

Helene Wind Fallesen

Undersøgelse af uddannelseskløften: Haler børn med indvandrerbaggrund ind på de etnisk danske børn?

Hvorfor eksisterer der en uddannelseskluft mellem etnisk danske børn og børn med indvandrerbaggrund fra ikke-vestlige lande? Og er kløften blevet reduceret over tid? Denne artikel bidrager til den eksisterende danske forskning ved at undersøge, hvad der kan forklare uddannelseskluftens eksistens samt kløftens udvikling over tid. Uddannelseskluftens undersøges ved at sammenligne gennemførelsesprocenten på ungdomsuddannelser for børn, der har færdiggjort 9. klasse i 1990 til 2007. Uddannelsesproduktionsfunktionen estimeres som en lineær sandsynlighedsmodel med OLS. Artiklen finder, at uddannelseskluftens er blevet reduceret over de seneste 20 år, og at forskelle i familiebaggrund forklarer en stor del af kløften. Yderligere findes en stærk sammenhæng mellem barnets alder på indvandringstidspunktet og uddannelsesmæssig præstation.

Ikke-vestlige indvandrere og deres børn udgør en voksende andel af befolkningen i flere europæiske lande – heriblandt Danmark. Vellykket integration af indvandrere er vigtig af flere grunde. Først og fremmest er det vigtigt, at alle mennesker bosat i Danmark har lige muligheder for at gennemføre en uddannelse og opnå succes på arbejdsmarkedet uafhængigt af socioøkonomisk baggrund og oprindelsesland. Dette er kernen i den danske velfærdsstat. Derudover står Danmark over for en udfordring i form af en ældende befolkning med færre mennesker i den arbejdsdygtige alder til at understøtte det voksende antal af ældre (Danmarks Statistik, 2014). Dette kan resultere i problemer med den finanspolitiske holdbarhed, men vellykket integration af indvandrere kan være en måde at afhjælpe problemet på (Schou, 2006).

I denne artikel fokuseres på børn med indvandrerbaggrund, og hvor godt de er integreret i det danske skolesystem. Dette undersøges ved at sammenligne gennemførelsesprocenter på ungdomsuddannelserne for etnisk danske børn og børn med indvandrerbaggrund, og hvordan denne forskel i gennemførelsesprocenterne har udviklet sig over tid. Hvis der er tegn på, at uddannelseskluftens er blevet reduceret over tid, og børn med indvandrerbaggrund konvergerer mod de etnisk danske børns uddannelsesmæssige præstation, kan det ses som en indikator på vellykket integration. At uddannelseskluftens eksisterer i Danmark er dokumenteret af blandt andre Colding, Husted og Hummelgaard (2009)

og Rangvid (2010), som også undersøger mulige forklaringer på kløftens eksistens. Udviklingen i uddannelseskløften er dog ikke tidligere undersøgt i en dansk kontekst. Forskning fra Norge viser, at børn med indvandrerbaggrund haler ind på de etnisk norske børn (Bratsberg, Raaum og Røed, 2012), og det er derfor interessant at undersøge, om den samme positive udvikling gælder i Danmark. Derudover har artiklen et særligt fokus på børn med indvandrerbaggrund, som er født i udlandet, og sammenhængen mellem deres uddannelsesmæssige præstation og alder på indvandringstidspunktet. Forskning fra Sverige viser, at børn, der ankommer senere, har en signifikant lavere uddannelsesmæssig præstation, og at udfordringen er særlig stor for børn, der ankommer, efter de er fyldt ni år (Böhlmark, 2008). I denne artikel undersøges det, om der kan findes en tilsvarende kritisk indvandringialder i Danmark.

I næste afsnit præsenteres den eksisterende forskning om uddannelseskløften samt denne artikels bidrag til forskningen. Herefter beskrives og diskuteres den empiriske tilgang herunder de estimerede modeller, de inkluderede indikatorer og den empiriske metode samt metodens implikationer. Datagrundlag og stikprøveudvælgelsen præsenteres, og afslutningsvis gennemgås den beskrivende analyse samt den empiriske analyses resultater.

Eksisterende viden om uddannelseskløften

Eksisterende forskning har både haft fokus på, om uddannelseskløften eksisterer, og hvad der kan forklare dens eksistens. Forskelle i familiebaggrund findes ofte som en vigtig del af forklaringen på kløften. Dustmann, Frattini og Lanza (2012) har undersøgt uddannelseskløften i flere OECD-lande og finder en stærk sammenhæng mellem børns testscore og forældres uddannelsesmæssige præstation. I visse lande forsvinder kløften helt, når der kontrolleres for forskelle i familiebaggrund. Todd og Wolpin (2007) får lignende resultater i deres undersøgelse af testcorekløften i USA. Rangvid (2010) undersøger forskelle i testscorer for etnisk danske børn og børn med indvandrerbaggrund, og hendes resultater viser, at forskelle i socioøkonomisk status for børn med indvandrerbaggrund og etnisk danske børn kan forklare en stor del af kløften, men at testscoren, selv efter der er kontrolleret for socioøkonomisk status, er lavere for både børn med indvandrerbaggrund, som er født i Danmark og i udlandet, end for etnisk danske børn.

Samtidig peger flere studier på, at familiebaggrund ikke har samme betydning for den uddannelsesmæssige præstation for den nationale majoritetsbefolkning og børn med indvandrerbaggrund. Støren og Helland (2010) undersøger forskelle i gennemførelsesprocenten på ungdomsuddannelserne i Norge og finder, at forældres uddannelsesniveau har mindre betydning for børn med

ikke-vestlig baggrund. Deres resultater indikerer, at etnisk norske børn har en større positiv effekt af at have højtuddannede forældre, hvorimod børn med indvandrerbaggrund oplever en mindre negativ effekt af at have forældre med lavere uddannelsesniveauer. At forældrevariable har mindre betydning for børn med indvandrerbaggrund understøttes også i studier af Schneeweis (2011), Jakobsen og Smith (2006) og Colding, Husted og Hummelgaard (2009).

Det er vigtigt at skelne mellem, om børn med indvandrerbaggrund er født i Danmark eller i udlandet, når uddannelseskløften undersøges. Børn, som er født i udlandet, har ikke nødvendigvis modtaget hele deres uddannelse i Danmark, og dette må påvirke deres uddannelsesmæssige præstation. Flere internationale studier har fokus på betydningen af barnets indvandringssalder for den uddannelsesmæssige præstation. I Holland finder Van Ours og Veenman (2006), at den kritiske indvandringssalder både afhænger af køn og oprindelsesland. En lignende analyse i Sverige viser, at der er en stærk sammenhæng mellem barnets alder ved indvandring og barnets uddannelsesmæssige præstation (Böhlmark, 2008). Børn, der ankommer senere, har en signifikant lavere præstation, og udfordringen er særlig stor for børn, der ankommer, efter de er fyldt ni år.

