

Segregering i danske børnehaver 2005-2014



Hans Henrik Sievertsen

Segregering i danske børnehaver 2005-2014

© VIVE og forfatteren, 2020

e-ISBN: 978-87-7119-791-4

Modelfoto: Cathrine Kjærø Ulf Ertmann/VIVE

Projekt: 100621

VIVE – Viden til Velfærd

Det Nationale Forsknings- og Analysecenter for Velfærd

Herluf Trolles Gade 11, 1052 København K

www.vive.dk

VIVEs publikationer kan frit citeres med tydelig kildeangivelse.

Forord

Både dansk og international samfundsvidenskabelig forskning har i de seneste årtier givet os langt større viden om betydningen af tidlige indsatser for børns udvikling. Center for Research on Universal Child Policies (CRUNCH), som er støttet af Innovationsfonden, har til formål at udbygge denne viden med fokus på de universelle indsatser i førskolealderen i Danmark. En vigtig del af CRUNCH er at udnytte de eksisterende data til at beskrive den aktuelle tilstand. Denne del fokuserer ikke på effektstudier og ny viden om konsekvenser af specifikke indsatser eller programmer, men derimod på at forbedre vores viden om, hvordan situationen ser ud i de danske universelle programmer til førskolebørn. Denne rapport indgår i dette projekt ved at beskrive omfanget af segregering i danske børnehaver.

Rapportens primære formål er at bidrage til vores viden om segregering i danske institutioner for perioden 2005 til 2014. Da man så vidt vides ikke tidligere har undersøgt segregeringen i danske institutioner, har rapporten det sekundære formål at præsentere metoder til at måle segregeringen samt belyse, om konklusionerne afhænger af metode og definitionsvalg.

Der har ikke været nogen følgegruppe tilknyttet undersøgelsen, men rapporten er blevet kommenteret af en ekstern reviewer. Konklusioner samt eventuelle fejl og mangler i rapporten er forfatterens ansvar.

Carsten Strømbæk Pedersen

Forsknings- og analysechef for VIVE Børn og Uddannelse

2020

Indhold

Sammenfatning	5
1 Indledning	7
1.1 Hvad er segregering?	7
1.2 Hvorfor undersøge segregering i børnehaver?	7
1.3 Viden om segregering i børnehaver	8
2 Hvordan måler vi segregeringen?	10
2.1 Duncans segregationsindeks.....	10
2.2 R ² -metoden	10
2.3 Sammenligning af Duncans segregationsindeks og R ² -metoden	10
2.4 Et randomiseret sammenligningsgrundlag	11
3 Data	12
3.1 Datakilder og dataudvælgelse	12
3.2 Variabeldefinitioner	13
4 Resultater	15
4.1 Segregering i børnehaver 2005-2014.....	15
4.2 Segregering i børnehaver 2005-2014 baseret på R ² -metoden	16
4.3 Har definitionen af forældrenes indkomst, uddannelse og forældrenes oprindelse betydning for omfanget af segregering?	18
4.4 Segregering gennem barndommen.....	19
5 Konklusion	22
Litteratur.....	23
Bilag 1 Ikke-teknisk introduktion til metoderne	24
Bilag 2 Teknisk beskrivelse af metoderne.....	31
Bilag 3 Tabeller om dataudvælgelsen	35
Bilag 4 Betydningen af, hvordan og hvornår økonomiske ressourcer måles	36
Bilag 5 Segregering med hensyn til indkomst	43

Sammenfatning

Baggrund og formål

En opdeling af befolkningen efter deres sociale baggrund betegnes som segregering. Formålet med denne rapport er at undersøge segregeringen i danske børnehaver i perioden 2005 til 2014. Rapporten belyser dermed, om børnehavebørn er blandet på tværs af deres socioøkonomiske baggrund. Med andre ord, så undersøger vi, om fx børn fra højindkomstfamilier går i børnehave med børn fra lavindkomstfamilier, eller om institutionerne er opdelt, så børn fra velhavende familier går i én børnehave og børn fra lavindkomstfamilier går i en anden.

Tidligere studier har undersøgt, hvordan specifikke udsatte grupper af børn er fordelt på tværs af institutioner i Danmark, men der er så vidt vides ikke undersøgelser af den overordnede segregering i institutioner. Rapporten adskiller sig fra undersøgelser af udsatte børnegrupper ved at anvende enklere, men dermed også mere upræcise mål til at identificere den enkelte gruppe. Dette gøres bl.a. for at skabe en høj grad af gennemsigtighed, men også fordi vi her er interesseret i segregeringen per se – altså, hvor blandet er børnesammensætningen i danske børnehaver? I studier af udsatte børn er vi typisk mest interesseret i netop denne gruppe børn og ikke i lige så høj grad de øvrige børn. I denne rapport er vi lige interesseret i alle grupper, idet vi netop kun vil undersøge, hvor meget grupperne blander sig.

Det primære formål med denne rapport er at give et enkelt og transparent mål for segregeringen i danske børnehaver baseret på barnets socioøkonomiske baggrund. Da rapporten er en af de første undersøgelser af segregeringen i børnehaver i Danmark, er et sekundært formål med rapporten at introducere og forklare to internationalt gængse metoder til at måle segregering.

Metode og forbehold

Vi undersøger segregeringen baseret på henholdsvis indkomst, uddannelse og forældrenes oprindelse og ved brug af Duncan Segregationsindeks og R2-metoden. Endelig undersøger vi betydningen af forskelle på tværs af kommuner ved hjælp af en randomiseringsmetode. Vi indleder rapporten med en kort beskrivelse af disse metoder og giver en mere detaljeret beskrivelse af dem i bilagene. Rapportens bilag indeholder en lang række analyser, som har til formål at undersøge, hvor følsomme konklusionerne er med hensyn til, hvordan vi måler barnets baggrund, samt undersøge, om konklusionerne afhænger af, hvilken metode vi bruger.

Formålet med rapporten er udelukkende at dokumentere omfanget af segregering i danske børnehaver. Den belyser ikke årsager til segregeringen og mulige konsekvenser heraf.

Det primære segregeringsmål i rapporten er Duncans segregationsindeks. Dette segregeringsmål kræver, at vi opdeler børnene i to grupper, og indekset angiver så, hvor mange børn fra en af grupperne der skulle skifte institution for at opnå en tilstand, hvor der ikke er nogen segregering. En tilstand uden segregering er defineret ved, at det for alle børnehaver gælder, at børnehavens andel af børn fra den ene gruppe svarer til børnehavens andel af børn fra den anden gruppe. Hvis en børnehave har 10 % af alle børn fra velhavende familier, så skal børnehaven også have 10 % af alle børn fra lavindkomstfamilier for, at der ikke er nogen segregering. Det betyder dermed ikke, at der er lige mange børn fra hver gruppe i børnehaven, idet gruppernes andel af den samlede befolkning kan variere meget.

Resultater

Rapportens resultater viser, at hvis vi opdeler alle børn i to lige store grupper efter den disponible indkomst i familien, så skal et ud af tre børn i højindkomstgruppen (eller lavindkomstgruppen) flytte institution for at opnå en situation uden segregering. Resultaterne for segregeringen med hensyn til forældrenes uddannelse ligner meget resultaterne af segregeringen baseret på de økonomiske ressourcer. Derimod er segregeringen med hensyn til forældrenes oprindelse¹ til sammenligning noget større.

En betydelig del af segregeringen i danske børnehaver skyldes forskelle på tværs af kommuner. Omtrent to tredjedele af segregeringen kan forklares ved forskelle på tværs af kommuner. Hvis vi fordelte børn helt tilfældigt på børnehaver inden for deres bopælskommune, ville vi kunne opnå et segregeringsniveau svarende til, at 20 % af alle børn fra en gruppe skulle skifte institution for at opnå en situation uden segregering.

Segregeringsniveauet har været meget stabilt i perioden 2005 til 2014, og segregeringsniveauet i børnehaverne er på niveau med segregeringsniveauet i grundskolerne.

Konklusioner

Det gennemgående resultat er, at vi finder et segregeringsniveau, der går ud over, hvad vi ville forvente, hvis børn blev tilfældigt fordelt på tværs af børnehaver. Det gælder, uanset om vi måler segregering ved husstandsindkomsten, forældrenes uddannelse, eller forældrenes oprindelse. Hvis vi kunne flytte børn på tværs af hele landet, skulle hvert tredje barn skifte institution for at opnå en situation uden segregering, såfremt segregeringen er målt med hensyn til husstandsindkomsten eller forældrenes uddannelse. En væsentlig del af segregeringen kan dog forklares ved forskelle i børns baggrund på tværs af kommunerne. Selv hvis vi fjernede al segregering i børnehaver inden for hver kommune, så ville vi kun kunne reducere den samlede segregering med en tredjedel. De resterende to tredjedele skyldes forskelle på tværs af kommunegrænser.

Segregeringen med hensyn til forældrenes oprindelse er omtrent dobbelt så høj som segregeringen med hensyn til husstandsindkomsten og forældrenes uddannelse, men igen kan omtrent to tredjedele forklares ved forskelle på tværs af kommuner.

Endelig viser resultaterne, at omfanget af segregering i børnehaver på tværs af alle segregeringsdimensioner har været meget stabilt i perioden 2005 til 2014. Segregeringen i børnehaver svarer nogenlunde til segregeringen i grundskolen. De overordnede konklusioner afhænger ikke af, hvordan vi måler segregeringen, eller hvordan og hvornår vi måler husstandsindkomst og forældrenes uddannelsesbaggrund.

¹ Med hensyn til oprindelse opdeler vi alle børnehavebørn i to grupper: Børn grupperes til at have ikke-vestlig oprindelse, hvis de er efterkommere eller indvandrere fra et ikke-vestligt land. Danmarks Statistik definerer efterkommere som børn, hvor ingen af forældrene er både danske statsborgere og født i Danmark. Vi følger Danmarks Statistik ved opdelingen af lande i vestlige og ikke-vestlige lande. De vestlige lande er ifølge den definition samtlige 28 EU-lande samt Andorra, Island, Liechtenstein, Monaco, Norge, San Marino, Schweiz, Vatikanstaten, Canada, USA, Australien og New Zealand. Da definitionen er baseret på forældrenes fødselsland og statsborgerskab anvendes så vidt muligt begrebet "forældrenes oprindelsesland".

1 Indledning

1.1 Hvad er segregering?

Der har i de seneste år været øget debat om opdelingen af samfundet. Der har fx været diskussioner af opdeltede klasser inden for skoler, hvor nogle klasser har en stor andel af børn med anden etnisk baggrund end dansk, og andre klasser har en meget lille andel. Der har ligeledes været diskussioner på tværs af skoler, hvor en skole har en høj andel af indvandrere og en anden har en lav andel. En sådan adskillelse af befolkningen i grupper kaldes for segregering.

Segregering drejer sig ikke nødvendigvis om opdeling af samfundet med hensyn til oprindelse, men dækker over enhver opdeling af befolkningsgrupper. Det kunne fx være i rig og fattig, højt og lavt uddannet eller piger og drenge. Opdelingen af befolkningen i grupper sker ikke kun i institutioner, som fx i skoler, men også i andre sociale sammenhænge som fritidsaktiviteter og på arbejdspladsen såvel som geografisk i bosætningsmønstre.

Der er flere grunde til, at segregering opstår. For det første kan en del af segregeringen muligvis forklares ved smag og præferencer. Vi vælger at bo i områder, som vi kan lide, og de personer, der kan lide det samme som os, ligner os også. Vi er derfor segregeret i grupper efter, hvad vi godt kan lide. For det andet kan segregering skyldes markedsmekanismer: Huspriser påvirker, hvor vi bosætter os. For det tredje kan institutionelle rammer påvirke segregeringen. Hvilken skole et barn går på afhænger bl.a. af reglerne om allokering af børn til skoler.

Formålet med denne rapport er at undersøge segregeringen i børnehaver i Danmark. Er det sådan, at alle børn fra hjem med højtuddannede forældre går i én børnehave, og alle børn fra hjem med lavtuddannede forældre går i en anden? Eller er børnene blandet på tværs af børnehaver, så børnehavebørn møder børn fra forskellige sociale grupperinger i institutionen? Ovenstående diskussion taget i betragtning, så er det forventeligt, at en del af segregeringen i børnehaver skyldes segregeringen i, hvor vi bor. Et barn, der bor i Esbjerg, går selvfølgelig ikke i børnehave i København. Vi vil senere gennemgå resultater, der både viser det overordnede segregeringsniveau samt det niveau, vi kunne opnå uden at flytte på tværs af kommunegrænser.

1.2 Hvorfor undersøge segregering i børnehaver?

Hvorfor overhovedet undersøge segregering? Formålet med denne rapport er ikke at pege på et problem eller et behov for indgreb. Vi vil her hverken afdække årsager til segregeringen eller konsekvenser af segregeringen. Formålet med rapporten er derimod blot at dokumentere omfanget af segregering i danske børnehaver. Vi vil dog her kort opsummere en række årsager til, hvorfor det er relevant at undersøge omfanget af segregering i danske børnehaver.

For det første viser en række internationale studier, at vi allerede tidligt bliver påvirket af dem, vi omgås med. Børns omgangskreds påvirker både deres kognitive og ikke-kognitive færdigheder (Henry & Rickman, 2007; Neidell & Waldfogel, 2010) såvel som deres præferencer og samfundsdeltagelse (Rao, 2019).

En anden grund til at interessere sig for segregering er, at børns sociale baggrund kan påvirke ressourcebehovet i en institution. Hvis børns baggrund varierer meget fra daginstitution til daginstitution, så kan ressourcebehovet derfor også variere meget fra institution til institution. Der

er flere institutionelle og politiske redskaber til at håndtere denne udfordring, fx sociale normeringer og socioøkonomiske ressourcefordelinger.

Endelig bør det nævnes, at det ikke kan udelukkes, at der også er fordele ved segregering. Det kan muligvis være lettere at organisere dagtilbud til en gruppe børn, der ligner hinanden mere. Det kan føre til færre udgifter, men muligvis også et mere givende tilbud til børnene, fordi børnehaven kan indrette sig specifikt til barnets baggrund.

1.3 Viden om segregering i børnehaver

Som nævnt indledningsvis findes der flere undersøgelser, der har berørt aspekter af segregeringen i Danmark ved at fokusere på udsatte grupper. Disse undersøgelser har typisk anvendt detaljerede beskrivelser af udsatte børn og undersøgt segregeringen som en del af en større undersøgelse om udsatte børn og udfordringer i daginstitutionerne. For at give et indtryk af, hvordan disse undersøgelser adskiller sig fra denne rapport, vil vi her nævne to eksempler.

Det første eksempel er en rapport fra Socialforskningsinstituttet i 1996, hvor Christensen (1996) undersøger udfordringer for daginstitutioner baseret på en spørgeskemaundersøgelse blandt 769 daginstitutioner. I rapporten anvendes udtrykket "truede børn". Truede børn er børn, som enten formodes at være udsat for vanrøgt eller nedværdigende behandling, eller som lever under forhold, der bringer deres sundhed eller udvikling i fare. Rapportens resultater viser bl.a., at 22 % af institutionerne ikke har børn, der betragtes som "truede børn", og 23 % af institutionerne angiver, at mere end 10 % af deres børn betegnes som "truede børn". Rapporten viser desuden, at det pædagogiske arbejde i børnehaver med en høj andel af truede børn afviger fra børnehaver med få eller ingen truede børn.

Det andet eksempel er en rapport udarbejdet i 2004 af AERådet for Ugebrevet A4 med støtte fra BUPL og PMF, hvor Niels Glavind dokumenterer omfanget af social segregering i daginstitutioner i Danmark (Glavind, 2004). I rapporten defineres børn med "svag hjemmebaggrund" som børn, hvor mindst to af følgende kriterier var opfyldt: 1) at barnet ikke boede hos begge forældre, 2) at ingen af forældrene havde en erhvervsuddannelse, og 3) at en af forældrenes primært levede af kontanthjælp. Baseret på denne definition var halvdelen af børn i Danmark med svag hjemmebaggrund i 2002 koncentreret i 20 % af institutionerne. I en nyere rapport (Glavind & Pade, 2015) anvendes igen begrebet "svag hjemmebaggrund", og det konkluderes, at de 25 % mest belastede daginstitutioner dækker over 50 % af børnene med "svag hjemmebaggrund".

