

Bilag S.1: Beskrivelse af beregningen af koefficienten på indvandrerbaggrund

Det er kun i model (1) i artiklen, at den gennemsnitlige betydning af at have indvandrerbaggrund (α_1) direkte kan estimeres. Så snart dummyvariable for oprindelseslande inkluderes, vil de fange betydning af at have baggrund fra Tyrkiet, Iran osv., og det vil derfor medføre lineær afhængighed, hvis man både inkluderer en dummyvariabel for indvandrerbaggrund og dummyvariable for hvert oprindelsesland. I praksis betyder det, at IB_i som variabel udelades fra estimeringerne, når oprindelsesland indgår, og koefficienten på IB_i beregnes i stedet som et gennemsnit af koefficienterne på oprindelsesland vægtet med indvandrerpopulationens andel.

Når betydningen af forældres uddannelse og socioøkonomiske status tillades at variere med indvandrerbaggrund, gør det beregningerne af koefficienten mere kompleks. Beregningerne skal nu yderligere inkludere koefficienterne fra de interagerede led, og spørgsmålet er, hvordan betydningen af indvandrerbaggrund skal evalueres. En mulighed er at evaluere den i den etniske danske forælderfordeling. Dette er relateret til Blinder-Oaxaca dekomponeringsmetoden, hvor en kløft mellem to grupper forklares ved at dekomponere den i to dele Oaxaca (1973). En del af kløften forklares af forskelle i størrelsen på faktorerne og den anden af forskelle i betydningen af faktorerne. Når der i artiklen inkluderes et interaktionsled mellem indvandrerbaggrund og familiebaggrund, så tillades betydningen af familiebaggrund at være forskellig for etnisk danske børn og børn med indvandrerbaggrund. Når den gennemsnitlige forskel, evalueres i den etnisk danske forælderfordeling, svarer det til at undersøge om uddannelseskløften består, når børn med indvandrerbaggrund gives de samme forældreressourcer som etnisk danske børn.

Tabel S.2: Stikprøvereduktion for hver stikprøveudvælgelseskriterie

Beskrivelse af udvælgelseskriterier	Reduktion	Stikprøvestørrelse efter kriterie
Antallet af etniske danske børn og børn med indvandrerbaggrund fra ikke-vestlige lande, som afslutter 9. klasse mellem 1990 og 2012		219.581
Observeres fortsat fem år efter 9. klasses afslutning	65.911	153.670
- Tilhører årgange der har færdiggjort 9. klasse mellem 2008 og 2012	61.363	
- Emigreret eller død	4.548	
Ingen børn med én etnisk dansk forælder og én forælder med indvandrerbaggrund ikke i udlandet i mere end et kalenderår i løbet af grundskolen eller de fem efterfølgende år	3.869	149.801
Højeste ankomstald er 15 år	3.086	146.715
Mellem 15 og 18 år når man afslutter 9. klasse	1.759	144.956
	586	144.370

NOTE: Registerdata fra Danmarks Statistik for stikprøve af ikke-vestlige indvandrere og etniske danskere der afslutter 9. klasse mellem 1990 og 2007. Børn med indvandrerbaggrund er udtrukket fra et datasæt med hele den danske befolkning, mens de etniske danske børn er udtrukket fra et datasæt med en 10 pct. stikprøve af den danske befolkning.

Tabel S.3: Stikprøvesammensætning for hver grundskoleårgang

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Etniske danskere	6.112	6.043	5.554	5.251	5.289	4.908	4.769	4.409	4.439	4,186
Indvandrerbaggrund født i Danmark	491	579	568	691	785	928	1.122	1.132	1.090	1,047
Indvandrerbaggrund født i udlandet	793	897	935	1.164	1.183	1.274	1.506	1.706	1.857	1,920
I alt	7.396	7.519	7.057	7.106	7.257	7.110	7.397	7.247	7.386	7,153

	2000	2001	2002	2003	2004	2005	2006	2007	Total
Etniske danskere	4.296	4.445	4.529	4.557	4.825	5.033	5.145	5.125	88.915
Indvandrerbaggrund født i Danmark	1.149	1.346	1.476	1.834	2.011	2.349	2.523	2.884	24.005
Indvandrerbaggrund født i udlandet	2.075	2.180	2.196	2.180	2.325	2.501	2.396	2.362	31.450
I alt	7.520	7.971	8.201	8.571	9.161	9.883	10.064	10.371	144.370

NOTE: Registerdata fra Danmarks Statistik for en stikprøve af børn med indvandrerbaggrund fra ikke-vestlige lande og etniske danske børn som gennemfører 9. klasse mellem 1990 og 2007 (Stikprøveudvælgelse beskrives i afsnittet 'Datagrundlag og stikprøveudvælgelse').