Størrelsen på uddannelseskløften kan også variere med køn. Rangvid (2010) finder en signifikant større kløft mellem tyrkiske piger og etnisk danske piger end mellem tyrkiske drenge og etnisk danske drenge, men finder ikke signifikante kønsforskelle, når hun undersøger børn med andre oprindelseslande. I Sverige taler kønsforskellen til pigernes fordel (Støren og Helland, 2010). De kvindelige studerende har højere gennemførelsesprocenter på ungdomsuddannelserne end de mandlige studerende, og kønsforskellen er større blandt studerende med indvandrerbaggrund.

Colding, Husted og Hummelgaard (2009) undersøger uddannelseskløften i Danmark ved hjælp af en dynamisk diskret model for uddannelsesvalg. De placerer børn, som er indvandret til Danmark efter seksårsalderen, i en separat gruppe for at undersøge, om de oplever større udfordringer end børn med indvandrerbaggrund, som er født i Danmark. De finder, at frafaldsprocenten på erhvervsuddannelserne er væsentligt højere for børn med indvandrerbaggrund end for etnisk danske børn, og at børn, der er indvandret efter seksårsalderen, oplever særlige udfordringer, da færre af dem formår at påbegynde en ungdomsuddannelse. Yderligere viser resultaterne, at frafaldsprocenten reduceres, hvis børn med indvandrerbaggrunds forældrekaraktistika styrkes.

Uddannelseskløften og mulige forklaringer på dens eksistens er blevet undersøgt i flere danske og internationale studier, og ovenstående liste er langt fra udtømmende. Denne artikel bidrager dog til litteraturen på flere områder.

Ingen af de danske studier undersøger, mig bekendt, den direkte betydning af indvandringsalder. Derudover bidrager denne artikel til den eksisterende forskning ved ikke kun at undersøge størrelsen og kilderne til uddannelseskløften, men også hvordan kløften har udviklet sig over tid. Udviklingen i uddannelseskløften er blevet undersøgt i både Norge og Tyskland (Bratsberg, Raaum og Røed, 2012; Riphahn, 2003). I Tyskland findes det lettere overraskende resultat, at uddannelseskløften mellem etnisk tyske børn og børn med indvandrerbaggrund, som er født i Tyskland, er blevet signifikant større over tid. Resultaterne fra Norge er dog mere positive, og de finder, at kløften er blevet væsentligt reduceret på tværs af de seneste to årtier. De forskellige resultater fra Norge og Tyskland understreger yderligere, hvorfor det er interessant at undersøge udviklingen i Danmark.

Empirisk tilgang

Uddannelsesproduktionsfunktionen

Uddannelsesproduktionsfunktionen anvendes i flere studier, hvor formålet er at undersøge den empiriske effekt af forskellige variable på barnets uddannelsesmæssige præstation. Tankegangen er, at den uddannelsesmæssige præstation skabes ved hjælp af en input-output-proces, hvor forskellige input i fællesskab skaber outputtet. Outputtet er et uddannelsesmæssigt mål, og input er fx skole-, forældre- og individkarakteristika.

Når uddannelsesproduktionsfunktionen skal opskrives og estimeres, opstår der en række væsentlige spørgsmål. Hvad er et passende mål for barnets præstation? Hvilke inputvariable bør indgå, og hvordan skal disse måles? Disse emner diskuteres blandt andet af Hanushek (1979) i hans artikel om de konceptuelle og empiriske udfordringer vedrørende estimeringen af uddannelsesproduktionsfunktionen. Der er dog konsensus om en model for uddannelsesmæssig præstation i stil med nedenstående (se fx Hanushek (1973) eller Todd og Wolpin (2003) for lignende ligninger):

$$A_{it} = f(F_i(t), S_i(t), I_i)$$

Den uddannelsesmæssige præstation for barn i på tidspunkt t er en funktion af en række individ- og forældre-karakteristika (F_i), skolevariable (S_i), som er relevante for barn i og barn i s medfødte evner (I_i). Forældre- og skolevariable skal være kumulative til tidspunkt t , dvs. at alle nuværende og tidligere forældre- og skoleinputs skal inkluderes i estimeringen.

Ideelt set estimeres uddannelsesproduktionsfunktionen i ovenstående form, når man skal undersøge effekten af forskellige faktorer på barnets uddannel-

sesmæssige præstation, men i praksis er det sjældent muligt. For det første er barnets medfødte evner pr. definition uobserverbare, og at finde et passende mål for disse kan være vanskeligt. For det andet er det sjældent muligt at inkludere alle tidligere og nuværende mål for forældre- og skolevariable. Derudover er udfordringen ved at anvende klassisk produktionsfunktionsteori på uddannelse, at der ikke er et homogent output, og at et passende mål for output derfor skal vælges på samme måde, som inputvariablene vælges.

De estimerede modeller i denne artikel tager udgangspunkt i uddannelsesproduktionsfunktionen, men med modifikationer. Da formålet er at sammenligne den uddannelsesmæssige præstation for etnisk danske børn og børn med indvandrerbaggrund, inkluderes dummyvariable for indvandrerbaggrund. En anden modifikation er, at skolevariable ikke er tilgængelige og derfor er udeladt i modellerne. Yderligere er der intet mål for barnets medfødte evner. Hvis skolevariable og barnets medfødte evner har betydning for barnets uddannelsesmæssige præstation, kan udeladelse af disse variable resultere i endogenitetsproblemer, hvilket diskuteres længere nede i afsnittet.

Ovenstående uddannelsesproduktionsfunktion lægger samtidig op til, at alle tidligere og nuværende familievariable skal inkluderes. Dette er muligt med det tilgængelige data, men da den samme inputvariabel målt på forskellige tidspunkter højst sandsynligt er stærkt korreleret med sig selv, vælges et måletidspunkt i stedet. Alle forældrevariable i modellerne i denne artikel er derfor målt, når barnet er 15 år. Baseret på ovenstående modifikationer kan de estimerede modeller opstilles.