Hvor Christensen (1996) er et eksempel på en undersøgelse, hvor man har bedt fagprofessionelle identificere "truede børn", så er Glavind (2004) et eksempel på et studie, der anvender en finkornet definition af børn med "svag hjemmebaggrund" baseret på registerdata. Disse metoder står i stor kontrast til dette studie, hvor børnegrupper dannes baseret på, om barnet kommer fra et hjem i den laveste eller højeste halvdel baseret på husstandsindkomsten eller forældrenes uddannelse, eller om forældrenes oprindelse er ikke-vestlig (jf. Danmarks Statistiks definition af vestlige og ikke-vestlige lande). Grupperne i denne rapport er derfor langt større, og vi tager heller ikke stilling til, om børn fra den ene eller den anden gruppe udgør en risikogruppe. Børn fra familier med en indkomst under medianen er ikke nødvendigvis i højere grad en risikogruppe end børn fra en familie med en ækvivaleret disponibel husstandsindkomst over medianen. Vi ser derfor blot på, om børn er blandet på tværs af social baggrund uden at antage, at den ene eller anden baggrund udgør nogen risiko. Fordelen ved de mere simple metoder og mål i denne rapport er, at det gør det muligt let at sammenligne segregeringen på tværs af årgange og kalenderår. Ulempen ved denne simple tilgang er, at vi opdeler børn i

midten af indkomstfordelingen (og fordelingen af husstandsindkomsten), hvilket betyder, at børnehaver med familier lige under og lige over midten bidrager til segregeringen, selvom børnene kommer fra forhold, der er meget ens. Vi vil i den senere analyse derfor også bruge en alternativ opdeling, hvor vi anvender et relativt fattigdomsmål.

2 Hvordan måler vi segregeringen?

I dette kapitel vil vi kort gennemgå, hvordan vi måler segregeringen på tværs af børnehaver i Danmark. Vi giver her kun en kort introduktion til metoderne. En mere detaljeret og ikke-teknisk beskrivelse af metoderne findes i bilag 1, og en mere teknisk beskrivelse af metoderne findes i bilag 2.

2.1 Duncans segregationsindeks

Det primære mål for segregeringen er baseret på Duncans segregationsindeks (Duncan & Duncan, 1955). Dette indeks måler segregeringen med hensyn til karakteristika, som opdeler befolkningen i to grupper. Duncans segregationsindeks angiver, hvor stor en andel af børnene fra en af grupperne der skal flytte institution, for at vi opnår en situation uden segregering. Værdien vil altid være mellem 0 og 1 (begge tal inkluderet). En større værdi angiver en større grad af segregering.

En situation uden segregering er ifølge Duncans segregationsindeks den, hvor det i alle børnehaver er gældende, at andelen af børnene fra en af grupperne er lig andelen af børn fra den anden gruppe. Det vil med andre ord sige, at en tilstand uden segregering er en tilstand, hvor en børnehaver, der har 1 % af alle børn i befolkningen, også har 1 % af alle rige børn og 1 % af alle fattige børn.

2.2 R^2 -metoden

Det sekundære segregationsmål er baseret på R^2 (R-i-anden) -metoden. R^2 kendes normalt som mål for "Goodness of fit" for en estimeret model. R^2 angiver i den sammenhæng, løst sagt, hvor meget af variationen i en variabel vi kan forklare med den anden variabel. Vi kan her overføre dette til børnehaver: Hvis vi kan "forklare" en stor del af variationen i forældrenes baggrund baseret på, hvilken børnehaver barnet går i, så vil børnehaver være meget segregerede.

Med andre ord, hvis vi kan gætte barnets sociale baggrund baseret på, hvilken daginstitution det går i, så må børnenes sociale baggrund variere på tværs af børnehaver. R^2 -værdien angiver, hvor godt vi kan gætte barnets baggrund baseret på viden om, hvilken børnehaver barnet går i. Værdien vil altid være mellem 0 og 1 (begge tal inkluderet). En større værdi angiver en større grad af segregering.

En situation uden segregering er ifølge R^2 -metoden en situation, hvor den gennemsnitlige sociale baggrund er ens på tværs af alle børnehaver. Det kunne fx være, at den gennemsnitlige ækvivalerede disponible indkomst er ens på tværs af børnehaver.

2.3 Sammenligning af Duncans segregationsindeks og R^2 -metoden

De to metoder bygger på ret forskellige tilgange, og som vi viser i bilag 2, så kan der være situationer, hvor den ene metode viser, at der ikke er nogen segregering, og den anden metode viser, at der er et væsentligt niveau af segregering. Det er netop derfor, at vi har valgt at undersøge segregationen baseret på begge disse mål, da de kan give ret forskellige konklusioner.

Duncans segregationsindeks er internationalt set et langt mere brugt segregationsmål sammenlignet med R^2 -metoden. Det skyldes bl.a., at Duncans segregationsindeks har en direkte fortolkning. Ulempen ved denne metode er dog, at den ser bort fra al variation i barnets baggrund som går ud over, hvad der kan klassificeres i to grupper.

2.4 Et randomiseret sammenligningsgrundlag

For at opnå en situation uden segregering skulle vi muligvis flytte alle børn på tværs af landet. En situation uden segregation er derfor ikke nødvendigvis et retvisende sammenligningsgrundlag. Vi skaber derfor i stedet et sammenligningsgrundlag, som tager højde for, i hvilken kommune barnet bor samt institutionernes størrelse og alderssammensætning. Dette gør vi ved at spørge: Hvordan ville segregeringen være, hvis vi fordelte børn tilfældigt på tværs af institutioner inden for kommunen uden at ændre på institutionsstørrelse, type og alderssammensætning?

For at skabe et retvisende sammenligningsgrundlag skaber vi derfor et såkaldt "randomiseret sammenligningsgrundlag" ved følgende metode:

1. Vi skaber et "kunstigt datasæt", hvor alle børn allokeres til en tilfældig børnehave.
2. I det kunstige datasæt tager vi højde for, i hvilken kommune barnet bor, børnehavernes størrelse og deres alderssammensætning. Dette gør vi ved at allokere børnene til en tilfældig børnehave inden for den kommune, hvor de bor. Derudover skal der i den "tilfældige" børnehave være plads til et barn på det alderstrin, som barnet har.
3. Vi udregner dernæst segregeringsomfanget i dette kunstige datasæt.
4. En enkelt tilfældig fordeling af børn kan naturligvis også give den faktiske fordeling eller en situation med højere segregering. Vi skaber derfor 50 "kunstige datasæt" med tilfældige fordelinger af børnehaver og udregner den gennemsnitlige segregering på tværs af disse 50 tilfældige fordelinger.²
5. Det gennemsnitlige segregeringsomfang på tværs af de 50 "kunstige datasæt" udgør vores "randomiserede sammenligningsgrundlag".
6. Til sidst sammenligner vi segregeringsomfanget i det rigtige datasæt med det gennemsnitlige segregeringsomfang på tværs af de 50 kunstige datasæt.

² Vi har for enkelte specifikationer undersøgt, hvor følsomme konklusionerne er med hensyn til antal tilfældige allokeringer. Vi har afprøvet op til 500 tilfældige allokeringer, uden at det har påvirket konklusionerne i væsentlig grad.

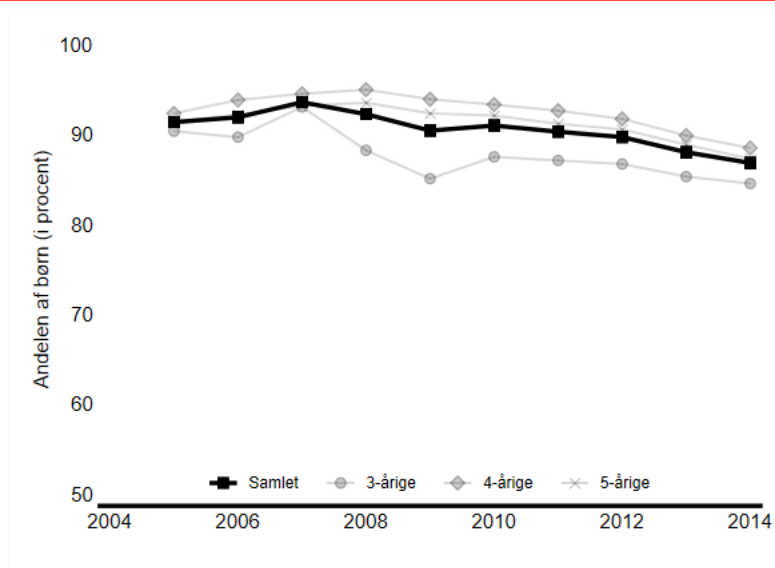
3 Data

3.1 Datakilder og dataudvælgelse

Udgangspunktet for analysen er Danmarks Statistiks daginstitutionsregister (DAGI). Registeret dækker over perioden 1995 til og med 2014. For perioden 1995 til og med 2004 mangler der en række oplysninger, og denne periode udelades derfor af analysen, og vi undersøger derfor segregeringen for årene 2005 til 2014. Registeret er et årligt register, som angiver, hvilken institution barnet er indskrevet i på opførelsestidspunktet i september eller oktober måned, og hvilken type ydelse det drejer sig om (fx om det er en børnehave eller en aldersintegreret institution). Vi kobler alle børn, som er dækket af daginstitutionsregisteret, med befolkningsregisteret fra Danmarks Statistik ved brug af et anonymiseret cpr-nummer.

Daginstitutionsregistret fra Danmarks Statistik – DAGI – er baseret på data fra kommunernes it-system. Registerets dækningsgrad kan derfor variere på tværs af kommuner og institutions-type. Hvis der er en systematisk sammenhæng mellem dækningsgraden og befolkningssammensætningen i kommunen og kommunens institutioner, så kan det påvirke resultaterne i analysen.

Figur 3.1 Andelen af børn, der er registreret i en børnehave eller en aldersintegreret institution i DAGI-registeret



Anm.: Figuren viser andelen af børn i befolkningen, som er indskrevet i en børnehave eller en aldersintegreret institution ifølge DAGI-registeret.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Figur 3.1 viser, hvor stor en andel af børnene i befolkningen, som ifølge DAGI-registeret enten er registreret i en børnehave eller i en aldersintegreret institution. Ifølge vores egne beregninger ved brug af registeret er samlet set omtrent 90 % af alle børn indskrevet i en institution. Andelen har været svagt faldende, og i 2014 var 86 % af de 3-5-årige indskrevet. Denne andel er i tråd med tal fra andre kilder. For eksempel rapporterer Glavind og Pade (2018) en andel på 86 % i 2016 baseret på direkte forespørgsel hos kommunerne. Vi har derfor ingen grund til at forvente væsentlige målefejl i vores data, om end det bør bemærkes, at der blandt de 10-14 %, som ifølge vores data ikke er indskrevet i en institution, kan være tale om målefejl.

Fra befolkningsregisteret får vi yderligere oplysninger om barnets fødselsdag, forældrenes oprindelse, bopælskommune, husstand og forældre. Endelig kobler vi husstands- og forældreoplysninger med data fra indkomst- og uddannelsesregistre for at få informationer om husstands- og forældreindkomst såvel som forældrenes gennemførte uddannelse.

I kalenderårene 2005 til 2014 medtager vi alle børn, som fylder henholdsvis 3, 4 og 5 år i det kalenderår. Det vil fx sige, at vi for 2007 medtager børn, der er født i årene 2002 (de fylder 5 år i 2007), 2003 (de fylder 4 år i 2007) og 2004 (de fylder 3 år i 2007). I praksis betyder det, at børnene er mellem 2 år og 9 måneder og 5 år og 9 måneder på opførelsestidspunktet. Derudover medtager vi kun børn, som er indskrevet i enten børnehave eller i en aldersintegreret institution. Ved vores udvælgelse vil der være børn, som i praksis endnu ikke er startet i børnehave på opførelsestidspunktet, og i de medtagne årgange er der også børn, som er begyndt i grundskolen på opførelsestidspunktet. Dette kan give en målefejl, men da vi, som beskrevet ovenfor, tager højder for alderssammensætningen i sammenligningsscenerierne, vil målefejlen ikke påvirke den overordnede konklusion.

Bilagstabel 1.1 viser antallet af børn og institutioner der er medtaget i analyserne. I beregningen af indkomstgrænserne har vi medtaget hele børnekohorten, dvs. også de børn, der ikke er indskrevet i en institution.

3.2 Variabeldefinitioner

I rapporten ser vi på segregationen i børnehaver med hensyn til indkomst, uddannelse og forældrenes oprindelse. I det følgende beskriver vi, hvordan disse variable er målt.

Indkomst

Det primære mål for indkomst er den ækvivalerede samlede disponible indkomst i husstanden i det pågældende kalenderår. Den disponible indkomst er indkomsten efter skatter og overførsler. For at gøre ressourcerne sammenlignelige på tværs af grupper justerer vi den samlede indkomst for husstandsstørrelse ved at dividere den samlede indkomst med kvadratroden af antal medlemmer i husstanden. Husstandsindkomsten justeret for husstandsstørrelse kaldes den ækvivalerede husstandsindkomst. Dette er en standard ækvivaleringsmetode, som bl.a. også benyttes af OECD i deres rapporter. Med det pågældende kalenderår mener vi det år, for hvilket segregationen udregnes. Undersøger vi fx segregeringen i 2005, så vil vi se på den ækvivalerede disponible husstandsindkomst i 2005.

I hvert kalenderår opdeler vi alle børn i hver fødselskohorte i to lige store grupper efter den ækvivalerede husstandsindkomst og danner dermed en høj- og lavindkomstgruppe. Igen bør det dog bemærkes, at langt fra alle i højindkomstgruppen har en høj indkomst og alle i lavindkomstgruppen en lav indkomst; der vil i begge grupper være en del husstande med indkomster omkring medianen. I 2014 skulle en husstand bestående af to børn og to voksne altså have en samlet disponibel indkomst på omkring 500.000 kr. eller derover for at indgå i højindkomstgruppen. Det skal desuden nævnes, at der inden for et år kan være overlap i beløbsgrænserne, idet grænserne er opgjort separat for hver fødselskohorte. Således vil der for kalenderåret 2014 være en beløbsgrænse for alle børn født i 2009, en grænse for alle børn født i 2010 og en grænse for alle børn født i 2011. De præcise indkomstgrænser for hvert år fremgår af bilagstabel 3.1.

Ud over den samlede husstandsindkomst i det samme kalenderår undersøger vi også segregationen med hensyn til forældrenes gennemsnitlige indkomst i året, før barnet blev født, og i det kalenderår barnet fyldte 5 år. Bemærk, at vi i de analyser ikke ser på husstandsindkomst,

men på den gennemsnitlige indkomst på tværs af de biologiske forældre. Endelig så opgør vi også segregering med hensyn til økonomiske ressourcer baseret på, om den ækvivalerede disponible husstandsindkomst er under 60 % af medianindkomsten, som er en gængs definition for relativ fattigdom. Formålet med disse analyser er at undersøge, hvor følsomme konklusionerne er over for definitionen af økonomiske ressourcer.

Vi udregner medianen baseret på hele fødselskohorten. Det vil sige, at opdelingen af børnene også inddrager børn, som ikke går i børnehave.

Uddannelse

Det primære mål for forældrenes uddannelse er det gennemsnitlige antal års uddannelse, som de biologiske forældre har gennemført, målt i oktober i det kalenderår, som vi undersøger. Antal års uddannelse er opgjort som det mindste antal år, som den højst gennemførte uddannelse kunne gennemføres på.

Ligesom med indkomst opdeler vi i hvert kalenderår og for hver fødselskohorte børnene i to lige store grupper efter forældrenes uddannelse og betegner de to grupper som børn af henholdsvis lavt- og højtuddannede forældre. Grænserne for at indgå i højtuddannelsesgruppen er forholdsvis konstante over tid på omkring 14 år i hele perioden. De præcise grænser for hvert år kan aflæses i bilagstabel 3.1.

Ligesom ved de økonomiske ressourcer udregnes medianen baseret på hele fødselskohorten. Det vil sige, at opdelingen af børnene også inddrager børn, som ikke går i børnehave.

Det er værd at bemærke, at det ikke nødvendigvis er dårligere eller bedre at være i den ene eller anden gruppe. Formålet med rapporten er blot at opføre, om børnehaver er systematisk opdelt efter forældrebaggrund.

Oprindelse

Det primære mål for oprindelse er baseret på forældrenes statsborgerskab og fødselsland. Børn grupperes som at have en ikke-vestlig oprindelse, hvis de er efterkommere eller indvandrere fra et ikke-vestligt land. Danmarks Statistik definerer efterkommere som børn, hvor ingen af forældrene er både danske statsborgere og født i Danmark. Vi følger Danmarks Statistik ved opdelingen af lande i vestlige og ikke-vestlige lande. De vestlige lande er ifølge definitionen samtlige 28 EU-lande samt Andorra, Island, Liechtenstein, Monaco, Norge, San Marino, Schweiz, Vatikanstaten, Canada, USA, Australien og New Zealand.

4 Resultater

4.1 Segregering i børnehaver 2005-2014

Figur 4.1 viser udviklingen i segregeringen i danske børnehaver i perioden 2005 til 2014 målt på Duncans segregationsindeks og baseret på henholdsvis forældrenes indkomst, deres uddannelsesniveau og deres forældres oprindelse.

