 3-5	0.098	0.000	0.128	0.039	0.088	0.000	0.533
Age 6-10	0.068	0.000	0.103	0.061	0.060	0.001	0.458
Age 11-13	0.012	0.592	0.051	0.311	0.004	0.881	0.409
Age $\geq$ 14	ref.	ref.	ref.	ref.	ref.	ref.	
Observations	94,727		94,727				
Family clusters	69,894		69,894				
Covariates	All*		All*				

Regression coefficients, p-values in parentheses

Table S16: Random-effects and between-within models for the standardised test scores obtained at the end of primary school by children of immigrants born in the Netherlands, by age at naturalisation (6 instead of 7 clusters)
*I include all the covariates described in Table S2.

7 References

References

Allison, P. D. (2009). *Fixed effects regression models*. Number 160 in Quantitative applications in the social sciences. SAGE, Los Angeles.

Altonji, J., Elder, T., and Taber, C. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1):151–184.

Bell, A. and Jones, K. (2015). Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. *Political Science Research and Methods*, 3(1):133–153.

Brumback, B. A., Li, L., and Cai, Z. (2017). On the use of between–within models to adjust for confounding due to unmeasured cluster-level covariates. *Communications in Statistics - Simulation and Computation*, 46(5):3841–3854.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46(1):69–85.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Schunck, R. (2013). Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. *The Stata Journal*, 13(1):65–76.

Schunck, R. and Perales, F. (2017). Within- and between-cluster effects in generalized linear mixed models: A discussion of approaches and the xthybrid command. *The Stata Journal*, 17(1):89–115.

Sjölander, A., Lichtenstein, P., Larsson, H., and Pawitan, Y. (2013). Between-within models for survival analysis. *Statistics in Medicine*, 32(18):3067–3076.

Snijders, T. A. B. and Berkhof, J. (2008). Diagnostic Checks for Multilevel Models. In Leeuw, J. d. and Meijer, E., editors, *Handbook of multilevel analysis*, pages 141–175. Springer, New York.