De estimerede modeller og indikatorer

Den uddannelsesmæssige præstation for barn i er formuleret som:

$$y_i = \alpha_0 + \alpha_1 IB_i + \alpha_2 IB_i \cdot UDLAND_i + \alpha_3 IB_i \cdot t + \alpha_4 IB_i \cdot UDLAND_i \cdot t + \alpha_5 X_i + \alpha_6 F_i + \varepsilon_i$$

hvor IB_i er en dummyvariabel, som angiver, om barn i har indvandrerbaggrund, og $UDLAND_i$ er en dummyvariabel, som angiver, om barn i er født i udlandet. t svarer til den grundskoleårgang,¹ barn i tilhører. X_i er en vektor af individkarakteristika, F_i er en vektor af familiekarakteristika, og ε_i er et fejlsled. Den estimerede model siger derfor, at den uddannelsesmæssige præstation for barn i bestemmes af en række individuelle karakteristika, familiekarakteristika og af, om barn i har indvandrerbaggrund og er født i udlandet.

Der er flere måder at måle barnets uddannelsesmæssige præstation (y_i) på. I undersøgelser, som sammenligner unge i forskellige lande, er det særlig brug-

bart at bruge standardiserede testscorer fra PISA eller TIMSS tests, da de er sammenlignelige på tværs af lande. At anvende resultater fra en specifik test eller eksamen giver dog et øjebliksbillede af den enkelte elevs uddannelsesmæssige præstation. En anden tilgang er derfor at måle den uddannelsesmæssige præstation som et opnået uddannelsesniveau i stedet for præstationen i en specifik test eller eksamen. I denne artikel bruges *gennemførelse af ungdomsuddannelse inden for fem år efter afslutning af grundskolen* som mål for præstation. Et lignende mål anvendes blandt andet af Riphahn (2003), Støren og Helland (2010) og Bratsberg, Raaum og Røed (2012). I Danmark er der stor politisk fokus på ungdomsuddannelserne og et mål om, at 95 pct. af en ungdomsårgang skal gennemføre en ungdomsuddannelse (Velfærdsaftalen, 2006). En gennemført ungdomsuddannelse giver unge de nødvendige kompetencer til at kunne fortsætte i et faglært job eller videreudanne sig. Med dette outputmål er det dermed muligt at undersøge, om børn med indvandrerbaggrund i lige så høj grad som etnisk danske børn formår at fortsætte i uddannelsessystemet efter grundskolens afslutning.

Artiklens formål er at undersøge den uddannelsesmæssige præstation for etnisk danske børn sammenlignet med børn med indvandrerbaggrund, og derfor inkluderes en dummyvariabel for at have indvandrerbaggrund. Baseret på om de er født i udlandet eller Danmark, kan børn med indvandrerbaggrund opdeles i to grupper: indvandrere og efterkommere. For at gøre det muligt at undersøge om der er forskel på disse to grupper, inkluderes en dummyvariabel for at være født i udlandet. Det vil sige, at den estimerede koefficient på IB_i giver den gennemsnitlige forskel i gennemførelsesprocenten på ungdomsuddannelserne mellem efterkommere og etnisk danske børn, mens koefficienten på $IB_i \cdot UDLAND_i$ giver den yderligere forskel for indvandrere. Det er dermed koefficienten på disse to dummyvariable, der angiver størrelsen på uddannelseskløften.

Trendvariabel (t) interageret med indvandrerbaggrund inkluderes for at kunne undersøge, om børn med indvandrerbaggrund haler ind på de etnisk danske børn, dvs. om senere grundskoleårgange af børn med indvandrerbaggrund klarer sig bedre i forhold til etnisk danske børn end tidligere grundskoleårgange. Koefficienterne på trendvariablene afspejler den gennemsnitlige årlige ændring i effekten af at have indvandrerbaggrund og være født i udlandet. Det er dermed de interagerede trendvariable, der angiver, om børnene med indvandrerbaggrund haler ind på de etnisk danske børn, eller modsat om kløften bliver større over tid.

Som kontrolvariable inkluderes en lang række individ- og familiekarakteristika. Som individuelle karakteristika inkluderes køn, grundskoleårgang

og oprindelsesland. Flere studier viser, at piger uddannelsesmæssigt klarer sig bedre end drenge, og kønsvariablen inkluderes for at kontrollere for dette (se fx Machin og McNally, 2005). Grundskoleårgangen er inkluderet for at kontrollere for en generel trend i gennemførelsesprocenten som fx ændringer i skolesystemet, som er ens for de etnisk danske børn og børnene med indvandrerbaggrund. Flere studier viser betydelige uddannelsesmæssige forskelle på børn med indvandrerbaggrund med forskellige oprindelseslande, og oprindelsesland inkluderes derfor for at tage højde for disse forskelle (Rangvid, 2010; Schneeweis, 2011).

Familiekaraktistika kan i denne artikel opdeles i familiestruktur og familiebaggrund. I forhold til familiestruktur kontrolleres der for, om forældrene er skilt eller ej, antallet af børn i familien og forældrenes alder ved barnets fødsel. Antallet af børn kan både påvirke det enkelte barns uddannelsesmæssige præstation positivt og negativt. Det kan være positivt med søskende, der kan hjælpe med lektier osv., men modsat kan flere børn betyde mindre opmærksomhed fra forældrene til det enkelte barn, hvilket kan påvirke negativt. Det sidstnævnte understøttes af Adli, Louichi og Tamouh (2010). Familiebaggrund inkluderer mors og fars års skolegang og socioøkonomiske status. Et tæt forhold mellem forældrenes uddannelse og socioøkonomiske status og barnets uddannelsesmæssige præstation er veldokumenteret (Deding og Hussain, 2005; Dickson, Gregg og Robinson, 2013; Weinberg, 2001). Variablene inkluderes for at kontrollere for dette. Forskningsresultater indikerer, at etnisk danske børn og børn med indvandrerbaggrund påvirkes forskelligt af familiebaggrund og for at tillade dette, indgår familiebaggrund både som en variabel for sig selv og interageret med indvandrerbaggrund.

Indledningsvis estimeres model (1), som kun inkluderer de fire primære variable (IB_p , $IB_i \cdot UDLAND_p$, $IB_i \cdot t$ og $IB_i \cdot UDLAND_i \cdot t$) samt køn og årgang for at få et indtryk af størrelsen på den "rå" uddannelseskloft.² Herefter inkluderes alle individ- og familiekaraktistika i model (2), og det bliver muligt at se, hvor stor uddannelseskloften er, når der tages højde for forskelle i blandt andet forældrenes socioøkonomiske status. For at undersøge om betydningen af at have indvandrerbaggrund og indhentningsraten varierer med køn, estimeres model (3), hvor de fire primære variable både indgår for sig selv og interageret med køn.³ Derudover fokuseres på betydningen af barnets indvandringialder, og derfor estimeres en model, hvor dummyvariablen for at være født i udlandet erstattes med dummyvariable for hver mulig indvandringialder.