Det fremgår af figur 4.1, at segregationsniveauet har været meget konstant igennem hele perioden. Vi ser ingen tydelig tendens til stigende eller faldende segregering i danske børnehaver. Derudover ser vi for både segregeringen med hensyn til forældrenes indkomst (figur 4.1.a) og deres uddannelse (figur 4.1.c) et segregeringsniveau på omtrent 0,3. Det betyder, at vi skulle flytte omtrent hvert tredje barn (30 %) fra en af grupperne for at opnå en situation uden segregering. Med andre ord, så skulle hvert tredje barn i lavindkomstgruppen flytte institution for at opnå en situation uden segregering. Vi ser dog også, at selv hvis børn var tilfældigt fordelt på tværs af institutioner (de grå kurver) inden for kommuner, så ville segregeringsniveauet være 0,2. Forskellen mellem det scenarie, hvor børn er tilfældigt fordelt på tværs af institutioner, og det faktiske segregeringsniveau er også forholdsvis konstant (figur 4.1.b og 4.2.d). Vi kan altså konkludere, at omtrent to tredjedele af den observerede segregation i børnehaver med hensyn til forældrenes indkomst og uddannelse kan forklares ved forskelle på tværs af kommuner, fordi selv hvis børn blev tilfældigt fordelt inden for en kommune, ville vi kun kunne reducere segregeringen med en tredjedel.

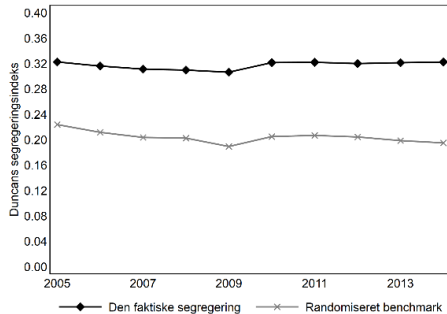
Undersøger vi segregeringen med hensyn til forældrenes oprindelse (figur 4.1.e), så ser vi et noget andet niveau sammenlignet med segregeringen baseret på baggrund af forældreindkomst og uddannelse. Her får vi en indekssværdi på omkring 0,6, svarende til at 60 % af alle børn i en gruppe skulle flytte institution for at opnå en situation uden segregering. Igen ser vi dog, at omkring to tredjedele af denne segregering skyldes forskelle på tværs af kommuner, idet den grå kurve i Figur 4.1.e viser et segregationsniveau på omkring 0,4 i det randomiserede sammenligningsgrundlag. Forskellen mellem det faktiske og det randomiserede segregeringsniveau er igen meget stabilt, som det fremgår af figur 4.1.f.

For at belyse usikkerheden i de beregnede segregeringsmål har vi udregnet 95 % konfidensintervallet, som er vist ved de stiplede linjer i figur 4.1. Vi ser, at konfidensintervallet ligger tæt omkring linjen, hvilket indikerer, at der ikke blot er tale om tilfældig variation. Gennemgående vil vi kun vise usikkerheden ved hjælp af konfidensintervallet, når vi sammenligner det faktiske segregeringsniveau med det randomiserede sammenligningsgrundlag.

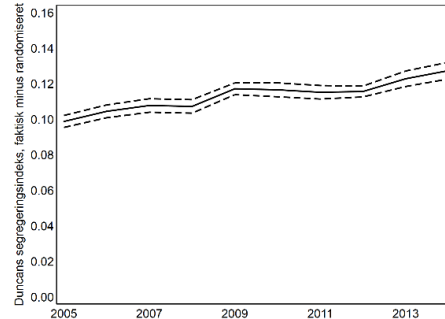
Vi kan altså ud fra figur 4.1 konkludere, at segregationsniveauet har været meget stabilt, at det ligger på omkring 0,3 for uddannelse og indkomst og på 0,6 for oprindelse, og at cirka to tredjedele af segregeringen kan forklares ved forskelle på tværs af kommuner. Der er dog flere åbne spørgsmål. For det første, er konklusionerne følsomme over for, hvordan vi måler segregeringen? For det andet, ville vi opnå samme konklusion, hvis vi opgjorde indkomsten og uddannelsen på en anden måde? For det tredje, er et Duncans segregationsindeks på 0,3 højt eller lavt? Vi vil undersøge disse spørgsmål i de følgende afsnit.

Figur 4.1 Segregeringen i børnehaver i Danmark målt ved Duncans segregationsindeks for årene 2005-2014

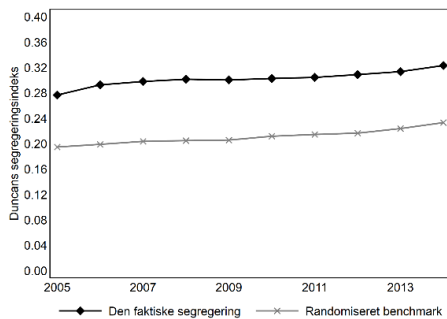
(a) Indkomst, det faktiske og det randomiserede segregeringsniveau



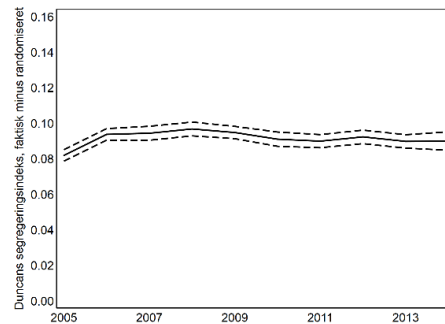
(b) Indkomstsegregering inden for kommunerne



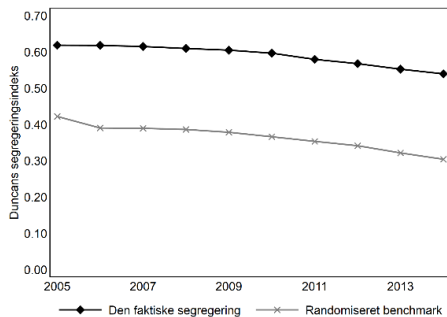
(c) Uddannelse, det faktisk og det randomiserede segregeringsniveau



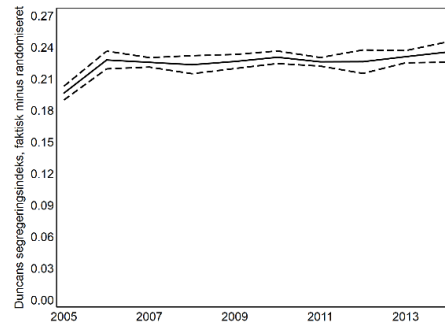
(d) Uddannelsessegregering inden for kommunerne



(e) Oprindelse, det faktiske og det randomiserede segregeringsniveau



(f) Oprindelsessegregering inden for kommunerne



Note.: Figurerne viser Duncans segregeringsindeks for årene 2005-2014 for kohorterne 2002-2011 baseret på den ækvivalerede husstandsindkomst i det pågældende kalenderår, de biologiske forældres gennemsnitlige uddannelsesniveau i det pågældende kalenderår og oprindelse.

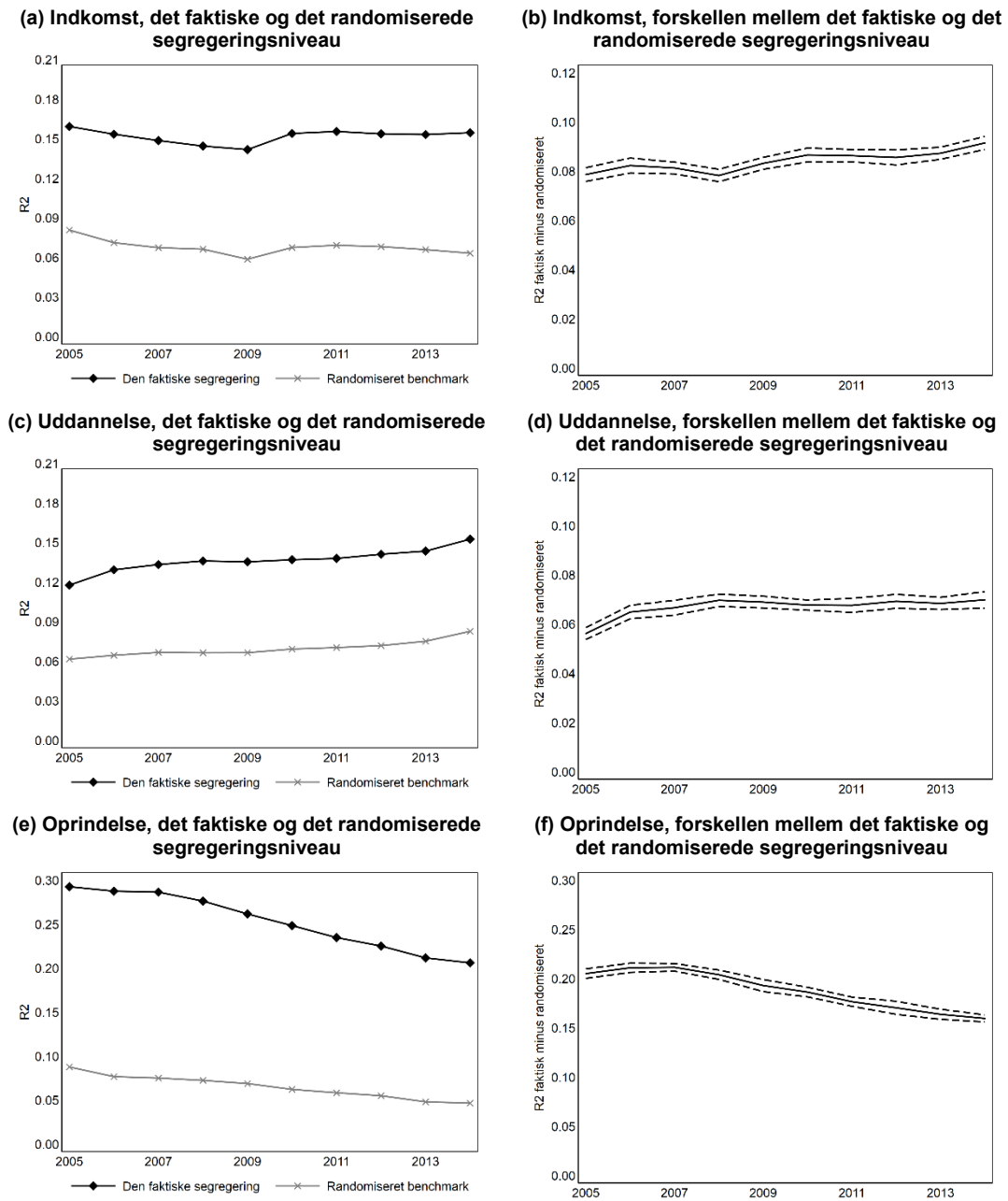
Det randomiserede sammenligningsgrundlag viser det gennemsnitlige segregeringsniveau på tværs af 50 tilfældige allokeringer af børn til børnehaver, hvor vi fastholder barnets alder, bopælskommune, institutionstyper og institutionsstørrelse. De stiplede kurver viser 95 % konfidensintervallet, som er baseret på 200 bootstrap-iterationer.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

4.2 Segregering i børnehaver 2005-2014 baseret på R^2 -metoden

Vi vil først undersøge, om konklusionerne ændrer sig, hvis vi anvender et alternativt segregeringsmål. Figur 4.2 viser segregeringen i børnehaver baseret på R^2 -metoden. Sammenlignet med resultaterne i figur 4.1 ser vi igen et meget stabilt niveau for segregeringen baseret på forældrenes indkomst og uddannelse hen over hele perioden. For segregeringen baseret på forældrenes oprindelse ser vi nu lidt stærkere tegn på en aftagende segregering.

Figur 4.2 Segregeringen i børnehaver i Danmark målt ved R²-metoden



Note.: Figurerne viser segregeringen målt ved R² for årene 2005-2014 for kohorterne 2002-2011 baseret på den ækvivalerede husstandsindkomst i det pågældende kalenderår, de biologiske forældres gennemsnitlige uddannelsesniveau i det pågældende kalenderår og oprindelse.

Det randomiserede sammenligningsgrundlag viser det gennemsnitlige segregeringsniveau på tværs af 50 tilfældige allokeringer af børn til børnehaver, hvor vi tager højde for barnets alder, bopælskommune, institutionstyper og institutionsstørrelse. De stiplede kurver viser 95 % konfidensintervallet, som er baseret på 200 bootstrap-iterationer.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Ligesom segregeringen baseret på Duncans segregationsindeks, så viser resultaterne baseret på R²-metoden en segregering, der går ud over hvad vi ville forvente i et scenarie, hvor børn blev fordelt tilfældigt på børnehaver inden for kommunen. Dette ses ved at sammenligne de grå og sorte kurver i figurerne 4.2.a, 4.2.c og 4.2.e. Den andel af segregationen, der kan forklares ved forskelle på tværs af kommuner, er nu lidt lavere og udgør omtrent 50 til 60 %. Denne forskel er dog igen meget konstant, som det fremgår af figurer 4.2.b, 4.2.d og 4.2.f.

Det er værd at bemærke, at vi med R^2 -metoden ikke opdeler børn i grupper baseret på forældrenes indkomst eller uddannelse. I stedet undersøger vi blot, hvor meget af variationen i forældrenes indkomst og uddannelse vi kan "forudsige" baseret på forskelle på tværs af børnehaver. Vi kan altså dermed inddrage en større del af variationen i indkomst og uddannelse.

4.3 Har definitionen af forældrenes indkomst, uddannelse og forældrenes oprindelse betydning for omfanget af segregering?

I figur 4.1 og 4.2 har vi vist resultater, hvor forældrenes indkomst og uddannelse er målt i det samme kalenderår. Vi vil her kort gennemgå, hvor følsomme disse konklusioner er med hensyn til, hvordan variablene er defineret og målt.

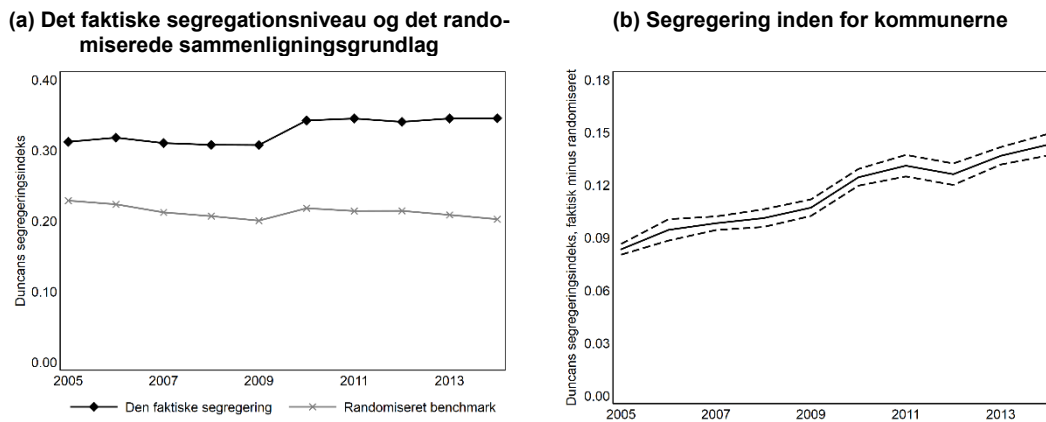
I bilag 4 præsenterer vi resultater, hvor vi anvender alternative mål for de økonomiske ressourcer. For det første undersøger vi, om konklusionerne afhænger af, om vi måler ressourcerne i det samme kalenderår, i kalenderåret før barnet blev født, eller i kalenderåret, hvor barnet fyldte 5 år. For det andet ser vi på, om resultaterne er følsomme over for, om vi måler husstandsindkomsten eller indkomsten på tværs af de biologiske forældre. For det tredje undersøger vi, om konklusionen er følsom over for, hvordan vi har opdelt de to grupper ved at se på, om barnet er fra en familie med en husstandsindkomst under 60 % af medianindkomsten (et gængs mål for relativ fattigdom). Overordnet set er konklusionerne ikke afhængige af definitionerne. Vi ser dog en lidt højere grad af segregering, når de økonomiske ressourcer måles senere i barnets liv.

Anvender vi fattigdomsdefinitionen til at opdele børnegrupperne, så ser vi, at det overordnede segregeringsniveau er nogenlunde konstant som vist i figur 4.3.a. Derimod har forskellen mellem det faktiske segregeringsniveau og de tilfældige sammenligningsgrundlag været støt stigende gennem perioden – fra mindre end 9 % til omkring 14 % – som vist i figur 4.3.b. Dette tyder altså på, at segregeringen inden for kommunerne er steget noget, baseret på en opdeling af børn med hensyn til relativ fattigdom.

I bilag 4 præsenterer vi desuden resultater, hvor vi måler forældrenes indkomst henholdsvis i kalenderåret, før barnet blev født, og i kalenderåret, hvor barnet fyldte 5 år. Ligesom ved de økonomiske ressourcer ser vi, at også for uddannelsesniveaet bliver de overordnede konklusioner ikke påvirket af, hvornår vi måler forældrenes uddannelsesniveau.