Tabel S.4: Oprindelsesland for børn med
indvandrerbaggrund

Variable	Middelværdi	Std. afv.
Tyrkiet	0,231	0,421
Pakistan	0,082	0,274
Libanon	0,074	0,262
Tidligere Jugoslavien	0,069	0,253
Bosnien og Hercegovina	0,068	0,252
Irak	0,056	0,230
Vietnam	0,048	0,213
Iran	0,045	0,206
Sri Lanka	0,037	0,189
Marokko	0,032	0,177
Somalia	0,031	0,173
Afghanistan	0,027	0,161
Sydkorea	0,026	0,159
Andre lande	0,175	0,380
Antal observationer		55.455

NOTE: Registerdata fra Danmarks Statistik for en stikprøve af børn med indvandrerbaggrund fra ikke-vestlige lande og etniske danske børn som gennemfører 9. klasse mellem 1990 og 2007 (Stikprøveudvælgelse beskrives i afsnittet 'Datagrundlag og stikprøveudvælgelse').

Tabel S.5: Beskrivende statistik for øvrige variable: Middelværdi (standardafvigelse i parentes)

Variabel	Indvandrerbaggrund		
	Etnisk dansk	Født i Danmark	Født i udlandet
Mors socioøkonomiske status			
<i>Beskæftiget</i>			
Topledere	0,008 (0,091)	0,002 (0,044)	0,001 (0,031)
Lønmodtager på højeste niveau	0,129 (0,335)	0,029 (0,168)	0,022 (0,146)
Lønmodtager på mellemniveau	0,275 (0,446)	0,050 (0,219)	0,033 (0,179)
Lønmodtager på grundniveau	0,215 (0,411)	0,115 (0,319)	0,072 (0,258)
Ufaglærte	0,065 (0,247)	0,033 (0,177)	0,018 (0,134)
Andre lønmodtagere	0,092 (0,289)	0,188 (0,391)	0,103 (0,304)
Selvstændige og medarbejdende ægtefælle	0,054 (0,225)	0,044 (0,204)	0,024 (0,152)
<i>Arbejdsløs</i>			
Arbejdsløs	0,046 (0,209)	0,138 (0,345)	0,102 (0,303)
<i>Uden for arbejdsstyrken</i>			
Under uddannelse	0,004 (0,064)	0,032 (0,177)	0,060 (0,237)
Andre uden for arbejdsstyrken	0,101 (0,301)	0,359 (0,480)	0,487 (0,500)
Missing	0,010 (0,101)	0,010 (0,101)	0,078 (0,268)
Mors uddannelse			
0-9 år	0,196 (0,397)	0,258 (0,438)	0,228 (0,420)
10-12 år	0,274 (0,446)	0,192 (0,394)	0,206 (0,404)
13+ år	0,510 (0,500)	0,232 (0,422)	0,342 (0,474)
Missing	0,020 (0,142)	0,318 (0,466)	0,224 (0,417)
Fars uddannelse			
0-9 år	0,198 (0,398)	0,287 (0,452)	0,160 (0,366)
10-12 år	0,182 (0,386)	0,159 (0,366)	0,127 (0,333)
13+ år	0,570 (0,495)	0,327 (0,469)	0,389 (0,488)
Missing	0,050 (0,219)	0,226 (0,419)	0,324 (0,468)
Mors alder ved barnets fødsel			
Under 18 år	0,004 (0,065)	0,020 (0,140)	0,033 (0,178)
18-22 år	0,149 (0,356)	0,274 (0,446)	0,259 (0,438)
23-27 år	0,387 (0,487)	0,313 (0,464)	0,311 (0,463)
28-32 år	0,313 (0,464)	0,224 (0,417)	0,192 (0,394)
33-37 år	0,115 (0,319)	0,116 (0,320)	0,094 (0,292)
38-42 år	0,024 (0,153)	0,037 (0,188)	0,028 (0,164)
Over 42 år	0,002 (0,044)	0,005 (0,069)	0,009 (0,094)
Missing	0,006 (0,079)	0,013 (0,112)	0,074 (0,261)
Fars alder ved barnets fødsel			
Under 18 år	0,001 (0,028)	0,002 (0,041)	0,003 (0,059)
18-22 år	0,052 (0,221)	0,111 (0,315)	0,066 (0,248)
23-27 år	0,274 (0,446)	0,251 (0,433)	0,218 (0,413)
28-32 år	0,361 (0,480)	0,243 (0,429)	0,241 (0,428)
33-37 år	0,191 (0,393)	0,193 (0,395)	0,150 (0,357)
38-42 år	0,066	0,098	0,072