Metode

Modellerne estimeres som *lineære sandsynlighedsmodeller* med OLS. Den lineære sandsynlighedsmodel er kritiseret af mange økonometrikere men er samtidig ofte anvendt i empirisk arbejde grundet dens enkelhed (Wooldridge, 2003). Den primære udfordring med modellen er, at den kan forudsige sandsynligheder under 0 og over 1. Samtidig betyder lineariteten, at den marginale effekt af hver forklarende variable er konstant. Disse udfordringer kan undgås ved at bruge mere sofistikerede binære responsemodeller såsom den ikke-lineære probit- eller logitmodel, der anvendes i flere studier om uddannelseskloften (se fx Riphahn, 2003; Støren og Helland, 2010).

I denne artikel fokuseres på de marginale effekter, dvs. på den marginale betydning af at have indvandrerbaggrund fremfor at være etnisk dansk, så udfordringen med forudsagte sandsynligheder under 0 eller over 1 er mindre relevant. Yderligere håndteres udfordringen med den lineære effekt ved at bruge dummyvariable som forklarende variable. Da alle de forklarende variable i denne analyse indgår som dummyvariable, er den lineære effekt i den lineære sandsynlighedsmodel ikke en udfordring. Hvis modellen yderligere er *saturated*, dvs. at hvis alle de forklarende variable er dummyvariable for gensidigt udelukkende og udtømmende kategorier, er den lineære sandsynlighedsmodel fuldstændig generel, og estimation med logit, probit og den lineære sandsynlighedsmodel vil give samme resultater (Wooldridge, 2010; Angrist og Pischke, 2009).

Modellen estimeres med OLS, hvilket giver unbiased estimater, hvis der ikke er endogenitet. Antagelsen om ingen endogenitet siger, at fejleddet skal være ukorreleret med de forklarende variable (Wooldridge, 2003). Alle variable, som ikke inkluderes eksplicit som en forklarende variabel, vil være indeholdt i fejleddet. Det vil sige, at hvis der er variable, som påvirker barnets uddannelsesmæssige præstation, som ikke er inkluderet i de estimerede modeller, og disse variable er korrelerede med en eller flere af de forklarende variable, vil der eksistere et endogenitetsproblem. I denne type estimationer er endogenitetsproblemet især knyttet til barnets egne evner. I de estimerede modeller indgår forældrenes uddannelsesniveau og socioøkonomiske status, som højst sandsynligt er korrelerede med barnets egne evner. Derfor kan effekten af familiebaggrund være biased opad, fordi den også fanger effekten af barnets egne evner. Når der anvendes OLS, skal man derfor være opmærksom på muligheden for biased estimater, når man fortolker resultaterne. Endogenitetsproblemet er dog forsøgt reduceret ved at inkludere en lang række relevante forklarende variable i estimationerne.

Når man estimerer en lineær sandsynlighedsmodel med OLS, opstår et heteroskedasticitetsproblem som gør, at standardafvigelserne og de tilhørende t -statistikker bliver ugyldige. Den simpleste måde at løse heteroskedasticitetsproblemet er fortsat at bruge OLS, men samtidig bruge heteroskedasticitetsrobuste standardafvigelser (Wooldridge, 2010). Alle rapporterede standardafvigelser i denne artikel er derfor heteroskedasticitetsrobuste standardafvigelser.

Datagrundlag og stikprøveudvælgelse

Datagrundlaget stammer fra tre registre: befolkningsregisteret, uddannelsesregisteret og arbejdsstyrkeregisteret. Befolkningsregistret giver adgang til demografiske karakteristika såsom køn, alder, indvandrerstatus, familiestruktur, oprindelsesland og indvandringsdato. Uddannelsesregisteret indeholder information om højest fuldførte uddannelse,⁴ og data om beskæftigelsesstatus er tilgængelig i arbejdsstyrkeregisteret.

Børn med indvandrerbaggrund er i denne artikel begrænset til børn fra ikke-vestlige lande, da deres uddannelsesmæssige præstation typisk varierer mest fra de etnisk danske børns.⁵ Flere studier af uddannelseskloften vælger ligeledes at fokusere på børn fra ikke-vestlige eller mindre udviklede lande (se fx Colting, Husted og Hummelgaard, 2009; Bratsberg, Raaum og Røed, 2012). Der er tilgængelig data for børn, der gennemfører 9. klasse mellem 1980 og 2012, men da der var et begrænset antal ikke-vestlige børn i Danmark i 1980'erne, indeholder analysen børn, som har færdiggjort 9. klasse i 1990 og fremefter.

I artiklen følges børn efter færdiggørelsen af 9. klasse for at undersøge, om de gennemfører en ungdomsuddannelse inden for fem år. Den typiske varighed for en ungdomsuddannelse er omkring tre år, og den femårige periode giver dermed de unge mulighed for at gå i 10. klasse inden ungdomsuddannelsen og for de fleste også mulighed for at ændre uddannelsesretning undervejs. Derfor vælges fem år som perioden inden for hvilken, børnene skal have gennemført en ungdomsuddannelse efter afslutningen af 9. klasse, og analysen er baseret på de 18 grundskoleårgange, som afslutter 9. klasse mellem 1990 og 2007.

Børn med indvandrerbaggrund defineres som børn, hvor ingen af forældrene både er dansk statsborger og født i Danmark. Disse børn inddrages yderligere efter, om de selv er i født i Danmark, dvs. er efterkommere, eller i udlandet, dvs. er indvandrere (jf. Danmarks Statistiks definition). I litteraturen er det forskelligt, hvordan man klassificerer børn, hvor den ene forælder er indvandrere. I visse danske studier inkluderes de som etnisk danske børn (se fx Rangvid, 2007). I et internationalt sammenligningsstudie af Dustmann, Frattini og Lanzara (2012) ekskluderes denne gruppe børn i størstedelen af analysen, og når de inkluderes, defineres de som børn med indvandrerbaggrund. Bratsberg,

Raaum og Røed (2012), som har lavet en analyse af uddannelseskløften i Norge, ekskluderer disse børn i deres analyse. For at gøre inddelingen i denne artikel så tydelig som mulig ekskluderes børn, hvor den ene forælder er indvandrer. Det vil sige, at etnisk danske børn her refererer til børn, hvor begge forældre er danske statsborgere og født i Danmark, hvilket er en snævrere definition end Danmarks Statistiks definition af etniske danskere.