Endelig viser bilaget segregeringen baseret på forældrenes oprindelse, hvor vi opdeler børn i to grupper efter børn med dansk oprindelse og børn med ikke-dansk oprindelse. Det vil med andre ord sige, at den ene gruppe består af børn, hvor mindst en af forældrene er født i Danmark og har dansk statsborgerskab, og den anden gruppe består af alle andre børn, uanset hvilket land barnet eller barnets forældre kommer fra. Segregeringen er en smule lavere ved denne definition af oprindelsesland, men overordnet set er konklusionen uændret.

Figur 4.3 Segregeringen i børnehaver i Danmark målt ved Duncans segregationsindeks for årene 2005-2014 baseret på, om barnet kommer fra en husstand med en disponibel indkomst under 60 % af medianindkomsten



Anm.: Figurene viser Duncans segregeringsindeks for årene 2005-2014 for kohorterne 2002-2011 baseret på den ækvivalerede disponible husstandsindkomst i det pågældende kalenderår. De to grupper består af alle børn fra husstande med en ækvivaleret disponibel husstandsindkomst på under 60 % af medianindkomsten.

Det randomiserede sammenligningsgrundlag viser det gennemsnitlige segregeringsniveau på tværs af 50 tilfældige allokeringer af børn til børnehaver, hvor vi tager højde for barnets alder, bopælskommune, institutionstyper og institutionsstørrelse. De stiplede kurver viser 95 % konfidensintervallet, som er baseret på 200 bootstrap-iterationer.

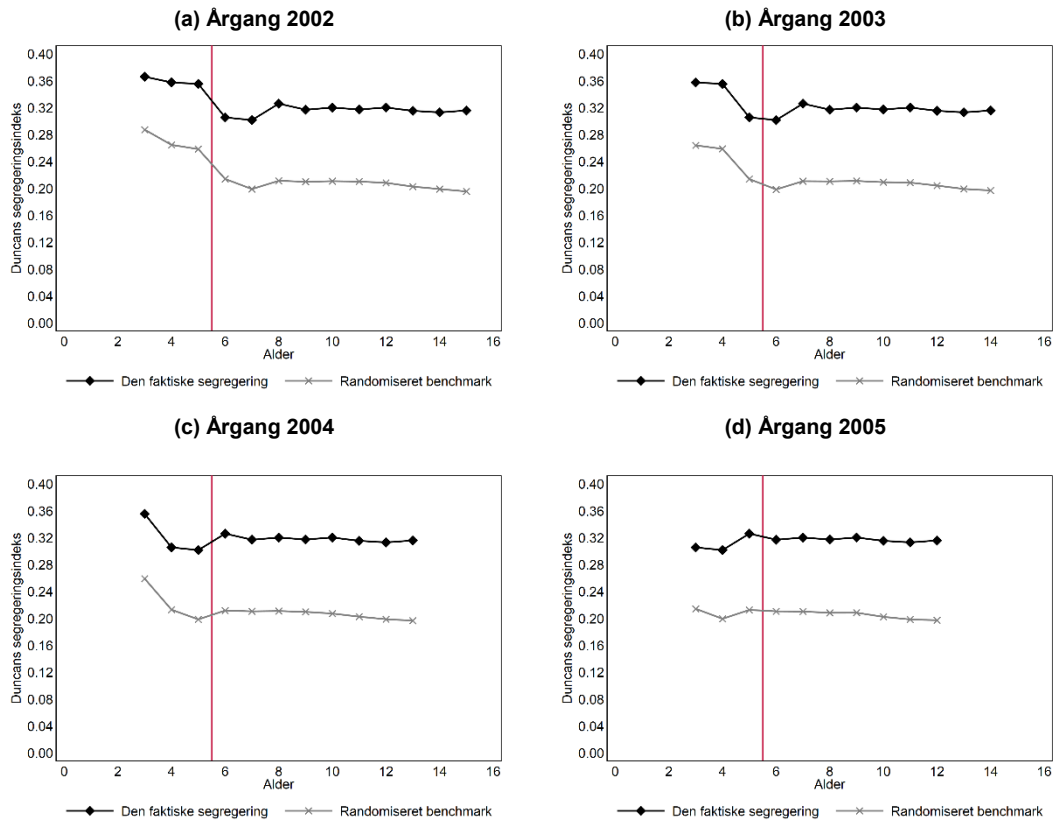
Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

4.4 Segregering gennem barndommen

Vi har vist, at segregeringen i de danske børnehaver har været forholdsvis konstant hen over perioden 2005-2014, og at segregeringen går ud over, hvad vi ville forvente baseret på forskelle på tværs af kommuner. Men er et Duncans segregationsindeks på 0,3 højt eller lavt? Det er svært at sammenligne segregationsmål på tværs af lande, idet værdien vil afhænge af, hvordan grupperne er defineret. Vi kan dog sammenligne segregeringen på tværs af institutionstyper og undersøge segregeringen gennem barndommen og se på segregeringen i børnehaver sammenlignet med segregeringen i grundskolen.

Figur 4.4 viser segregeringen med hensyn til økonomiske ressourcer gennem barndommen for fire børnekoherter. I disse figurer tager vi kun udgangspunkt i en fødselskohorte og undersøger segregeringen for hvert alderstrin. De første tre punkter i graferne svarer nogenlunde til resultaterne vist i figur 4.1 – med den lille afvigelse, at vi her kun sammenligner børn født i samme år.

Figur 4.4 Segregeringen gennem barndommen målt ved Duncans segregationsindeks, baseret på den ækvivalerede disponible husstandsindkomst



Anm.: Figurene viser Duncans segregeringsindeks for årene 2005-2017 for kohorterne 2002-2005 baseret på den ækvivalerede disponible husstandsindkomst i det kalenderår, hvor barnet fylder det antal år angivet på den horisontale akse. De to grupper består af alle børn fra husstande med en ækvivaleret disponibel husstandsindkomst henholdsvis under og over medianen. Segregeringen udregnes i forhold til alle andre børn i samme kohorte.

Det randomiserede sammenligningsgrundlag viser det gennemsnitlige segregeringsniveau på tværs af 50 tilfældige allokeringer af børn til børnehaver og grundskoler, hvor vi tager højde for barnets alder, bopælskommune, institutionstyper (for børnehaver) og institutionsstørrelse.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Den sorte kurve i figur 4.4.a antyder, at segregeringen er højest, når barnet er 3-4 år, dvs. i børnehavealderen, og så falder noget, når barnet starter i grundskolen i det år, hvor barnet fylder 6 år. Derefter er segregeringen forholdsvis konstant. Sammenligner vi denne udvikling med de tre øvrige kohorter i figur 4.4.b-4.4.d, ser vi dog, at denne udvikling sandsynligvis skyldes års-effekter. Alderstrinnet, hvor segregeringen falder, falder med et år for hvert år senere, at børnene er født, hvilket tyder på, at faldet skyldes en begivenhed i 2008. Faldet i 2008 kan skyldes strukturelle ændringer, men det kan også være økonomiske konjunkturer og tilfældigheder. Finanskrisen kan påvirke, hvilke børn der bliver placeret i henholdsvis høj- og lavindkomstgrupperne. Samlet set tyder resultaterne i dette afsnit på, at segregeringen i børnehaver svarer nogenlunde til niveauet af segregeringen i grundskolen. Det er muligvis lidt højere. Dette bliver især tydeligt, når vi betragter kohorten født i 2005, som er vist i figur 4.3.d.

Vi genfinder konklusionen for segregeringen gennem barndommen baseret på økonomiske ressourcer, når vi undersøger segregeringen gennem barndommen baseret på forældrenes uddannelse og forældrenes oprindelsesland, som er vist i bilag 5. Samlet set er segregeringen

i børnehaver nogenlunde på niveau med i grundskolen. Endelig bør det bemærkes, at grundskoler generelt er noget større sammenlignet med børnehaver, hvilket burde trække imod mindre segregering.

At segregeringsomfanget i børnehaver svarer til segregeringen i grundskolen, giver i sig selv ikke et direkte svar på, om segregeringen i børnehaver er høj eller lav. På grundskoleniveauet har der i flere år været offentlig debat og analyse af segregeringsomfanget, især med hensyn forholdet mellem folkeskolen og privatskolerne (se Vasiljeva et al, 2017). Den offentlige debat om segregering i daginstitutioner har derimod været mere begrænset, selvom omfanget – ifølge vores resultater – er meget tilsvarende. Endelig kan det nævnes, at Massey (1990) i en videnskabelig artikel om segregering med hensyn til oprindelse i USA angiver et segregeringsomfang på under 0,33 på Duncans segregationsindeks som lavt. Det vil altså sige, at segregeringsomfanget i de danske børnehaver er på grænsen til lav. Det bør dog bemærkes, at dette kun er en tommelfingerregel.

5 Konklusion

Ved brug af registerdata fra Danmarks Statistik har vi undersøgt graden af segregation baseret på forældrenes indkomst, uddannelse og forældrenes oprindelse.

Hvis vi opdeler børnene i to grupper – høj og lav – finder vi for både forældrenes uddannelse og indkomst, at omkring hvert tredje barn i en af grupperne skulle flytte daginstitution for at opnå en helt ligelig fordeling på tværs af institutioner. En væsentlig del af denne ulige fordeling kan dog forklares ved forskelle på tværs af kommuner. Vores analyser viser, at omtrent to tredjedele af den samlede ulighed skyldes forskelle på tværs af kommuner og institutionelle rammer (institutionsstørrelser mv.). Med hensyn til oprindelse viser resultaterne, at segregeringen er dobbelt så stor, men igen kan omtrent to tredjedele af segregeringen forklares ved forskelle på tværs af kommuner og gennem institutionelle faktorer.

Overordnet set har graden af segregeringen været forholdsvis konstant i perioden 2005-2014. En lille undtagelse er segregeringen med hensyn til, om barnet kommer fra en husstand med en disponibel indkomst på under 60 % af medianindkomsten. For dette mål finder vi tegn på en svagt stigende grad af segregering inden for kommunerne.

Konklusionerne er ikke følsomme over for, ved hvilket alderstrin vi måler segregeringen, og hvilken metode vi anvender til at måle den. Samlet set viser vores resultater, at der er en ikke ubetydelig grad af segregation på tværs af børnehaver i Danmark, og at denne segregering har været meget konstant.

Det bør bemærkes, at vi her i rapporten har anvendt meget simple mål for barnets socioøkonomiske baggrund. Det er gjort med det formål at gøre analyserne og resultaterne så enkle og transparente som muligt. Det betyder dog også, at målene indeholder en del støj. Blot fordi et barn kommer fra den nedre eller øvre halvdel af forældreindkomstfordelingen, så betyder det ikke nødvendigvis, at barnet kommer fra en anden socioøkonomisk gruppe. Denne afvejning mellem simple og let gennemskuelige mål og præcise, men mere komplicerede mål, er vanskelig at undgå. Det kunne være interessant at undersøge, hvordan konklusionerne ville afvige, hvis vi anvendte mere præcise mål for udsatte børn i samme periode. Her i rapporten har vi dog valgt at fokusere på følsomheden over for definitioner og metoder ved disse simple mål.

Vi har valgt at måle segregationen baseret på forældrenes indkomst og uddannelse. Dette er gjort ud fra en forventning om, at disse variable er relateret til ressourcebehovet i institutionen og til børnenes sociale baggrund. Samtidig anvendes disse variable også i såkaldte socioøkonomiske ressourcetildelinger i kommuner, hvor kommunerne allokerer en del af budgettet til institutionerne baseret på den socioøkonomiske sammensætning (Pedersen et al., 2017). Der kan dog være andre variable, som er mere afgørende (fx børnenes og forældrenes sundhed eller forældrenes arbejdsmarkedstilknytning).

Litteratur

- Christensen, E. (1996). *Daginstitutionen som forebyggende tilbud til truede børn – en undersøgelse af 769 daginstitutioner*. København: SFI – Det Nationale Forskningscenter for Velfærd.
- Duncan, O.D. & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20(2), 210-217.
- Glavind, N. (2004). *Det Delte Danmark: Daginstitutioner: Ulige vilkår for indsats mod "negativ social arv"*. København: AErådet.
- Glavind, N. & Pade, S. (2015). *Leger lige børn bedst? Ulighed og adskillelse i daginstitutioner og skoler*. København: FOA.
- Glavind, N. & Pade, S. (2018). *Flere børn i dagtilbud: Udviklingstendenser for dagpleje og daginstitutioner 2018*. København: FOA.
- Henry, G.T. & Rickman, D.K. (2007). Do Peers Influence Children's Skill Development in Pre-school? *Economics of Education Review*, 26(1), 100-112.
- Massey, D.S. (1990). American apartheid: Segregation and the making of the under-class. *American Journal of Sociology*, 96(2), 329-357.
- Neidell, M. & Waldfogel, J. (2010). Cognitive and Noncognitive Peer Effects in Early Education. *The Review of Economics and Statistics*, 92(3), 562-576.
- Pedersen, H.S., Flarup, L.H. & Petersen, J.S. (2017). *Pædagogisk praksis og socioøkonomisk ressourcefordeling: En undersøgelse i 24 daginstitutioner*. København: KORA – Det Nationale Institut for Kommuners og Regioners Analyse og Forskning.
- Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools. *American Economic Review*, 109(3), 774-809.
- Söderström, M. & Uusitalo, R. (2010). School choice and segregation: Evidence from an admission reform. *Scandinavian Journal of Economics*, 112(1), 55-76.
- Vasiljeva, K., Wandsøe, R., Weissert, C. & Naur, S. (2017). *Segregering i de danske grundskoler*. København: Kraka.

Bilag 1 Ikke-teknisk introduktion til metoderne

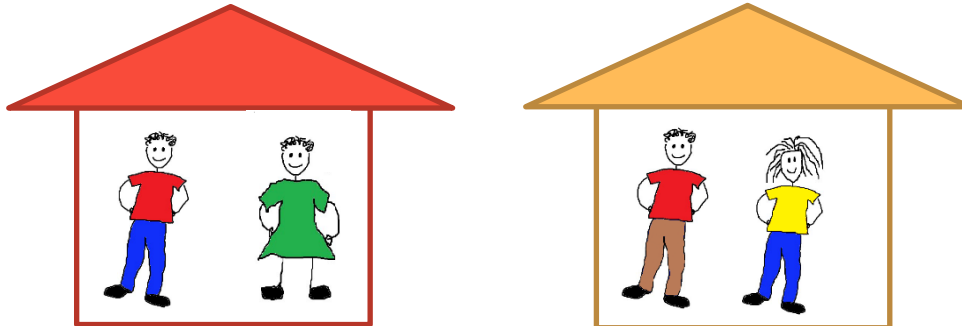
Duncans segregeringsindeks

I den akademiske litteratur er et af de mest anvendte mål for, hvor opdelt samfundet er, Duncans segregationsindeks (Duncan & Duncan, 1955). I dette kapitel vil vi introducere indekset ved en række simple eksempler. Som sagt vil eksemplerne forekomme banale og overflødige, men hvor de første eksempler har fokus på forståelsen af indekset, er de senere eksempler designet på en sådan måde, at de svarer til den virkelige segregation i de danske børnehaver. Af hensyn til illustrationen er eksemplerne baseret på segregationen med hensyn til køn, og vi undersøger derfor, om børnehaver er opdelt efter børnenes biologiske køn. Dette sker udelukkende, fordi det gør fremstillingen lettere. Det er kortere og lettere forståeligt at tale om drenge og piger end om børn fra familier med en ækvivaleret disponibel husstandsindkomst henholdsvis over eller under medianen.

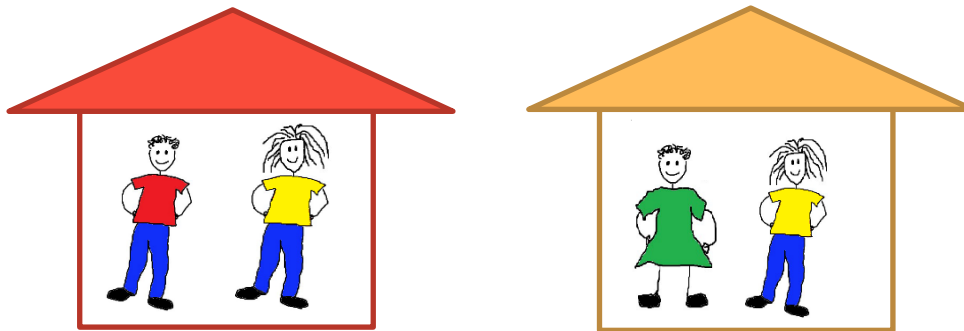
Vi vil altså gerne måle, hvor opdelt børnehaver er med hensyn til børnenes køn. Et nyttigt udgangspunkt er, som ovenfor, at forestille sig scenarierne, hvor der er henholdsvis den mindst mulige og størst mulige opdeling af kønnene. I det første scenarie har alle børnehaver en lige stor andel af det samlede antal piger i samfundet og en lige stor andel af det samlede antal drenge i samfundet. Dette eksempel er illustreret ved scenarie A i bilagsfigur 1.1. Der er i det scenarie to piger og to drenge i samfundet, og de er fordelt på to børnehaver. I hver børnehave går der en pige og en dreng. Det er den mindst mulige opdeling af kønnene. I det andet ekstrem, hvor der er en fuldstændig opdeling efter køn, er der ikke én børnehave, hvor der både er piger og drenge. Dette eksempel er illustreret ved scenarie B i bilagsfigur 1.1. Igen er der to piger og to drenge i samfundet, som er fordelt på to børnehaver, men denne gang går begge drenge i den røde børnehave og begge piger i den gule børnehave. Der er altså ikke en børnehave med både drenge og piger.