	(0,248)	(0,297)	(0,258)
Over 42 år	0,025	0,052	0,044
	(0,156)	(0,222)	(0,206)
Missing	0,031	0,050	0,206
	(0,174)	(0,219)	(0,404)
Antal søskende			
Ingen søskende	0,423	0,241	0,221
	(0,494)	(0,427)	(0,415)
1 søskende	0,395	0,296	0,273
	(0,489)	(0,456)	(0,446)
2 søskende	0,144	0,241	0,203
	(0,351)	(0,428)	(0,402)
3 søskende	0,031	0,138	0,132
	(0,173)	(0,344)	(0,338)
4 eller flere søskende	0,007	0,084	0,141
	(0,083)	(0,277)	(0,348)
Missing	0,001	0,001	0,030
	(0,029)	(0,027)	(0,172)
Forældre er skilt eller ej			
Skilt	0,182	0,114	0,095
	(0,386)	(0,318)	(0,294)
Missing	0,001	0,001	0,030
	(0,029)	(0,027)	(0,172)
Barnets alder ved indvandring			
0 år			0,025
			(0,157)
1 år			0,060
			(0,238)
2 år			0,051
			(0,221)
3 år			0,048
			(0,215)
4 år			0,057
			(0,232)
5 år			0,060
			(0,237)
6 år			0,060
			(0,238)
7 år			0,064
			(0,245)
8 år			0,065
			(0,246)
9 år			0,066
			(0,248)
10 år			0,069
			(0,254)
11 år			0,072
			(0,259)
12 år			0,071
			(0,257)
13 år			0,068
			(0,251)
14 år			0,061
			(0,239)
15 år			0,052
			(0,222)
Missing			0,051
			(0,219)
Antal observationer	88.915	24.005	31.450

Tabel S.6: Resultater fra OLS estimationer efter trinvis kontrol af en række kontrolvariable (inden indvandrerbaggrund interageres med familiebaggrundsvariable)

Afhængig variabel: Gennemførelse af ungdomsuddannelse inden for fem år efter grundskole						
	(1)	(2)	(3)	(4)	(5)	(6)
Indvandrerstatus (Etniske danske børn som referencekategori)						
Indvandrerbaggrund	-0,169*** (0,009)	-0,138*** (0,004)	-0,133*** (0,004)	0,009** (0,004)	-0,078*** (0,004)	0,009** (0,004)
Indvandrerbaggrund · født i udlandet	-0,163*** (0,011)	-0,190*** (0,011)	-0,166*** (0,011)	-0,133*** (0,011)	-0,183*** (0,011)	-0,147*** (0,011)
Indvandrerbaggrund · trend	0,005*** (0,001)	0,004*** (0,001)	0,006*** (0,001)	0,003*** (0,001)	0,006*** (0,001)	0,004*** (0,001)
Indvandrerbaggrund · født i udlandet · trend	0,006*** (0,001)	0,004*** (0,001)	0,004*** (0,001)	0,003*** (0,001)	0,005*** (0,001)	0,004*** (0,001)
Kvinde	0,104*** (0,003)	0,103*** (0,003)	0,105*** (0,002)	0,106*** (0,002)	0,105*** (0,002)	0,106*** (0,002)
Kontrolvariable						
Grundskoleårgang	Ja	Ja	Ja	Ja	Ja	Ja
Oprindelsesland	Nej	Ja	Ja	Ja	Ja	Ja
Familiestruktur	Nej	Nej	Ja	Ja	Ja	Ja
Mor og fars socioøkonomiske status	Nej	Nej	Nej	Ja	Nej	Ja
Mor og fars uddannelse	Nej	Nej	Nej	Nej	Ja	Ja
Antal observationer	144.370	144.370	144.370	144.370	144.370	144.370
Tilpasset R ²	0,048	0,063	0,090	0,131	0,121	0,141