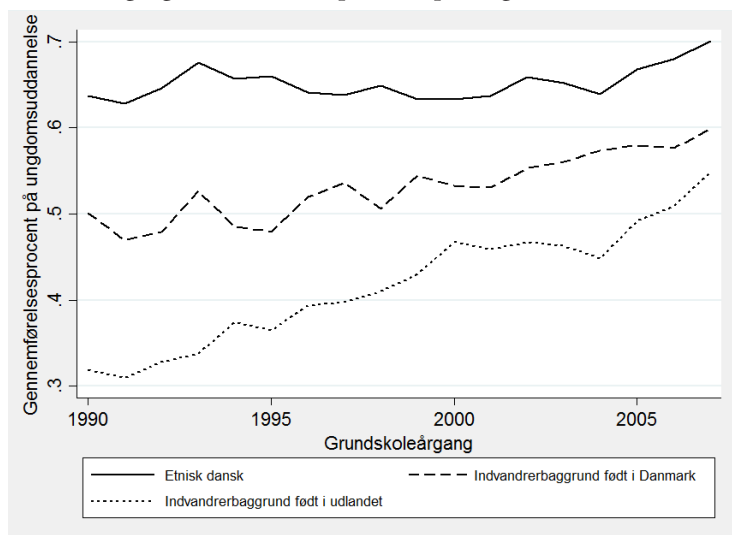
Børn er ekskluderet fra analysen, hvis de har tilbragt et helt kalenderår eller mere i udlandet i løbet af grundskolen eller de fem efterfølgende år. Dette gøres, fordi børn, der har været i udlandet og derefter vender tilbage til Danmark, kan være påvirket af andre faktorer, end børnene der blev i Danmark, hvilket gør det sværere at sammenligne deres uddannelsesmæssige præstation. For børn med indvandrerbaggrund, som er født i udlandet, vælges 15 år som den højeste alder ved indvandring. For at undgå at stikprøven inkluderer voksne, som gennemfører 9. klasse senere i deres liv, inkluderer stikprøven udelukkende elever, som er mellem 15 og 18 år, når de færdiggør 9. klasse. Tabel S.2 i det supplerende materiale viser reduktionen i stikprøven for hvert af ovenstående udvælgelseskriterier.

Børn med indvandrerbaggrund udtrækkes fra et datasæt med hele den danske befolkning, mens de etnisk danske børn udtrækkes fra et datasæt med en 10 pct. stikprøve af den danske befolkning. Baseret på ovenstående udvælgelseskriterier betyder det, at datasættet består af 144.370 børn, hvoraf 88.915 er etnisk danske børn, og 55.455 er børn med indvandrerbaggrund, som yderligere kan opdeles i 24.005 børn født i Danmark og 31.450 børn født i udlandet.

Beskrivende analyse

Figur 1 viser gennemførelsesprocenten på ungdomsuddannelserne for hver årgang, som gennemfører 9. klasse mellem 1990 og 2007, separat for etnisk danske børn og børn med indvandrerbaggrund som er født i Danmark og i udlandet. Gennemførelsesprocenten angiver, hvor stor en andel af børnene i hver grundskoleårgang der gennemfører en ungdomsuddannelse inden for fem år. Figur 1 understøtter resultaterne fra tidligere dansk forskning, da den viser, at børn med indvandrerbaggrund gennemsnitligt har en lavere gennemførelsesprocent end etnisk danske børn. Udviklingen i uddannelseskjøften over tid er dog positiv, og figuren viser, at den uddannelsesmæssige præstation for børn med indvandrerbaggrund nærmer sig de etnisk danske børns præstation. Spørgsmålet er, om den positive udvikling er en konsekvens af ændringer i indvandrerpopulationen over tid. Det kan være tilfældet, at børnene med indvandrerbaggrund i de seneste årgange har andre oprindelseslande eller mere

Figur 1: Udvikling i gennemførelsesprocent på ungdomsuddannelse



Note: Gennemførelsesprocenten på ungdomsuddannelserne afspejler andelen af en grundskoleårgang, som gennemfører en ungdomsuddannelse inden for fem år efter afslutningen af grundskolen (9. klasse). Registerdata fra Danmarks Statistik for en stikprøve af børn med indvandrerbaggrund fra ikke-vestlige lande og etnisk danske børn som gennemfører 9. klasse mellem 1990 og 2007 (stikprøveudvælgelse beskrives i afsnittet ”Datagrundlag og stikprøveudvælgelse”).

ressourcestærke forældre, og at dette er forklaringen på forbedringen i gennemførelsesprocenten over tid. Dette undersøges i den empiriske analyse.

Tabel 1 viser den beskrivende statistik for udvalgte forklarende variable, hvor det blandt andet fremgår, at børn med indvandrerbaggrund i gennemsnit har fædre med en lavere socioøkonomisk status end etnisk danske børn. Tabel S.3 til S.5 i det supplerende materiale viser stikprøvens fordeling på grundskoleårgange, oprindelsesland samt den beskrivende statistik for de resterende variable.

Tabel 1: Beskrivende statistik for udvalgte variable: middelværdi (standardafvigelse i parentes)

Variable	Indvandrerbaggrund		
	Etnisk dansk	Født i Danmark	Født i udlandet
Gennemført ungdomsuddannelse inden for fem år efter afslutning af 9. klasse	0,652 (0,476)	0,548 (0,498)	0,438 (0,496)
Kvinde	0,489 (0,500)	0,500 (0,500)	0,476 (0,499)
Fars socioøkonomiske status			
Beskæftiget			
Topledere	0,034 (0,182)	0,009 (0,093)	0,004 (0,059)
Lønmodtager højeste niveau	0,180 (0,384)	0,037 (0,189)	0,028 (0,165)
Lønmodtager mellemste niveau	0,126 (0,332)	0,035 (0,183)	0,020 (0,139)
Lønmodtager grundniveau	0,251 (0,434)	0,162 (0,368)	0,092 (0,289)
Ufaglærte	0,046 (0,209)	0,040 (0,197)	0,028 (0,166)
Ander lønmodtagere	0,112 (0,315)	0,181 (0,385)	0,091 (0,288)
Selvstændige og medarbejdende ægtefælle	0,117 (0,321)	0,133 (0,339)	0,058 (0,233)
Arbejdsløs			
Arbejdsløs	0,035 (0,183)	0,110 (0,313)	0,100 (0,300)
Uden for arbejdsstyrken			
Under uddannelse	0,001 (0,035)	0,009 (0,092)	0,030 (0,171)
Andre uden for arbejdsstyrken	0,062 (0,242)	0,243 (0,429)	0,344 (0,475)
Missing	0,036 (0,187)	0,041 (0,197)	0,205 (0,404)
Antal observationer	88.915	24.005	31.450