Bilagsfigur 1.1 En illustration af en situation uden segregering (scenarie A) og en situation med fuldstændig segregering (scenarie B)

Scenarie A: Piger og drenge er ligeligt fordelt på tværs af to børnehaver. Der er ingen segregering efter køn.



Scenarie B: Piger og drenge er opdelt i de to børnehaver. Der er den højest mulige segregering.



Ovenstående situationer er lette at forstå. Der er fire børn og to børnehaver. Vi kan hurtigt se, at de to scenarier giver meget forskellige opdelinger. I Danmark består en børneårgang af mere end 60,000 børn, og mere end 90 % af disse børn går i en af de mere end 4.000 børnehaver. Vi kan derfor ikke illustrere segregeringen i danske børnehaver på samme måde. Vi må i stedet bruge en metode til systematisk at måle segregeringen. Denne metode skal også gøre det muligt at sammenligne forskellige grader af segregering, som ikke er de ekstreme situationer, hvor der enten ingen segregering er overhovedet, eller hvor samfundet er helt opdelt.

Kan vi sammenfatte scenarierne i bilagsfigur 1.1 i to tal, der giver et direkte fortolkning af segregeringen? Ja, det kan vi med Duncans segregationsindeks (Duncan & Duncan, 1955). Udgangspunktet for indekset er, at den mindst mulige segregering opstår, hvis alle grupper udgør den samme andel inden for en børnehave. Hvis en børnehave fx har halvdelen af alle drenge i samfundet, så skal børnehaven også have halvdelen af alle piger i samfundet. Jo større andelen på tværs af grupperne afviger fra hinanden, desto større er segregeringen. I scenarie B i bilagsfigur 1.1 udgør drenge i den røde børnehave 100 % af den samlede drengebefolkning, mens pigeandelen udgør 0 % af den samlede pigebefolkning. De to andele afviger derfor med 100 procentpoint.

Det er underordnet, om afvigelsen er positiv eller negativ. Hvis afvigelsen er positiv, så udgør den første gruppe en større andel. Hvis afvigelsen er negativ, udgør den anden gruppe en større andel. Vi betragter i stedet kun den numeriske forskel, hvor vi fjerner et eventuelt negativt fortegn (et minus). I bilagsfigur 1.2 viser vi de samlede afvigelser for de to scenarier, som vi præsenterede i bilagsfigur 1.1. Vi ser, at den samlede afvigelse bliver 0 i scenarie A og 2 i

scenarie B. Dividerer vi denne samlede numeriske afvigelse i andelene med 2, får vi Duncans segregationsindeks. I scenarie A er Duncans segregationsindeks således 0, og i scenarie B er det 1. Hvad betyder tallet 1? Det betyder, at 1 ud af 1 børn fra en af de to grupper (piger eller drenge) skal flytte institution for at opnå en ligelig fordeling. Bemærk, at da vi taler om andele, så svarer 1 til 100 %, og et Duncans segregationsindeks på 1 (eller 100 %) betyder, at hvis vi ville opnå en situation uden segregering, så skulle vi flytte 100 % af børnene i en af grupperne (og kun i en af grupperne).

Bilagsfigur 1.2 Duncans segregationsindeks illustreret ved to scenarier

Scenarie A: Piger og drenge er ligeligt fordelt på tværs af to børnehaver. Det er den mindst mulige segregering.



$$\frac{\text{Antal drenge i børnehaven}}{\text{Antal drenge i alt}} - \frac{\text{Antal piger i børnehaven}}{\text{Antal piger i alt}} = \frac{1}{2} - \frac{1}{2} = 0$$



$$\frac{\text{Antal drenge i børnehaven}}{\text{Antal drenge i alt}} - \frac{\text{Antal piger i børnehaven}}{\text{Antal piger i alt}} = \frac{1}{2} - \frac{1}{2} = 0$$

- Den samlede afvigelse er afvigelsen i den røde børnehave plus afvigelsen i den gule børnehave = $0+0=0$.
- Duncans segregationsindeks er dermed $D = \frac{0}{2} = 0$

Scenarie B: Piger og drenge er opdelt i de to børnehaver. Det er den størst mulige segregering.



$$\frac{\text{Antal drenge i børnehaven}}{\text{Antal drenge i alt}} - \frac{\text{Antal piger i børnehaven}}{\text{Antal piger i alt}} = \frac{2}{2} - \frac{0}{2} = 1$$



$$\frac{\text{Antal drenge i børnehaven}}{\text{Antal drenge i alt}} - \frac{\text{Antal piger i børnehaven}}{\text{Antal piger i alt}} = \frac{0}{2} - \frac{2}{2} = -1$$

- Den samlede afvigelse er den numeriske afvigelse i den røde børnehave plus den numeriske afvigelse i den gule børnehave = $1+1=2$.
- Duncans segregationsindeks er dermed $D = \frac{2}{2} = 1$.

Duncans segregationsindeks angiver altså, hvor stor en andel af børnene fra en gruppe, som skulle flytte for at opnå en situation uden segregering. I scenariet ovenfor, hvor drengene og pigerne går i hver sin børnehave, er det da også gældende, at hvis vi flytter alle piger fra den gule børnehave til den røde børnehave, så er der ingen opdeling, idet alle piger og alle drenge går i samme børnehave. Det betyder dog i dette tilfælde, at den ene børnehave efter justeringen er tom, og den anden er blevet dobbelt så stor. Det er naturligvis ikke altid et relevant sammenligningsgrundlag, idet vi i praksis ikke frit kan justere institutionsstørrelserne. Vi vil dog tage højde for dette i analyserne.

Lad os nu betragte et eksempel, som er en anelse mere realistisk end ovenstående. Vi betragter nu en situation, hvor der er 12 børn i samfundet, 6 piger og 6 drenge. De 12 børn er fordelt på to børnehaver. I den røde børnehave er der 3 piger og 5 drenge, og i den gule børnehave er der 3 piger og 2 drenge. Dette scenarie er illustreret i den øverste del af bilagsfigur 1.3.

Bilagsfigur 1.3 Et scenarie med et Duncans segregationsindeks på 0,33

Scenarie A: 5 ud af 6 drenge går i den røde børnehave. Pigerne er ligeligt fordelt på tværs af de to børnehaver.



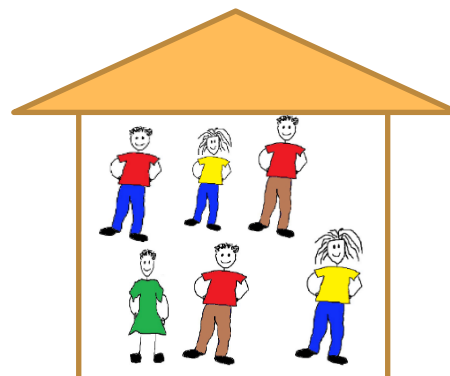
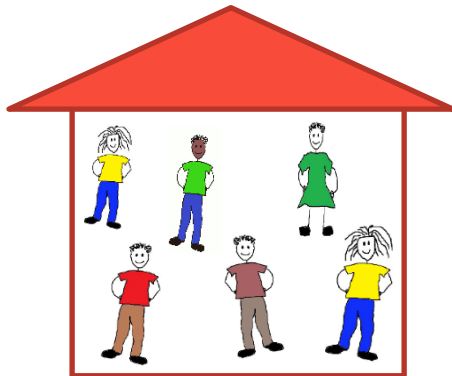
$$\frac{\text{Antal drenge i børnehaven}}{\text{Antal drenge i alt}} - \frac{\text{Antal piger i børnehaven}}{\text{Antal piger i alt}} = \frac{5}{6} - \frac{3}{6} = \frac{2}{6}$$



$$\frac{\text{Antal drenge i børnehaven}}{\text{Antal drenge i alt}} - \frac{\text{Antal piger i børnehaven}}{\text{Antal piger i alt}} = \frac{1}{3} - \frac{3}{6} = -\frac{2}{6}$$

$$\text{Duncan's segregationsindeks} = \frac{1}{2} \times \left(\left| \frac{2}{6} \right| + \left| -\frac{2}{6} \right| \right) = \frac{2}{6}$$

Scenarie A efter vi har flyttet 2 ud af 6 drenge fra den røde til den gule børnehave.



Hvorfor er scenariet i bilagsfigur 1.3 mere realistisk og relevant? For det første er det interessant, fordi det illustrerer en situation, hvor vi ikke er i ekstremerne. Segregationen er ikke fuldstændig, men det er stadig sådan, at det er langt mere sandsynligt, at du møder en dreng i den røde børnehave end i den gule børnehave. For det andet viser det, at vi kan have situationer med segregering, selvom den ene befolkningsgruppe er ligeligt fordelt på tværs af institutioner. Halvdelen af pigerne er i den røde børnehave og den anden halvdel i den gule børnehave. For det tredje er eksemplet udvalgt, fordi det giver et Segregationsindeksværdi på 0,33, som svarer til den segregation, vi har i danske børnehaver målt ved forældrenes indkomst, hvilket vi kommer tilbage i resultatafsnittet. Her er det dog værd at huske billedet af fordelingen af piger og drenge på tværs af den gule og den røde børnehave som svarende til den faktiske indkomst-segregering i Danmark.

Udregner vi Duncans segregationsindeks for scenariet i bilagsfigur 1.3, så får vi, som vist, en værdi på en tredjedel (eller to sjettedele). Det betyder altså, at vi skal flytte 2 ud af 6 piger eller 2 ud af 6 drenge for at opnå en situation uden segregering. Det har vi gjort i den nedre del af bilagsfigur 1.3, hvor vi ikke længere har en segregering efter køn. Vi flyttede i det tilfælde 2 af

drengene i den røde børnehave til den gule børnehave. Vi kunne også have flyttet 2 af pigerne i den gule børnehave til den røde børnehave og opnå en situation uden segregering.

Piger og drenge blev benyttet som eksempler her, men i rapporten undersøger vi opdelingen med hensyn forældrenes indkomst, uddannelse og oprindelse. Vi kan fx opdele alle børn i to lige store grupper efter forældrenes indkomst. Det betyder ikke, at alle børn fra den halvdel, som har den laveste forældreindkomst, lider afsavn eller kan betegnes som fattige. Vi kunne også have dannet to grupper af ulige størrelser, fx de rigeste 90 % og resten. Formålet er blot at undersøge, om der er segregering med hensyn til forældrenes indkomst. For at undersøge om konklusionerne er følsomme over for, hvordan vi har grupperet børnene, undersøger vi også segregeringen med en alternativ metode, som bliver beskrevet nedenfor.

Det er værd at bemærke, at en ligelig fordeling i praksis ikke altid vil være mulig. Betragt fx situationen, hvor der kun er 1 pige og 2 drenge og to børnehaver. Den eneste måde at skabe en ligelig fordeling er ved at lukke den ene børnehave, så alle børn kan gå i den ene børnehave. Alternativt skulle pigen deles i 2, fx ved at hun placeres i den ene børnehave i halvdelen af tiden og i den anden børnehave den anden halvdel af tiden.

Et alternativt segregationsmål: R^2

En væsentlig fordel ved Duncans segregationsindeks er den direkte fortolkning. Værdien angiver, hvor mange børn vi skulle flytte fra en gruppe til en anden for at opnå en situation uden segregering. En væsentlig ulempe ved Duncans segregationsindeks er, at det kun kan anvendes, når segregeringen betragtes med hensyn til to grupper. Vi kan fx bruge Duncans segregationsindeks til at måle segregeringen med hensyn til forældrenes uddannelsesbaggrund, men i så fald vil vi blot skulle opdele alle børn i to grupper bestående af børn fra henholdsvis højt- og lavtuddannede hjem. Vi vil altså med Duncans segregationsindeks ikke kunne tage højde for forskelle i forældrenes uddannelsesniveau.

For at illustrere ulempen ved Duncans segregationsindeks har vi i bilagstabel 1.1 vist to scenarier, hvor 8 børn er fordelt på to børnehaver. Der er 4 børn i hver børnehave, og tabellen viser forældrenes gennemsnitlige uddannelsesniveau. I den første børnehave er forældrenes gennemsnitlige uddannelsesniveau 16,25 år. I den anden børnehave er gennemsnittet 12 år. Det er derfor nærliggende at konkludere, at der er en betydelig grad af segregering. Børnene i den første børnehave kommer fra hjem, hvor forældrenes gennemsnitlige uddannelsesniveau er 4 år længere end i den anden børnehave. Men hvis beslutter os for at danne to lige store grupper af børn efter forældrenes uddannelsesniveau, så vil Duncans segregationsindeks fortælle os, at der ikke er nogen segregering. Begge børnehaver har 2 børn fra henholdsvis lavt- og højtuddannede hjem.

Bilagstabel 1.1 Sammenligning mellem Duncans segregationsindeks og R²-metoden

Scenarie A: Der er ingen segregering ifølge Duncans segregationsindeks, men ifølge R²-metoden er der segregering.

Forældrenes gennemsnitlige uddannelseslængde				
	Antal år		Højtuddannet (mindst 14 år)	
	Børnehave 1	Børnehave 2	Børnehave 1	Børnehave 2
Barn 1	12,50	9,00	Nej	Nej
Barn 2	13,50	9,50	Nej	Nej
Barn 3	18,00	14,50	Ja	Ja
Barn 4	21,00	15,00	Ja	Ja
Gennemsnit	16,25	12,00	To af hver	To af hver

R²: 0,32

Duncan's segregationsindeks: 0,00.

Scenarie B: Der er ingen segregering ifølge R²-metoden, men ifølge Duncans segregationsindeks er der segregering.

Forældrenes gennemsnitlige uddannelseslængde				
	Antal år		Højtuddannet (mindst 14 år)	
	Børnehave 1	Børnehave 2	Børnehave 1	Børnehave 2
Barn 1	12,50	18,00	Nej	Ja
Barn 2	13,50	9,50	Nej	Nej
Barn 3	9,00	14,50	Nej	Ja
Barn 4	21,00	15,00	Ja	Ja
Gennemsnit	14.125	14.125	En høj og tre lave	En lav og tre høje

R²: 0,00

Duncans segregationsindeks: 0,50.

Udfordringen opstår, de 2 børn fra gruppen af lavtuddannede meget tæt på at tilhøre gruppen af højtuddannede i den første børnehave. Der er meget variation inden for grupperne. Blandt gruppen af børn fra hjem med højtuddannede forældre er der forældre med 14,5 og 21 års uddannelse. Det samme er gældende for gruppen af børn fra lavtuddannede hjem. Her er der mellem 9 og 13,5 års uddannelse. Grænsen på 14 år for at tilhøre gruppen af højtuddannede er tilfældigt sat for at skabe to lige store grupper, men konklusionen afhænger meget af, hvor denne grænse sættes.

Kan vi måle segregeringen på en måde, hvor vi tager højde for al variation i baggrundsvariablene? Ja, vi kan måle, hvor stor en del af variationen i forældrenes uddannelsesniveau vi kan forklare ved at se på, hvor børnene går i børnehave. Et godt bud på forældrenes uddannelsesniveau i den første børnehave i scenarie 1 er gennemsnittet, som er på 16,25 år, mens et godt gæt for den anden børnehave er 12 år. Hvis vi ikke ved noget om, hvor barnet går i børnehave, må vi stedet bruge det samlede gennemsnit som vores bedste gæt. Vi kan nu sammenligne de to gæt ved at udregne den samlede afvigelse mellem gættene og det faktiske tal. Jo mindre den samlede afvigelse er, des bedre er metoden. Til sidst sammenligner vi den samlede afvigelse, når vi bruger børnehavens gennemsnit som bedste gæt, og når vi bruger det samlede gennemsnit som bedste gæt. Til sidst trækker vi forholdet mellem de to afvigelser fra tallet 1 og får så R². Jo mindre afvigelserne med de børnehavespecifikke gæt er i forhold til afvigelserne med det samlede gennemsnit, des større bliver R². Hvad betyder et stort R² så? Det betyder, at hvis vi bare ved, hvor barnet går i børnehave, så har vi også en god idé om forældrenes uddannelse, fordi afvigelserne ved kun at bruge børnehavespecifikke forældregennem-

snit er meget små. Det betyder altså, at jo større R^2 er, desto mere segregerede er børnehaverne. I scenarie A i bilagstabel 1.1 får vi da også en R^2 -værdi på 0,32. Den direkte fortolkning er, at 32 % af variationen i forældrenes uddannelsesniveau kan forklares ved at kende gennemsnittet for børnehaven. I de følgende afsnit giver vi en mere formel beskrivelse af R^2 -metoden.