NOTE: Robuste standardafvigelser er angivet i parentes. Registerdata fra Danmarks Statistik for en stikprøve af børn med indvandrerbaggrund fra ikke-vestlige lande og etniske danske børn som gennemfører 9. klasse mellem 1990 og 2007 (Stikprøveudvælgelse beskrives i afsnittet 'Datagrundlag og stikprøveudvælgelse'). Regressioner er baseret på en *lineær sandsynlighedsmodel*. I specifikation (2) til (6), er koefficienterne på *indvandrerbaggrund* beregnet som gennemsnittet af koefficienterne på oprindelsesland vægtes med indvandrerpopulationens andel. Kontrolvariablene inkluderer 18 årgangsdummies, 14 oprindelsesland dummies, 8 aldersgruppeddummies for forældres alder ved barnets fødsel, 5 dummies for antal søskende, en skilsmisseuddummy, 10 dummies for forskellige niveauer for forældrenes socioøkonomiske status og 3 dummies for forældrenes uddannelsesniveau samt dummies for missing værdier. Urapporert konstant er inkluderet i alle regressioner. *** 1 pct. signifikansniveau. ** 5 pct. signifikansniveau. * 10 pct. signifikansniveau.

Tabel S.7: Resultater fra OLS estimationer efter kontrol for indvandringsalder

Afhængig variabel: Gennemførelse af ungdomsuddannelse inden for fem år efter grundskole	
Indvandringsalder	
<i>(Født i Danmark som referencekategori)</i>	
0 år	-0,039** (0,020)
1 år	-0,088*** (0,016)
2 år	-0,074*** (0,016)
3 år	-0,071*** (0,017)
4 år	-0,091*** (0,016)
5 år	-0,090*** (0,016)
6 år	-0,118*** (0,016)
7 år	-0,106*** (0,016)
8 år	-0,131*** (0,015)
9 år	-0,156*** (0,015)
10 år	-0,151*** (0,015)
11 år	-0,183*** (0,015)
12 år	-0,175*** (0,015)
13 år	-0,194*** (0,015)
14 år	-0,240*** (0,015)
15 år	-0,280*** (0,015)
Indvandrerstatus	
<i>(Etniske danske børn som referencekategori)</i>	
Indvandrerbaggrund	-0,030*** (0,011)
Indvandrerbaggrund · trend	0,004*** (0,001)
Indvandrerbaggrund · født i udlandet · trend	0,001 (0,001)
Kvinde	0,106***

(0,002)

Kontrolvariable

Grundskoleårgang	Ja
Oprindelsesland	Ja
Familiestruktur	Ja
Mor og fars socioøkonomiske status	Ja
- plus interaktion med indvandrerbaggrund	Ja
Mor og fars uddannelse	Ja
- plus interaktion med indvandrerbaggrund	Ja