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

Empirisk analyse

Grundmodellen

Tabel 2 viser resultaterne af estimationerne. Resultaterne for model (1) viser, at børn med indvandrerbaggrund har en signifikant lavere gennemførelsesprocent end etnisk danske børn. Dette er som forventet og i tråd med, hvad figur 1 viste. Den estimerede gennemsnitlige forskel på gennemførelsesprocenten

Table 2: Resultater fra OLS estimationer

	(1)	(2)	(3)
Afhængig variabel: Gennemførelse af ungdomsuddannelse inden for fem år efter grundskole			
Indvandrerstatus			
(Etnisk danske børn som referencekategori)			
Indvandrerbaggrund	-0,169*** (0,009)	-0,029*** (0,011)	-0,039*** (0,014)
Indvandrerbaggrund · kvinde			0,021 (0,018)
Indvandrerbaggrund · født i udlandet	-0,163*** (0,011)	-0,145*** (0,011)	-0,113*** (0,016)
Indvandrerbaggrund · født i udlandet · kvinde			-0,066*** (0,023)
Indvandrerbaggrund · trend	0,005*** (0,001)	0,004*** (0,001)	0,003*** (0,001)
Indvandrerbaggrund · trend · kvinde			0,003* (0,001)
Indvandrerbaggrund · født i udlandet · trend	0,006*** (0,001)	0,003*** (0,001)	0,002 (0,001)
Indvandrerbaggrund · født i udlandet · trend · kvinde			0,003* (0,002)
Kvinde	0,104*** (0,003)	0,107*** (0,002)	0,084*** (0,003)
Kontrolvariable			
Grundskoleårgang	Ja	Ja	Ja
Oprindelsesland	Nej	Ja	Ja
Familiestruktur	Nej	Ja	Ja
Mor og fars socioøkonomiske status	Nej	Ja	Ja
- plus interaktion med indvandrerbaggrund	Nej	Ja	Ja
Mor og fars uddannelse	Nej	Ja	Ja
- plus interaktion med indvandrerbaggrund	Nej	Ja	Ja
Observationer	144.370	144.370	144.370
Justeret R ²	0,048	0,143	0,145

Note: Robuste standardafvigelse er angivet i parentes. Registerdata fra Danmarks Statistik for en stikprøve af børn med indvandrerbaggrund fra ikke-vestlige lande og etnisk 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) og (3), er koefficienterne på indvandrerbaggrund beregnet som gennemsnittet af koefficienterne på oprindelsesland vægget med indvandrerpopulationens andel. Yderligere i specifikation (2) og (3) er koefficienterne på indvandrerbaggrund beregnet som gennemsnittet på tværs af ti niveauer af socioøkonomisk status og tre uddannelsesniveauer plus missing indikatorer vægget med den etnisk danske fordeling. Koefficienten på indvandrerbaggrund · kvinde i specifikation (3) er beregnet som gennemsnittet af koefficienterne på oprindelsesland interageret med køn og vægget med den kvindelige indvandrerpopulationens andel og er yderligere beregnet som gennemsnittet på tværs af ti niveauer af socioøkonomisk status og tre uddannelsesniveauer plus missing indikatorer vægget med den etnisk danske fordeling for kvinder. Kontrolvariablene inkluderer 18 årgangsdummies, 14 oprindelsesland dummies, otte aldersgruppedummies for forældres alder ved barnets fødsel, fem dummies for antal søskende, en skilsmissegummy, ti dummies for forskellige niveauer for forældrenes socioøkonomiske status og tre dummies for forældrenes uddannelsesniveau samt dummies for missing værdier. Urapporтерet konstant er inkluderet i alle regressioner. *** 1 pct. signifikansniveau. ** 5 pct. signifikansniveau. * 10 pct. signifikansniveau.

for etnisk danske børn og børn med indvandrerbaggrund, som er født i Danmark, er 16,9 procentpoint. Forskellen mellem børnene med indvandrerbaggrund, som er født i udlandet, og etnisk danske børn er hele 33,2 procentpoint (0,169 + 0,163). Trendvariablene viser dog, at forskellene reduceres over tid. De estimerede trends er positive og signifikante for både børn med indvandrerbaggrund, som er født i Danmark og i udlandet. De danskfødte børn med indvandrerbaggrund nærmer sig de etnisk danske børn med 0,5 procentpoint om året, og børnene født i udlandet nærmer sig med 1,1 procentpoint (0,005 + 0,006). Resultaterne viser derfor, at der eksisterer en betydelig uddannelseskløft i Danmark, men at kløften reduceres over tid.

Resultaterne for model (2) viser, at en stor del af uddannelseskløften mellem børn med indvandrerbaggrund, som er født i Danmark, og etnisk danske børn kan forklares af forskelle i familiebaggrund.⁶ Efter kontrol af individ- og familiekaraktistika er den tilbageværende kløft på 2,9 procentpoint. Det vil sige, at hvis børn med indvandrerbaggrund, som er født i Danmark, havde samme familiebaggrund som de etnisk danske børn, ville de kun have en gennemførelsesprocent, der er 2,9 procentpoint lavere end de etnisk danske børns. For børn med indvandrerbaggrund, som er født i udlandet, forklarer forskelle i familiebaggrund også en del af kløften. Efter kontrol af individ- og familiekaraktistika har de dog fortsat en gennemførelsesprocent, som er 17,4 procentpoint lavere end de etnisk danske børn (0,029 + 0,145).

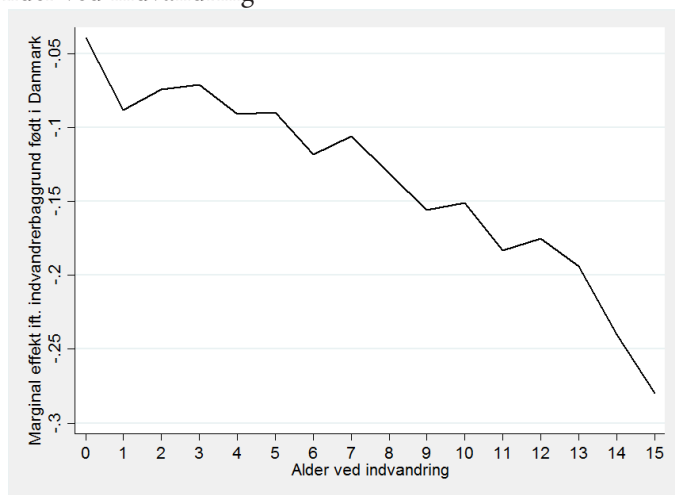
Det er vigtigt at bemærke, at trendvariablene fortsat er positive og signifikante, efter der er kontrolleret for en række karakteristika. Det betyder, at den positive udvikling ikke udelukkende kan forklares af ændringer i sammensætningen af børnene med indvandrerbaggrund over tid. Udviklingen viser, at der har været en forbedring i den uddannelsesmæssige præstation for børn med indvandrerbaggrund på tværs af grundskoleårgange, hvilket kan ses som en indikator på vellykket integration.