I scenarie B i bilagstabel 1.1 har 2 børn flyttet institution i forhold til i scenarie A. Efter vi har gjort dette, er Duncans segregationsindeks steget fra 0 til 0,5 og R^2 faldet fra 0,32 til 0. Forældrenes gennemsnitlige uddannelsesniveau er nu helt ens i de to børnehaver, hvilket gør, at R^2 falder til 0. Det giver os ikke meget information om forældrenes uddannelsesniveau, at vi ved, hvilken børnehave barnet går i, fordi gennemsnittet er helt ens. Omvendt, er der kun 1 barn fra et højtuddannet hjem i den første børnehave, hvilket giver et højt Duncan Segregationsindeks, og vi ville skulle flytte halvdelen af børnene fra en gruppe for at opnå en ligelig fordeling uden segregering ifølge Duncans segregationsindeks.

Eksemplerne i bilagstabel 1.1 viser, at graden af segregering meget afhænger af, hvordan vi måler segregeringen. Derudover viser eksemplerne, at der ikke er en entydig sammenhæng mellem de forskellige mål. Vi kan reducere segregeringen ifølge det ene mål og samtidig øge segregeringen ifølge det andet. Derfor vil rapporten vise resultater med begge mål. Ulempen ved R^2 -tilgangen er, at den er mindre intuitiv end Duncans segregationsindeks, og resultaterne kan ikke fortolkes helt så let.

Det relevante sammenligningsgrundlag

Når vi bruger Duncans segregationsindeks eller R^2 -metoden til at evaluere segregeringen, så tager vi ikke højde for institutionernes størrelse. Når Duncans segregationsindeks angiver, at vi skal flytte halvdelen af børnene i en gruppe, så betyder det, at størrelserne vil ændre sig. Nogle børnehaver vil blive større og nogle mindre. Det er ikke altid muligt at gennemføre sådan en justering i praksis. Vi kan ikke bare tilføje flere børn til en institution.

En yderligere udfordring er, at vi ikke blot kan flytte et barn fra Esbjerg til København for at reducere segregeringen. En væsentlig del af forskellene mellem børnehaver kan skyldes forskelle i, hvor børnene bor. Så selvom det overordnede segregationsindeks er relevant for at få en idé om det samlede omfang af segregering, er det ikke videre informativt om, hvad der i praksis ville være muligt.

Det er vanskeligt at tage højde for, hvor barnet bor, men for at give et bud på, hvor meget segregeringen potentielt set kan reduceres, vil vi tage højde for, hvilken kommune barnet bor i. Vi gør dette ved at spørge os selv: Hvad ville segregeringen være, hvis alle børn blev allokeret til en tilfældig børnehave inden for samme kommune og ved at holde børnehavernes størrelse konstant? En tilfældig allokering kan selvfølgelig give det samme udfald som den faktiske allokering eller en allokering med endnu mere segregering. Vi udfører derfor 100 tilfældige allokeringer og udregner den gennemsnitlige segregering på tværs af de 50 tilfældige allokeringer. I disse allokeringer tager vi desuden højde for alderssammensætningen i børnehaven samt børnehavetyper (dvs. om det er en aldersintegreret institution eller en almindelig børnehave).

Bilag 2 Teknisk beskrivelse af metoderne

Duncans segregeringsindeks

I dette afsnit vil vi kort formalisere Duncans segregationsindeks. Indekset er introduceret ovenfor, og gennemgangen i dette kapitel oversætter blot illustrationerne til en formel, der også viser, hvordan målet opgøres, hvis der er flere end to børnehaver. Duncans segregationsindeks udregnes ved følgende formel:

$$D = \frac{1}{2} \sum_{i=1}^N \left| \frac{g1_i}{G1} - \frac{g2_i}{G2} \right|$$

hvor D er segregeringsmålet, N er antal enheder eller institutioner (fx antal skoler, kommuner, børnehaver), $g1_i$ er antallet af individer fra befolkningsgruppe 1 i enhed i , og $g2_i$ er antallet af individer fra befolkningsgruppe 2 i enhed i . $G1$ og $G2$ er det samlede antal af individer i henholdsvis gruppe 1 og 2 (fx piger og drenge). Ligesom i eksemplerne ovenfor udregner vi altså først hver enheds afvigelse fra en proportional fordeling og bruger så summen over alle enheders afvigelse.

Duncans segregeringsmål har været anvendt til at kvantificere segregeringen gennem mange årtier og på mange områder. Metodens egenskaber og relation til andre mål gennemgås i den videnskabelige artikel "A Methodological Analysis of Segregation Indexes" af Otis Dudley Duncan og Beverly Duncan, som blev udgivet i the American Sociological Review i 1955. Ud over at metoden giver en direkte fortolkning, idet det angiver andelen af en befolkningsgruppe, som skal flyttes for at opnå en ligelig fordeling (eller ingen segregering), så kan målet også illustreres grafisk.

Bilagsfigur 2.1 Formlen for Duncans segregationsindeks illustreret ved scenarierne fra bilagsfigur 1.1

Scenarie A: Piger og drenge er ligeligt fordelt på tværs af to børnehaver. Der er ingen segregering efter køn.

$$G_1 = 2, G_2 = 2, g1_1 = 1, g2_1 = 1, g1_2 = 1 \text{ og } g2_2 = 1$$

$$D = \frac{1}{2} \sum_{i=1}^N \left| \frac{g1_i}{G1} - \frac{g2_i}{G2} \right| = \frac{1}{2} \left(\left| \frac{1}{2} - \frac{1}{2} \right| + \left| \frac{1}{2} - \frac{1}{2} \right| \right) = 0$$

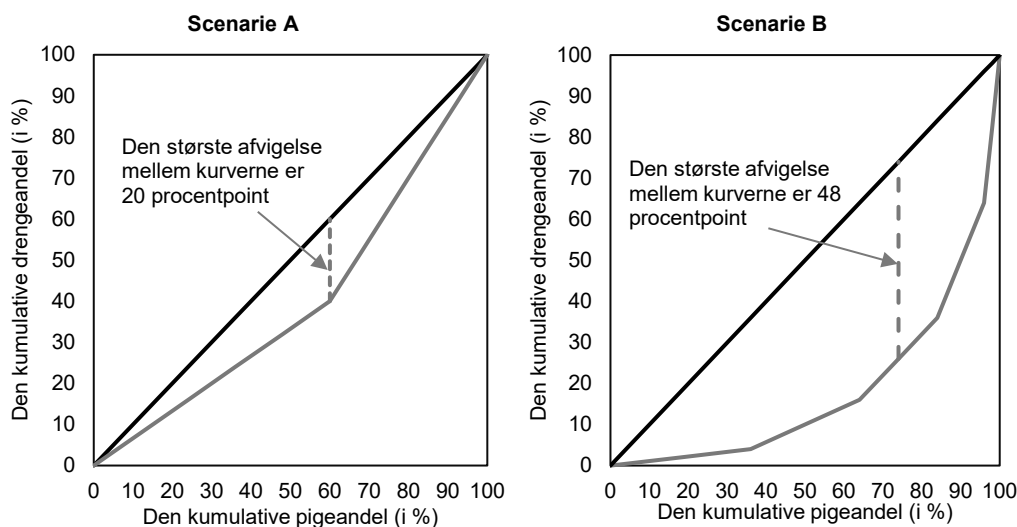
Scenarie B: Piger og drenge er opdelt i de to børnehaver. Det er den højeste mulige segregering.

$$G_1 = 2, G_2 = 2, g1_1 = 2, g2_1 = 0, g1_2 = 0 \text{ og } g2_2 = 2.$$

$$D = \frac{1}{2} \sum_{i=1}^N \left| \frac{g1_i}{G1} - \frac{g2_i}{G2} \right| = \frac{1}{2} \left(\left| \frac{2}{2} - \frac{0}{2} \right| + \left| \frac{0}{2} - \frac{2}{2} \right| \right) = 1$$

Bilagsfigur 2.2 En grafisk illustration af Duncans segregationsindeks med 10 børnehaver

	Børnehave										I alt
	BH 1	BH 2	BH 3	BH 4	BH 5	BH 6	BH 7	BH 8	BH 9	BH 10	
Scenarie A:											
Antal piger	6	6	6	6	6	4	4	4	4	4	50
Antal drenge	4	4	4	4	4	6	6	6	6	6	50
Duncans	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,20
Scenarie B:											
Antal piger	9	9	7	7	5	5	3	3	1	1	50
Antal drenge	1	1	3	3	5	5	7	7	9	9	50
Duncans	0,08	0,08	0,04	0,04	0,00	0,00	0,04	0,04	0,08	0,08	0,48



For at illustrere segregeringen af daginstitutioner grafisk starter vi med at sortere alle børnehaver efter deres andel af drenge. Den første børnehave er altså den børnehave med den mindste andel af drenge. Efter vi har sorteret børnehaverne, så udregner vi de kumulative andele af henholdsvis piger og drenge. Det betyder, at vi for den børnehave i rækken lægger andelene for den første og anden børnehave sammen. For den tredje børnehave lægger vi andelene for den første, anden og tredje børnehave sammen. Når vi ser på de kumulative andele for den tredje børnehave, kan vi altså aflæse de første tre børnehavers samlede andel af drenge i samfundet. Denne fremgangsmåde svarer nogenlunde til, hvad vi normalt gør, når vi er interesseret i at opgøre indkomst eller formueuligheden i samfundet med Lorenz-kurven og Gini-indekset. I stedet for andelen af landets samlede indkomst ser vi her på landets andel af drenge. Ligesom med Lorenz-kurven kan vi sammenligne den med en hypotetisk kurve uden ulighed (dvs. uden segregering). Jo større afvigelsen af kurven er fra denne hypotetiske kurve, desto større er segregeringen. Faktisk er det sådan, at Duncans segregeringsmål svarer til den største afvigelse mellem de to kurver. Dette er illustreret ved scenarie A og B i bilagsfigur 2.2.

Et alternativt segregationsmål: R^2

Det alternative mål, som vi anvender, er R^2 ("R-i-anden"), der kendes som redskab til at evaluere "Goodness-of-fit" i regressioner. Denne metode er bl.a. blevet brugt i et studie om skolevalg og segregation i Sverige (Söderström & Uusitalo, 2010). Ideen bag dette mål er følgende: I en simpel lineær model finder man den lige linje, som "bedst" beskriver sammenhængen

mellem to variable. Hvordan denne lige linje findes, afhænger af en række antagelser, bl.a. hvordan man definerer "bedst". Den mest almindelige metode er mindste kvadraters metode, hvor man finder den linje, som minimerer summen af de kvadrerede forskelle mellem linjen og de faktiske punkter. For at få et mål for, hvor godt denne linje passer til data, sammenligner man typisk summen af disse kvadrerede afvigelse med summen af den totale afvigelse i forhold til gennemsnitsværdien. Trækker man denne værdi fra 1, får man R^2 , som dermed kan fortolkes som andelen af variationen i den ene variabel, som forklares ud fra den anden variabel i denne model.

Da R^2 giver et mål for, hvor godt en variabel forklarer variationen i en anden variabel, kan vi direkte anvende dette mål på børnehaver. Hvis en børnehave forklarer en høj del af variationen i forældrenes baggrund, så betyder det, at der er stor forskel i forældrenes baggrund på tværs af daginstitutioner. Vi gør derfor det, at vi estimerer følgende regressionsmodel med mindste kvadraters metode:

$$y_i = \alpha + \beta' \mathbf{I}_i + u_i$$

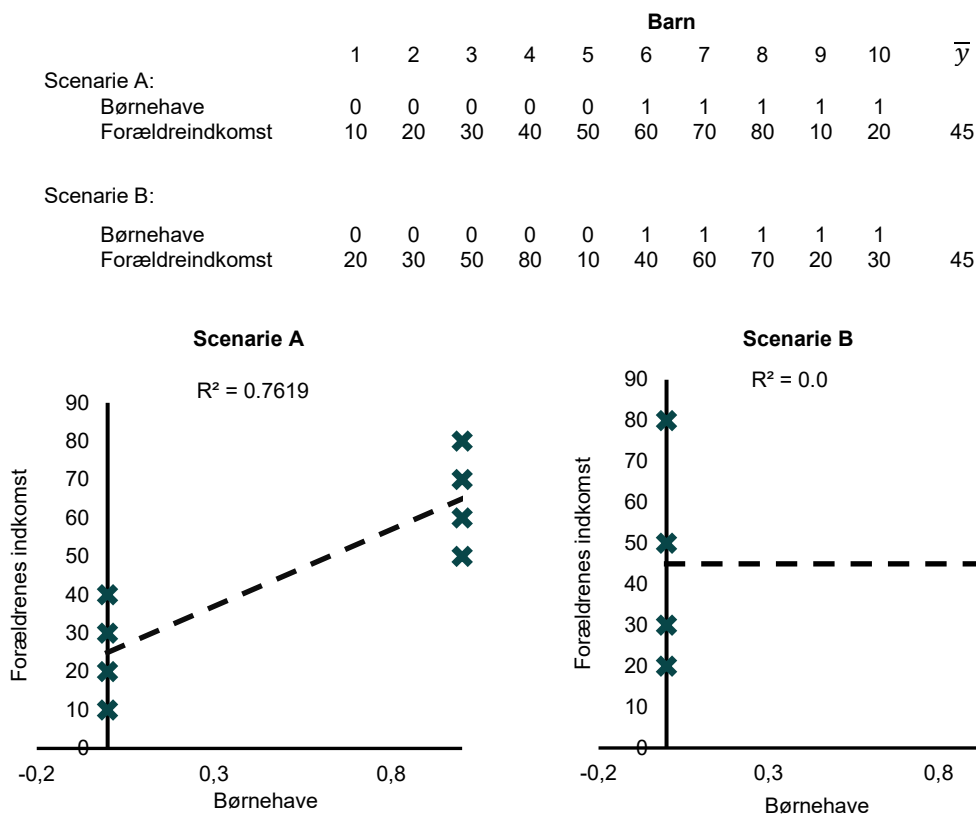
hvor y_i er et mål for forældrebaggrund (fx forældrenes indkomst) for barn i og \mathbf{I}_i er en række variable for alle institutioner, hvor variabelen har værdien 1, hvis barnet er indskrevet i den pågældende institution og 0 hvis ikke. u_i er et fejledd. Efter vi har estimeret ovenstående model udregner vi R^2 ved følgende udtryk:

$$R^2 = 1 - \frac{\sum_i^N (y_i - \hat{y})^2}{\sum_i^N (y_i - \bar{y})^2}$$

hvor \bar{y} er gennemsnittet af y og $\hat{y} = \hat{\alpha} + \hat{\beta}' \mathbf{I}_i$. Med andre ord er \bar{y} vores bedste bud på værdien af y givet vores estimerede parametre $\hat{\alpha}$ og $\hat{\beta}$. En højere værdi af R^2 betyder, at vi kan forklare en større andel af variationen i y med \mathbf{I}_i . Overført til vores eksempel med børnehaver betyder det, at R^2 giver et mål for andelen af variationen i forældrenes indkomst, som kan forklares ved forældrenes indkomst.

Bilagsfigur 2.3 illustrerer R^2 -metoden ved brug af to eksempler. I eksemplerne har vi to børnehaver og 8 børn. Vi kalder børnehaverne børnehave 0 og børnehave 1. Børnenes forældre har en indkomst (der altså er vores y), som går fra 10 til 80. I scenarie A har vi den maksimale segregering i scenariet. Den fattigste halvdel af børnene går i børnehave 0 og den rigeste halvdel i børnehave 1. Hvis vi bruger ovenstående formel, får vi, at $\hat{\alpha}$ er med 40, og $\hat{\beta}$ er 25. Det betyder med andre ord, at for børn i børnehave 0 er det bedste gæt på deres forældres indkomst, at den er 40, og for børnene i børnehave 1 er det bedste bud $40 + 25 = 65$. Kunne vi være nået frem til denne konklusion på en lettere måde? Ja, for de bedste bud svarer blot til det gennemsnitlige for forældrene i hver gruppe. Vores segregeringsmål – R^2 – er i det tilfælde 0,76. Modsat i scenarie B. Her er børnene ikke perfekt sorteret på tværs af børnehaver efter forældrebaggrund. Der er børn fra begge grupper i begge børnehaver. Faktisk er den gennemsnitlige forældreindkomst helt ens i de to grupper. Det gør også, at R^2 i dette tilfælde bliver 0. Det er altså en situation med minimal segregering.

Bilagsfigur 2.3 En grafisk illustration af R^2 -metoden



Hvis vi ser på krydserne i bilagsfigur 2.3, kan vi også få et indtryk af segregeringen. I scenarie A er alle krydserne for børnehave 0 længere nede end i børnehave 1. Det betyder, at så snart vi ved, at et barn går i børnehave 1, så ved vi, at barnet må være fra den øverste halvdel af børnene med hensyn til forældreindkomst.