Antal observationer	144.370
---------------------	---------

Tilpasset R ²	0,146
--------------------------	-------

NOTE: Robuste standardafvigelse er angivet i parentes. Registerdata fra Danmarks Statistik for en stikprøve af børn med indvandrerbaggrund fra ikke-vestlige lande og etniske danske børn som gennemfører 9. klasse mellem 1990 og 2007 (Stikprøveudvælgelse beskrives i afsnittet 'Datagrundlag og stikprøveudvælgelse'). Regressioner er baseret på en *lineær sandsynlighedsmodel*. Koefficienten på *indvandrerbaggrund* er beregnet som gennemsnittet af koefficienterne på oprindelsesland vægtet med indvandrerpopulationens andel. Yderligere er koefficienten på *indvandrerbaggrund* beregnet som gennemsnittet på tværs af 10 niveauer af socioøkonomisk status og tre uddannelsesniveauer plus missing indikatorer vægtet med den etniske danske fordeling. Kontrolvariablene inkluderer 18 årgangsdummies, 14 oprindelsesland dummies, 8 aldersgruppedummies for forældres alder ved barnets fødsel, 5 dummies for antal søskende, en skilsmisseuddummy, 10 dummies for forskellige niveauer for forældrenes socioøkonomiske status og 3 dummies for forældrenes uddannelsesniveau samt dummies for missing værdier. Urapporert konstant er inkluderet i alle regressioner. *** 1 pct. signifikansniveau. ** 5 pct. signifikansniveau. * 10 pct. signifikansniveau.

Tabel S.8: Resultater fra OLS estimationer efter børn med indvandrerbaggrund fra Tyrkiet er ekskluderet

Afhængig variabel: Gennemførelse af ungdomsuddannelse inden for fem år efter grundskole				
	(1)	(2)	(3)	(4)
Indvandrerbaggrund	-0,169*** (0,009)	-0,107*** (0,011)	-0,029*** (0,011)	0,001 (0,012)
Indvandrerbaggrund · født i udlandet	-0,163*** (0,011)	-0,182*** (0,013)	-0,145*** (0,011)	-0,150*** (0,014)
Indvandrerbaggrund · trend	0,005*** (0,001)	0,004*** (0,001)	0,004*** (0,001)	0,003*** (0,001)
Indvandrerbaggrund · født i udlandet · trend	0,006*** (0,001)	0,004*** (0,001)	0,003*** (0,001)	0,004*** (0,001)
Kvinde	0,104*** (0,003)	0,096*** (0,003)	0,107*** (0,002)	0,100*** (0,003)
Kontrolvariable				
Grundskoleårgang	Ja	Ja	Ja	Ja
Oprindelsesland	Nej	Nej	Ja	Ja
Familiestruktur	Nej	Nej	Ja	Ja
Mor og fars socioøkonomiske status	Nej	Nej	Ja	Ja
- plus interaktion med indvandrerbaggrund	Nej	Nej	Ja	Ja
Mor og fars uddannelse	Nej	Nej	Ja	Ja
- plus interaktion med indvandrerbaggrund	Nej	Nej	Ja	Ja
Antal observationer	144.370	131.571	144.370	131.571
Beskrivelse af stikprøve	Fuld	Uden Tyrkiet	Fuld	Uden Tyrkiet
Tilpasset R ²	0,048	0,038	0,143	0,138

NOTE: Robuste standardafvigelser er angivet i parentes. Registerdata fra Danmarks Statistik for en stikprøve af børn med indvandrerbaggrund fra ikke-vestlige lande og etniske danske børn som gennemfører 9. klasse mellem 1990 og 2007 (Stikprøveudvælgelse beskrives i afsnittet 'Datagrundlag og stikprøveudvælgelse'). Specifikation (2) og (4) er baseret på same stikprøveudvælgelseskriterier men ekskluderer børn med indvandrerbaggrund som har Tyrkiet som oprindelsesland. Regressioner er baseret på en *lineær sandsynlighedsmodel*. I specifikation (3) og (4) er koefficienterne på *indvandrerbaggrund* beregnet som gennemsnittet af koefficienterne på oprindelsesland vægget med indvandrerpopulationens andel. Yderligere er koefficienterne på *indvandrerbaggrund* beregnet som gennemsnittet på tværs af 10 niveauer af socioøkonomisk status og tre uddannelsesniveauer plus missing indikatorer vægget med den etniske danske fordeling. Kontrolvariablene inkluderer 18 årgangsdummies, 14 oprindelsesland dummies, 8 aldersgruppetummies for forældres alder ved barnets fødsel, 5 dummies for antal søskende, en skilsmisseuddummy, 10 dummies for forskellige niveauer for forældrenes socioøkonomiske status og 3 dummies for forældrenes uddannelsesniveau samt dummies for missing værdier. Urapporateret konstant er inkluderet i alle regressioner. *** 1 pct. signifikansniveau. ** 5 pct. signifikansniveau. * 10 pct. signifikansniveau.