I model (3) tillades betydningen af indvandrerbaggrund og trendvariablene at variere med køn. Resultaterne viser, at piger med indvandrerbaggrund haler hurtigere ind på de etnisk danske piger end drenge med indvandrerbaggrund haler ind på de etnisk danske drenge. De danskfødte piger med indvandrerbaggrund nærmer sig de etnisk danske piger med 0,6 procentpoint om året, mens de danskfødte drenge med indvandrerbaggrund har en indhentningsrate på 0,3 procentpoint. For børnene med indvandrerbaggrund som er født i udlandet er raten henholdsvis 0,9 procentpoint for pigerne og 0,3 procentpoint for drengene.

Betydning af alder på indvandringstidspunktet

Figur 2 viser den marginale effekt af at indvandre til Danmark i forskellige aldre sammenlignet med at være et barn med indvandrerbaggrund, som er født i Danmark. Figuren viser, at barnets indvandringssalder har stor betydning for barnets sandsynlighed for at gennemføre en ungdomsuddannelse. Der er ikke stor forskel på gennemførelsesprocenten for et barn, der indvandrer som 1-årig eller 5-årig – alt andet lige. Dette indikerer, at hvis børnene ankommer til Danmark før skolestartsalderen, betyder det ikke meget for deres uddannelsesmæssige præstation, om de kommer et par år efter fødslen. Efter skolestartsalderen bliver den negative betydning af at ankomme ét år senere større og større. Et barn, som er født i udlandet og indvandrer til Danmark som 7-årig, har en gennemførelsesprocent, som er 10 procentpoint lavere end de danskfødte børn med indvandrerbaggrund, og ulempen øges til 13 og 15 procentpoint, hvis barnet ankommer som 8- eller 9-årig. Resultaterne viser dermed, at børn, der ankommer til Danmark, efter de er fyldt syv år, har en særlig udfordring, som øges med senere indvandring. Børn, som indvandrer som 15-årige har en gennemførelsesprocent, som er 28 procentpoint lavere end børn med indvandrerbaggrund, som er født i Danmark.

Figur 2: Alder ved indvandring



Note: Referencekategori er barn med indvandrerbaggrund, som er født i Danmark. Fulde resultater fra estimationen findes i tabel S.7 i det supplerende materiale. Registerdata fra Danmarks Statistik for en stikprøve af børn med indvandrerbaggrund fra ikke-vestlige lande og etnisk danske børn som gennemfører 9. klasse mellem 1990 og 2007 (stikprøveudvælgelse beskrives i afsnittet ”Datagrundlag og stikprøveudvælgelse”).

Resultaterne for model (2) i tabel 2 viste, at trendvariablen for børn med indvandrerbaggrund, som er født i udlandet, var positiv og signifikant. Det betyder, at børn med indvandrerbaggrund, som er født i udlandet, haler hurtigere ind på de etnisk danske børn end danskfødte børn med indvandrerbaggrund. Dette er umiddelbart et overraskende resultat, og når børnene opdeles efter deres indvandringsalder viser det sig også, at trendvariablen for børnene, som er født i udlandet, bliver insignifikant (se supplerende materiale tabel S.7). Det tyder derfor på, at den positive trend var et resultat af ændringer i indvandringsalderen på tværs af grundskoleårgangene.

Robusthedstest og diskussion

Robustheden af analysens resultater tjekkes ved at estimere model (1) og (2) uden børn med indvandrerbaggrund fra Tyrkiet. En stor andel af børnene med indvandrerbaggrund stammer fra Tyrkiet, og hvis børn med tyrkisk baggrund har en lav uddannelsesmæssig præstation, kan hele uddannelseskløften være drevet af børnene fra Tyrkiet. Resultaterne af estimationerne, hvor børn med Tyrkiet som oprindelsesland er ekskluderet, findes i tabel S.8 i det supplerende materiale. Resultaterne viser, at uddannelseskløften reduceres, når Tyrkiet er ekskluderet, men fortsat eksisterer. Dog forklarer forskelle i familiebaggrund hele uddannelseskløften mellem etnisk danske børn og børn med indvandrerbaggrund, som er født i Danmark, når familievariable inkluderes i estimationerne, hvor Tyrkiet er ekskluderet. Derudover ændres den positive trend kun marginalt og er fortsat signifikant. Det vil sige, at uddannelseskløften i Danmark ikke udelukkende er drevet af børn med indvandrerbaggrund fra Tyrkiet, og trenden er signifikant og positiv, både når Tyrkiet er inkluderet og ekskluderet i analysen.

En anden væsentlig udfordring er endogenitetsproblemet, og hvordan det påvirker resultaterne. Dette er beskrevet i metodeafsnittet. Den anvendte empiriske tilgang løser ikke endogenitetsproblemet, og man skal derfor være opmærksom på risikoen for biased estimater, når man fortolker resultaterne. Resultaterne skal derfor tolkes med forsigtighed og ses som sammenhænge snarere end kausale forhold.

Konklusion

Vellykket integration af indvandrere bliver mere og mere vigtig, da de udgør en voksende del af den danske befolkning. Uddannelse er en vigtig del af integrationsprocessen, og denne artikel undersøger den uddannelsesmæssige præstation for børn med indvandrerbaggrund. Analysen er baseret på de 18 grundskoleårgange, der har afsluttet 9. klasse mellem 1990 og 2007 og viser,

at en lavere andel af børnene med indvandrerbaggrund gennemfører en ungdomsuddannelse inden for fem år end blandt de etnisk danske børn.

Udviklingen i uddannelseskløften over tid er dog positiv. Kløften er blevet væsentligt mindre, og den positive udvikling kan ikke blot tilskrives en ændring i sammensætningen af indvandrerbefolkningen over tid. Der er tegn på, at børn med indvandrerbaggrund haler ind på de etnisk danske børn – og især at pigerne med indvandrerbaggrund indhenter de etnisk danske piger.