Det modsatte gør sig gældende i scenarie B i bilagsfigur 2.3. Hvis vi kun ved, at et barn er fra børnehaven 1, men ikke har anden information om barnet, så kan vi ikke sige noget om, hvorvidt barnet kommer fra den rigeste eller fattigste halvdel. Der er ingen segregering i dette tilfælde.

Vi vil anvende samme fremgangsmåde i analyserne som i eksemplerne her, men blot på registerdata fra Danmarks Statistik med mere end 180.000 børn og mere end 4.000 børnehaver. Regressionsmodeller anvendes ofte til at afdække årsagssammenhænge, og dette eksempel illustrerer tydeligt mekanikken i metoden, og at den ikke nødvendigvis afslører årsagssammenhængen. Vi kan ikke sige, at forældrene har en højere indkomst, fordi deres børn går i en bestemt børnehaven ud fra denne analyse.

Bilag 3 Tabeller om dataudvælgelsen

Bilagstabel 3.1 Antal børn og institutioner

Kalenderår	Børn				Institutioner			
	3 år	4 år	5 år	I alt	3 år	4 år	5 år	I alt
2005	58.429	60.740	61.847	181.016	4.690	4.091	4.083	4.717
2006	58.473	60.660	60.641	179.774	4.585	4.102	4.083	4.607
2007	60.700	61.695	60.295	182.690	4.500	4.099	4.078	4.524
2008	57.576	62.148	61.214	180.938	4.325	4.060	4.058	4.357
2009	56.364	61.484	60.532	178.380	4.231	3.905	3.909	4.265
2010	57.235	62.024	60.362	179.621	4.120	3.853	3.843	4.138
2011	57.802	60.715	60.647	179.164	3.941	3.680	3.691	3.973
2012	55.502	61.065	59.453	176.020	3.713	3.492	3.490	3.748
2013	55.144	57.720	59.307	172.171	3.348	3.183	3.191	3.376
2014	51.051	57.493	56.349	164.893	3.157	3.031	3.037	3.182

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Bilagstabel 3.2 Indkomst og uddannelsesgrænser

År	Indkomst (DKK, 2015 niveau)				Uddannelse (år)		
	Ækvivaleret husstandsindkomst		Gennemsnit på tværs af forældre		Gennemsnit på tværs af forældre		
	P50 SÅ	0.6*P50 SÅ	P50 FÅ-1	P50 FÅ+5	P50 SÅ	P50 FÅ-1	P50 FÅ+5
2005	290.000	180.000	180.000	230.000	14	14	14
2006	300.000	180.000	180.000	230.000	14	14	14
2007	300.000	180.000	180.000	240.000	14	14	14
2008	300.000	180.000	180.000	230.000	14	14	14
2009	310.000	190.000	190.000	240.000	14	14	14
2010	240.000	140.000	200.000	250.000	14	14	14
2011	240.000	140.000	200.000	250.000	14	14	14
2012	240.000	140.000	200.000	250.000	14.2	14	14
2013	240.000	140.000	210.000	250.000	14.2	14	14
2014	240.000	140.000	210.000	260.000	14.3	14	14

Anm.: P50: medianen. SÅ: Samme år. FÅ: Fødselsår. Tallene angiver henholdsvis indkomstgrænser og uddannelsesgrænser for analyserne baseret på Duncans segregationsindeks. Alle tal dækker over mindst 10 personer.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Bilag 4 Betydningen af, hvordan og hvornår økonomiske ressourcer måles

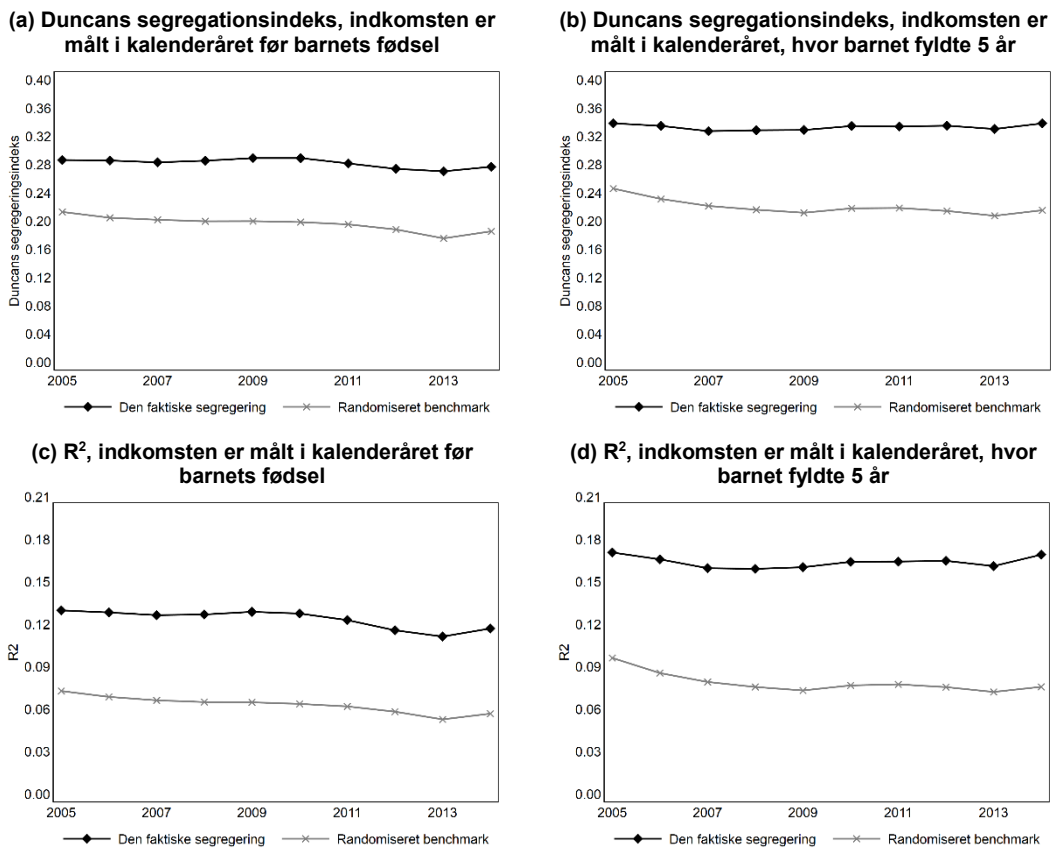
Segregeringen er baseret på forældrenes indkomst målt i året før barnet blev født og i det kalenderår, hvor barnet fyldte 5 år.

I bilagsfigur 4.1 viser vi segregeringen målt ved henholdsvis Duncans segregationsindeks (øverst) og ved R^2 -metoden (nederst), samt hvor indkomsten er målt henholdsvis i kalenderåret før barnets fødsel (til venstre) og i kalenderåret, hvor barnet fyldte 5 år (til højre). Figurene viser kun niveauerne for det faktiske segregeringsniveau og det randomiserede sammenligningsgrundlag, men forskellene mellem de to kan aflæses i bilagstabel 3.2.

Det er værd at bemærke, at målingen af økonomiske ressourcer i bilagsfigur 4.1 i flere aspekter afviger fra de tidligere metoder. For det første er indkomsten målt på et andet alderstrin for barnet. For mange familier med børn i den aldersgruppe vil indkomstmålet ikke være et godt mål for familiens ressourcer. Det kan fx være, at forældrene stadig er under uddannelse, og at deres indkomst derfor vil stige markant inden for den nærmeste fremtid, når de træder ind på arbejdsmarkedet. Det er derfor vigtigt at undersøge, hvor følsomme konklusionerne er med hensyn til, hvornår indkomsten måles. En anden ændring i forhold til tidligere er, at vi her måler den gennemsnitlige indkomst på tværs af forældrene og ikke den ækvivalerede husstandsindkomst. Vi tager derfor ikke højde for familiens størrelse, og vi benytter de biologiske forældres indkomst i stedet for indkomsten i den husstand, hvor barnet bor. Der er både faglige og praktiske grunde for at gøre dette. Den faglige grund er, at selvom et barn muligvis ikke bor sammen med begge biologiske forældre, så kan begge biologiske forældres økonomiske ressourcer påvirke barnet. Den anden grund er rent praktisk: Det er vanskeligt at vurdere, hvilken husstand barnet ville have tilhørt i kalenderåret, før barnet blev født.

Overordnet set bekræfter resultaterne i bilagsfigur 4.1 de øvrige resultater. Med hensyn til økonomiske ressourcer ser vi et meget stabilt segregeringsniveau af børn på tværs af børnehaver, som går ud over, hvad vi ville forvente givet de strukturelle rammer og de geografiske forskelle. Segregeringsniveauerne varierer lidt på tværs af, hvornår indkomsten måles, og vi kommer således frem til en lidt større segregering, hvis vi måler indkomsten ved 5-årsalderen sammenlignet med de øvrige resultater. Det er i god tråd med forventningen om mere støj og målefejl i indkomsten, når den er målt tidligere.

Bilagsfigur 4.1 Segregeringen i børnehaver i Danmark målt ved Duncans segregationsindeks og R^2 -metoden for årene 2005-2014 baseret på forældrenes gennemsnitlige disponible indkomst



Note.: Figurene viser Duncans segregeringsindeks og R^2 for årene 2005-2014 for kohorterne 2002-2011 baseret på den gennemsnitlige disponible indkomst for de biologiske forældre i henholdsvis kalenderåret, før barnet blev født, og kalenderåret, hvor barnet fyldte 5 år.

Det randomiserede sammenligningsgrundlag viser det gennemsnitlige segregeringsniveau på tværs af 50 tilfældige allokeringer af børn til børnehaver, hvor vi tager højde for barnets alder, bopælskommune, institutionstyper og institutionsstørrelse.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

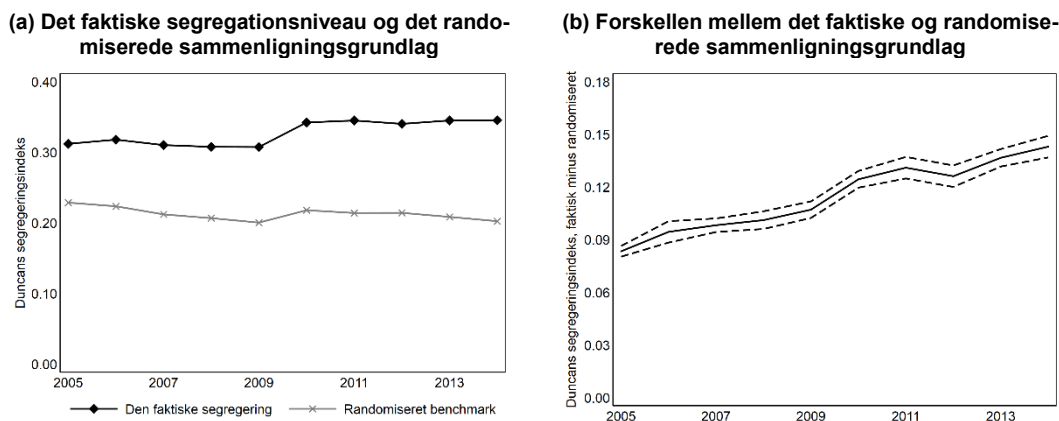
Segregering baseret på, om barnet kommer fra en husstand med en indkomst under 60 % af medianen

Indtil videre har vi opdelt børnene i to lige store grupper, når vi har anvendt Duncans segregationsindeks. Vi vil nu se på segregeringen med hensyn til økonomiske ressourcer, hvor den ene gruppe er noget mindre. Vi opdeler nu alle børn i to grupper, hvor den ene gruppe består af alle børn fra husstande med en ækvivaleret disponibel husstandsindkomst, som er under 60 % af medianindkomsten. Denne gruppe kan betegnes som lavindkomstgruppe, idet definitionen svarer til gængse definitioner for relativ fattigdom (denne definition svarer fx til Eurostats definition af "at-risk-of-poverty"). Denne definition minder i højere grad om definitionerne i Christensen (1996) og Glavind (2004), idet vi har indsnævret den ene gruppe til at være en mere "sårbar" lavindkomstgruppe.

Resultaterne for segregeringen baseret på lavindkomstgruppen er vist i bilagsfigur 4.2. Samlet set er resultaterne i tråd med de øvrige resultater. Vi ser et stabilt segregeringsniveau på omkring 0,33 baseret på Duncans segregationsindeks, hvor cirka to tredjedele af segregeringen kan forklares af forskelle på tværs af kommuner og institutionelle rammer. Ser vi på denne

forskel mellem det faktiske segregeringsniveau og det randomiserede sammenligningsgrundlag i bilagsfigur 4.2.b, så ser vi dog en lidt kraftigere stigning i forskellen sammenlignet med, hvad vi så tidligere.

Bilagsfigur 4.2 Segregeringen i børnehaver i Danmark målt ved Duncans segregationsindeks for årene 2005-2014, baseret på om barnet kommer fra en husstand med en disponibel indkomst under 60 % af medianindkomsten



Anm.: Figurene viser Duncans segregationsindeks for årene 2005-2014 for kohorterne 2002-2011 baseret på den ækvivalerede disponible husstandsindkomst i det pågældende kalenderår. De to grupper består af henholdsvis alle børn fra husstande med en ækvivaleret disponibel husstandsindkomst på under 60 % af medianindkomsten.

Det randomiserede sammenligningsgrundlag viser det gennemsnitlige segregationsniveau på tværs af 50 tilfældige allokeringer af børn til børnehaver, hvor vi tager højde for barnets alder, bopælskommune, institutionstyper og institutionsstørrelse. De stiplede linjer viser 95 % konfidensintervallet, som er baseret på 200 bootstrap-iterationer.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Vi har nu set, at de overordnede konklusioner vedrørende segregeringen på tværs af børnehaver baseret på familiens økonomiske ressourcer er forholdsvis robuste til, hvordan og hvornår vi måler den. Spørgsmålet er nu: Er segregeringen stor? Det er vanskeligt at svare på, om den er stor eller lille. Vi kan dog sige tre ting: For det første kan vi se, at segregeringen er betydeligt større, end hvis børn blev tilfældig, fordelt på tværs af institutioner. For det andet kan vi sammenligne segregeringen med den simple segregering i bilagsfigur 4.1, hvor der blot er to børnehaver og 12 børn. Her går 5 af drengene i den ene børnehave og kun 1 dreng i den anden børnehave. Det svarer til, at 5 af højindkomstbørnene går i den ene børnehave og kun 1 af højindkomstbørnene i den anden børnehave. For det tredje kan vi sammenligne segregeringen i børnehaverne med segregeringen i grundskolen.

Betydningen af, hvornår forældrenes uddannelsesniveau måles

Segregering baseret på forældrenes uddannelse målt i året, før barnet blev født, og i det kalenderår, hvor barnet fyldte 5 år

En væsentlig udfordring ved at kvantificere børns socioøkonomiske baggrund er, at forældrenes indkomst og uddannelsesniveau ikke nødvendigvis er retvisende for indkomsten og uddannelsen i hele barndommen. Dertil kommer, at indkomst og uddannelse er meget grove mål for barnets baggrund og ressourcerne i barnets hjem. Fordelen ved uddannelse og indkomst er, at det er forholdsvis let at måle. Det har betydning både for analyser og forskning, men også, hvis man vil indføre redskaber til at justere institutioners sammensætning af ressourcer.

En udfordring ved at måle segregering ved forældrenes gennemsnitlige uddannelse er, at forældrene muligvis ikke har afsluttet deres uddannelse på det pågældende tidspunkt. Det er

derfor vigtigt at undersøge, hvor følsomme konklusionerne er med hensyn til, hvornår forældrenes uddannelse måles.

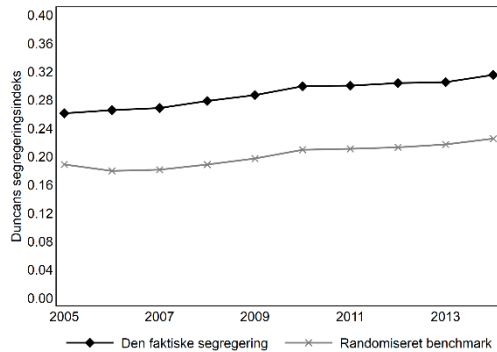
I bilagsfigur 4.3 viser vi segregeringen målt ved henholdsvis Duncans segregationsindeks (øverst) og ved R^2 -metoden (nederst), hvor forældrenes uddannelse er målt henholdsvis i kalenderåret før barnets fødsel (til venstre) og i kalenderåret, hvor barnet fyldte 5 år (til højre). Figurerne viser kun niveauerne for det faktiske niveau og det randomiserede sammenligningsgrundlag, men forskellene mellem de faktiske og det randomiserede sammenligningsgrundlag findes i bilagstabel 5.3.

Resultaterne i bilagsfigur 4.3 bekræfter de tidligere resultater: Vi ser et svagt stigende segregeringsniveau hen over perioden, uanset hvordan vi måler segregeringen (med Duncans segregationsindeks eller R^2 -metoden), og uanset hvornår vi måler forældrenes uddannelsesniveau (i kalenderåret, før barnet blev født, eller i kalenderåret, hvor barnet fyldte 5 år).