Forskelle i familiebaggrund mellem etnisk danske børn og børn med indvandrerbaggrund viser sig at være den primære forklaring på uddannelseskloftens eksistens. Når der kontrolleres for disse forskelle, er den resterende uddannelseskloft marginal. Der er dog fortsat en væsentlig uddannelseskloft mellem de etnisk danske børn og børnene med indvandrerbaggrund, som er født i udlandet, og størrelsen på denne kloft hænger sammen med barnets alder på indvandringstidspunktet. Senere indvandring betyder en lavere gennemførelsesprocent, og den kritiske indvandringsalder er syv år. Dette betyder, at sandsynligheden for at gennemføre en ungdomsuddannelse reduceres væsentligt for børn, der indvandrer til Danmark, efter de er blevet syv år.

Supplerende materiale

Supplerende materiale til artiklen findes [her](#).

Noter

1. t tager værdien 1, hvis barn i tilhører årgangen, der afslutter 9. klasse i 1990, 2, hvis barnet tilhører årgangen der afslutter 9. klasse i 1991 osv.
2. Det er kun i model (1), at koefficienten på indvandrerbaggrund kan estimeres. Så snart dummyvariable for oprindelseslande indgår i den estimerede model, vil der opstå lineær afhængighed, hvis dummyvariablen for indvandrerbaggrund også indgår. Derfor udelades variablen, og koefficienten for indvandrerbaggrund beregnes i stedet for at estimeres. Se bilag S.1 i det supplerende materiale for beskrivelse af beregningerne.
3. Dette vil ændre fortolkningen af variablene. Koefficienten på IB_i vil nu kun give den marginale effekt af at have indvandrerbaggrund for drenge, og den marginale effekt for piger fås ved at tilføje koefficienten, hvor køn er ganget på ($IB_i \cdot KVIN-DE_i$).
4. Uddannelsesoplysninger om uddannelse gennemført før indvandring stammer fra et surveybaseret register.
5. Definitionen af ikke-vestlige lande følger Danmarks Statistiks definition og inkluderer alle lande undtagen EU, Andorra, Australien, Canada, Island, Liechtenstein, Monaco, New Zealand, Norge, San Marino, Schweiz, USA og Vatikanstaten.

6. Model (2) findes at være den mest korrekt specificerede model, når kontrolvariablene er inkluderet trinvis i estimationerne, og det er undersøgt, om familiebaggrund skal interagere med indvandrerbaggrund eller ej. Derudover viser estimationerne, hvor kontrolvariable inkluderer trinvis, at det især er familiebaggrund, der forklarer forskellene i gennemførelsesprocenterne (se supplerende materiale tabel S.6).

Litteratur

- Adli, Rhonya, Ahmed Louichi og Nadia Tamouh (2010). The sibling size impact on the educational achievement in France. *Education Economics* 18 (3): 331-348.
- Angrist, Joshua David og Jörn-Steffen Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- Bratsberg, Bernt, Oddbjørn Raaum og Knut Røed (2012). Educating children of immigrants: Closing the gap in Norwegian schools. *Nordic Economic Policy Review* 211-251.
- Böhlmark, Anders (2008). Age at immigration and school performance: A siblings analysis using Swedish register data. *Labour Economics* 15: 1366-1387.
- Colding, Bjørg, Leif Husted og Hans Hummelgaard (2009). Educational progression of second-generation immigrants and immigrant children. *Economics of Education Review* 28: 434-443.
- Danmarks Statistik (2014). *Statistisk årbog*. København: Rosendahls-Schultz Grafisk A/S.
- Deding, Mette og Mohammad Azhar Hussain (2005). Educational attainment in Denmark: The role of parents' education and childhood living conditions. *Journal of Applied Social Science Studies* 125 (3): 1008-1019.
- Dickson, Matt, Paul Gregg og Harriet Robinson (2013). Early, late or never? When does parental education impact child outcomes? *IZA Discussion Paper* (7123).
- Dustmann, Christian, Tommaso Frattini og Gianandrea Lanzara (2012). Educational achievement of second-generation immigrants: an international comparison. *Economic Policy* 69 (69): 143-185.
- Hanushek, Eric Alan (1973). *Education and Race: An Analysis of the Educational Production Process* (2. udg.). Lexington: Lexington Books.
- Hanushek, Eric Alan (1979). Conceptual and empirical issues in the estimation of educational production function. *The Journal of Human Resources* 14 (3): 351-388.
- Jakobsen, Vibeke og Nina Smith (2006). The educational attainment of the children of the Danish "guest worker" immigrants. *Danish Journal of Economics* 144 (2): 18-42.
- Machin, Stephen og Sandra McNally (2005). Gender and student achievement in English schools. *Oxford Review of Economic Policy* 21 (3): 357-372.

- Rangvid, Beatrice Schindler (2007). Sources of immigrants' underachievement: Results from PISA-Copenhagen. *Education Economics* 15 (3): 293-326.
- Rangvid, Beatrice Schindler (2010). Source country differences in test score gaps: Evidence from Denmark. *Education Economics* 18 (3): 269-295.
- Riphahn, Regina Therese (2003). Cohort effects in the educational attainment of second generation immigrants in Germany: An analysis of census data. *Journal of Population Economics* 16 (4): 711-737.
- Schneeweis, Nicole (2011). Educational institutions and the integration of migrants. *Journal of Population Economics* 24 (4): 1281-1308.
- Schou, Poul (2006). Immigration, integration and fiscal sustainability. *Journal of Population Economics* 19 (4): 671-689.
- Støren, Liv Anne og Håvard Helland (2010). Ethnicity differences in the completion rates of upper secondary education: How do the effects of gender and social background variables interplay? *European Sociological Review* 26 (5): 585-601.
- Todd, Petra Elisabeth og Kenneth Wolpin (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113 (485): F3-F33.
- Todd, Petra Elisabeth og Kenneth Wolpin (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital* 1 (1): 91-136.
- Van Ours, Jan og Justus Veenman (2006). Age at immigration and educational attainment of young immigrants. *Economic Letters* 90 (3): 310-316.
- Velfærdsaftalen (2006). Aftale mellem regeringen (Venstre og Det Konservative Folkeparti) og Socialdemokraterne, Dansk Folkeparti og Det Radikale Venstre om initiativer til sikring af fremtidens velstand og velfærd og investeringer i fremtiden. <https://www.fm.dk/publikationer/2006/aftale-om-fremtidens-velstand-og-velfaerd>
- Weinberg, Bruce (2001). An incentive model of the effect of parental income on children. *Journal of Political Economy* 109 (2): 266-280.
- Wooldridge, Jeffrey (2003). *Introductory Econometrics: A Modern Approach* (2. udg.). Australia: Thomson/South-Western.
- Wooldridge, Jeffrey (2010). *Econometric Analysis of Cross Section and Panel Data* (2. udg.). Cambridge: the MIT press.