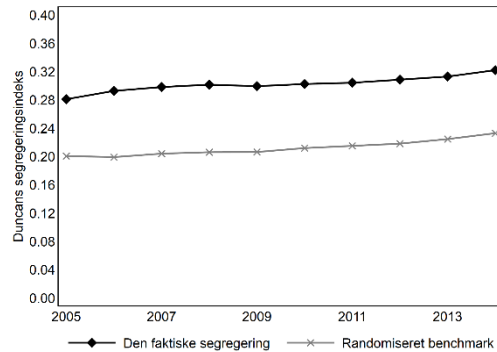
Betragter vi de præcise niveauer, så ser vi et mønster, der er i god tråd med konklusionerne for segregeringen baseret på økonomiske ressourcer. Overordnet set er segregeringen på nogenlunde samme niveau, uanset hvornår vi måler forældrenes uddannelse, men både for Duncans segregationsindeks og for R^2 -metoden ser vi, at segregeringen er lidt større, når vi måler forældrenes uddannelse senere. Det kunne tyde på, at der er nogen "målefejl", når vi måler segregeringen tidligere, fx hvor børn placeres i lavuddannelsesgruppen, fordi deres forældre endnu ikke har gennemført den lange uddannelse. Overordnet set er forskellene dog ikke store.

Bilagsfigur 4.3 Segregeringen i børnehaver i Danmark målt ved Duncans segregationsindeks for årene 2005-2014, baseret på forældrenes uddannelsesnivea

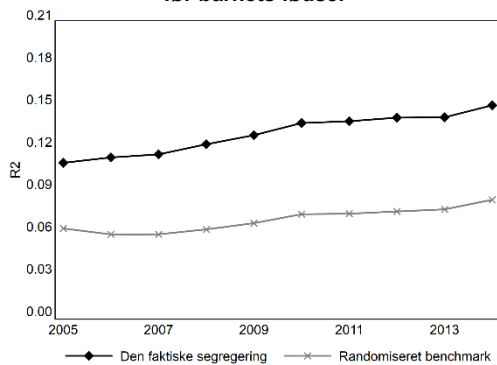
(a) Duncans segregationsindeks, uddannelsesniveaet er målt i kalenderåret før barnets fødsel



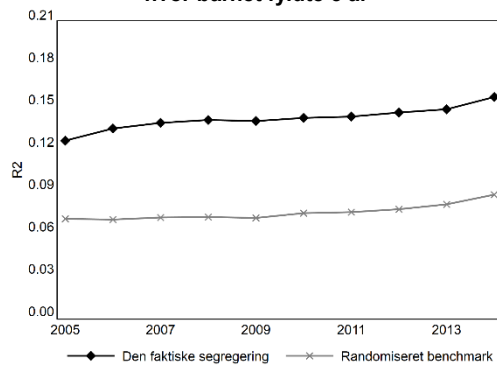
(b) Duncans segregationsindeks, uddannelsesniveaet er målt i kalenderåret, hvor barnet fyldte 5 år



(a) R², uddannelsesniveaet er målt i kalenderåret før barnets fødsel



(b) R², uddannelsesniveaet er målt i kalenderåret, hvor barnet fyldte 5 år



Note.: Figurene viser R² for årene 2005-2014 for kohorterne 2002-2011 baseret på forældrenes uddannelsesnivea. De to grupper består af alle børn af forældre med et gennemsnitligt uddannelsesnivea henholdsvis over og under medianen. Forældrenes uddannelsesnivea er målt i det oktober i henholdsvis kalenderåret, før barnet blev født, og i kalenderåret, hvor barnet fyldte 5 år.

Det randomiserede sammenligningsgrundlag viser det gennemsnitlige segregeringsnivea på tværs af 50 tilfældige allokeringer af børn til børnehaver, hvor vi tager højde for barnets alder, bopælskommune, institutionstyper og institutionsstørrelse.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

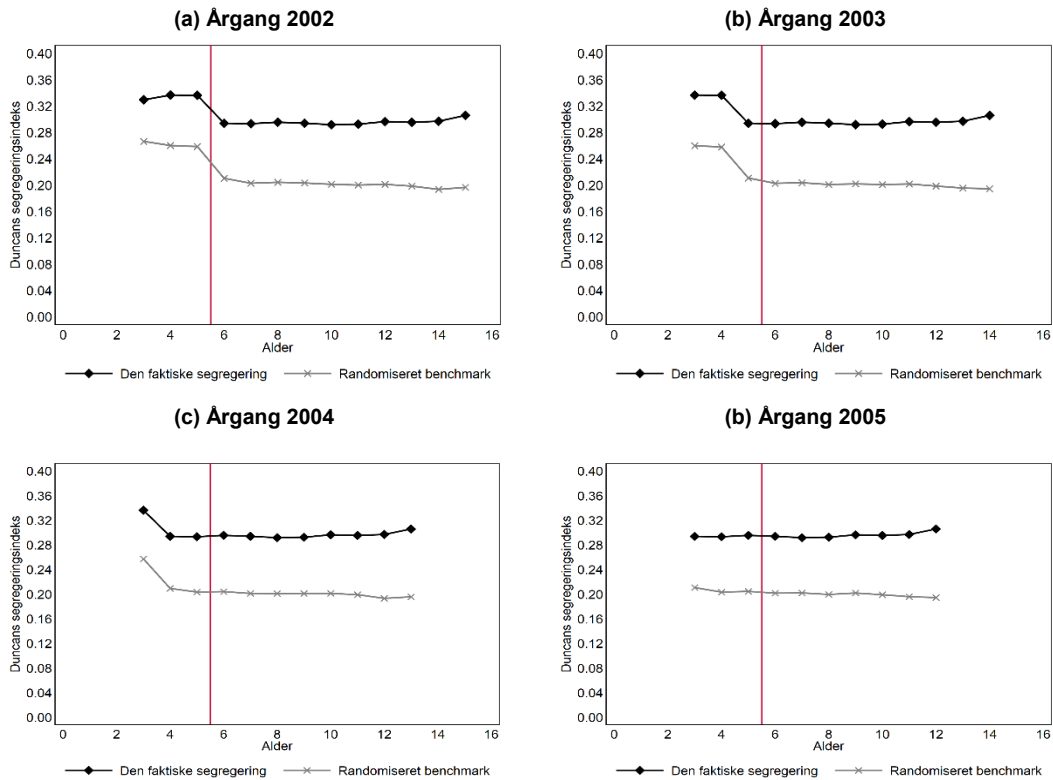
Segregering gennem barndommen, baseret på uddannelse og oprindelsesland

Segregering gennem barndommen baseret på forældrenes uddannelsesnivea

Vi har nu konkluderet, at segregeringen i danske børnehaver med hensyn til forældrenes uddannelsesnivea svarer nogenlunde til segregeringen med hensyn til økonomiske ressourcer. Det gælder uanset, hvordan vi måler segregeringen, og hvornår vi måler forældrenes uddannelse. Til sidst vil også belyse, om segregeringen igennem barndommen med hensyn til forældrenes uddannelsesnivea svarer til de resultater, vi så med hensyn til økonomiske ressourcer.

Bilagsfigur 4.4 viser segregeringen igennem barndommen for alle børn født i 2002-2005 baseret på Duncans segregationsindeks og forældrenes uddannelsesnivea. Betragter vi de sorte kurver i figuren, så ser vi, at segregeringen er forholdsvis konstant hen over barndommen. Der sker et lille udsving omkring skolestart, men det præcise tidspunkt for dette "udsving" ændrer sig med et par år fra kohorte til kohorte, hvilket igen kunne tyde på, at det er forårsaget af noget, der skete i et bestemt år og ikke, at det skyldes skolestarten.

Bilagsfigur 4.4 Segregeringen igennem barndommen målt ved Duncans segregationsindeks baseret på forældrenes gennemsnitlige uddannelsesnivea



Note.: Figurene viser Duncans segregeringsindeks for årene 2005-2017 for kohorterne 2002-2005 baseret på forældrenes uddannelsesnivea. De to grupper består alle af børn af forældre med et gennemsnitligt uddannelsesnivea henholdsvis over og under medianen. Forældrenes uddannelsesnivea er målt i oktober måned i det pågældende kalenderår. Segregeringen udregnes i forhold til alle andre børn i samme kohorte.

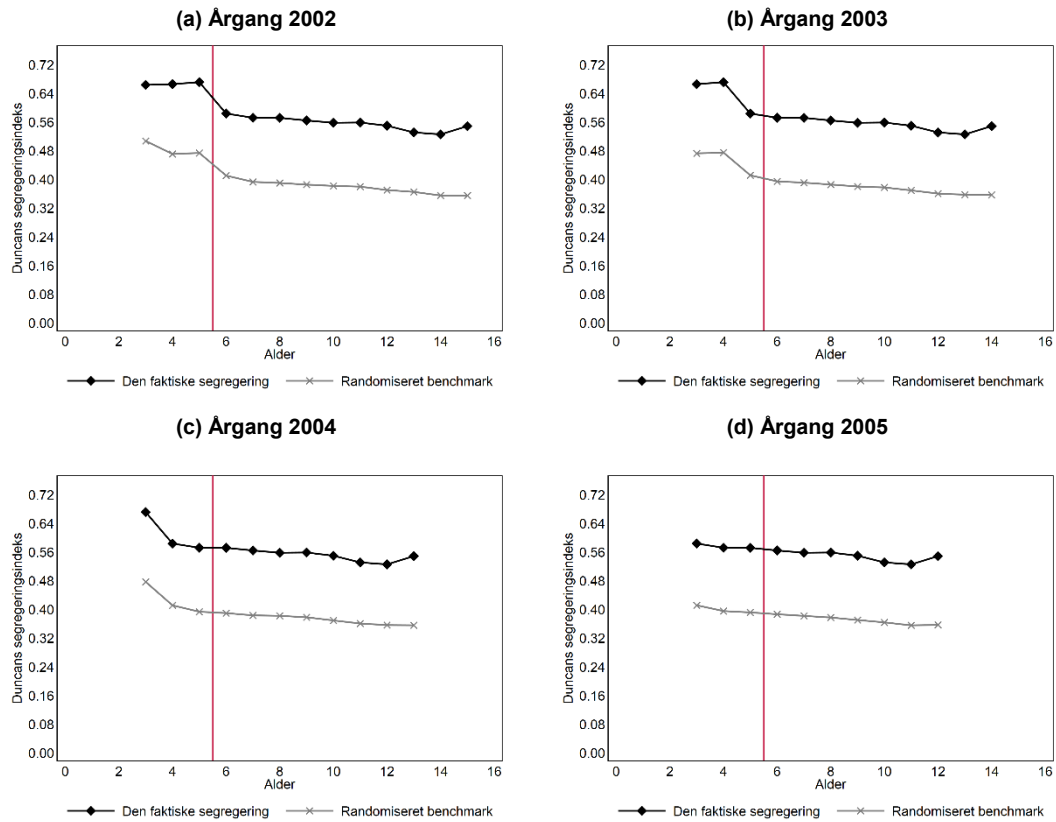
Det randomiserede sammenligningsgrundlag viser det gennemsnitlige segregeringsniveau på tværs af 50 tilfældige allokeringer af børn til børnehaver og grundskoler, hvor vi tager højde for barnets alder, bopælskommune, institutionstyper (for børnehaver) og institutionsstørrelse.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Segregering igennem barndommen baseret på oprindelsesland

Bilagsfigur 4.5 viser segregeringen gennem barndommen baseret på barnets oprindelse ved brug af Duncans segregationsindeks. Vi bruger her inddelingen af børn efter, om de er indvandrere eller efterkommere fra et ikke-vestligt land og alle andre. Resultaterne er meget i tråd med resultaterne for indkomst og uddannelse. Vi ser et lille knæk i kurven, som nærmere skyldes noget, der skete omkring 2008, end noget, der skete ved et specifikt alderstrin. Det er lidt overraskende, idet oprindelse ikke afhænger af konjunkturer. Det tyder på, at ændringen i 2008 ikke nødvendigvis er drevet af konjunkturer, men i stedet muligvis af institutionelle ændringer, som fx strukturreformen i 2007.

Bilagsfigur 4.5 Segregeringen gennem barndommen målt ved Duncans segregationsindeks baseret på barnets oprindelse



Note.: Figurerne viser Duncans segregeringsindeks for årene 2005-2017 for kohorterne 2002-2005 baseret på barnets oprindelse. De to grupper består af henholdsvis alle børn, som er indvandrere eller efterkommere fra et ikke-vestligt land, og alle andre børn. De vestlige lande er ifølge Danmarks Statistiks definition samtlige 28 EU-lande samt Andorra, Island, Liechtenstein, Monaco, Norge, San Marino, Schweiz, Vatikanstaten, Canada, USA, Australien og New Zealand. Segregeringen udregnes i forhold til alle andre børn i samme kohorte.

Det randomiserede sammenligningsgrundlag viser det gennemsnitlige segregeringsniveau på tværs af 50 tilfældige allokeringer af børn til børnehaver og grundskoler, hvor vi tager højde for barnets alder, bopælskommune, institutionstyper (for børnehaver) og institutionsstørrelse.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Bilag 5 Segregering med hensyn til indkomst

Bilagstabel 5.1 Segregeringen i danske børnehaver med hensyn til økonomiske ressourcer for årene 2005-2014

		2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Duncans segregationsindeks											
Ækvivaleret husstandsindkomst i samme år	F	0.32	0.32	0.31	0.31	0.31	0.32	0.32	0.32	0.32	0.32
	R	0.22	0.21	0.20	0.20	0.19	0.20	0.21	0.20	0.20	0.19
	D	0.10	0.10	0.11	0.11	0.12	0.12	0.12	0.12	0.12	0.13
Forældrenes indkomst i kalenderåret, før barnet blev født	F	0.29	0.29	0.28	0.29	0.29	0.29	0.28	0.27	0.27	0.28
	R	0.21	0.20	0.20	0.20	0.20	0.20	0.20	0.19	0.18	0.19
	D	0.07	0.08	0.08	0.09	0.09	0.09	0.09	0.09	0.10	0.09
Forældrenes indkomst i kalenderåret, hvor barnet fyldte 5 år	F	0.34	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.34
	R	0.25	0.23	0.22	0.22	0.21	0.22	0.22	0.21	0.21	0.21
	D	0.09	0.10	0.11	0.11	0.12	0.12	0.12	0.12	0.12	0.12
Husstandsindkomst under 60 % af medianindkomsten	F	0.31	0.32	0.31	0.31	0.31	0.34	0.34	0.34	0.35	0.35
	R	0.23	0.22	0.21	0.21	0.20	0.22	0.21	0.21	0.21	0.20
	D	0.08	0.09	0.10	0.10	0.11	0.12	0.13	0.13	0.14	0.14
R²-metoden											
Ækvivaleret husstandsindkomst i samme år	F	0.16	0.15	0.15	0.14	0.14	0.15	0.16	0.15	0.15	0.15
	R	0.08	0.07	0.07	0.07	0.06	0.07	0.07	0.07	0.07	0.06
	D	0.08	0.08	0.08	0.08	0.08	0.09	0.09	0.09	0.09	0.09
Forældrenes indkomst i kalenderåret, før barnet blev født	F	0.13	0.13	0.13	0.13	0.13	0.13	0.12	0.12	0.11	0.12
	R	0.07	0.07	0.07	0.06	0.06	0.06	0.06	0.06	0.05	0.06
	D	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
Forældrenes indkomst i kalenderåret, hvor barnet fyldte 5 år	F	0.17	0.17	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.17
	R	0.10	0.09	0.08	0.08	0.07	0.08	0.08	0.08	0.07	0.08
	D	0.07	0.08	0.08	0.08	0.09	0.09	0.09	0.09	0.09	0.09
Husstandsindkomst under 60 % af medianindkomsten	F	0.08	0.09	0.09	0.09	0.09	0.11	0.11	0.10	0.11	0.11
	R	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03
	D	0.04	0.05	0.05	0.05	0.05	0.07	0.07	0.07	0.07	0.08

Anm.: F: Det faktiske segregeringsniveau. R: Det randomiserede sammenligningsgrundlag (sammenligningsgrundlag). D: Forskellen mellem det faktiske segregeringsniveau og det randomiserede sammenligningsgrundlag. Indkomsten angiver den disponible indkomst efter skat og overførsler. Den ækvivalerede indkomst er den samlede indkomst divideret med kvadratroden af antal husstandsmedlemmer.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Bilagstabel 5.2 Segregeringen gennem barndommen med hensyn til økonomiske resourcer for kohorterne 2002-2005 baseret på ækvivaleret disponibel husstandsindkomst

		Fødselskohorte												
		2002			2003			2004			2005			
		F	R	D	F	R	D	F	R	D	F	R	D	
Duncans segregationsindeks														
Barnets alder (år)	3	0.37	0.29	0.08	0.36	0.26	0.09	0.36	0.26	0.10	0.31	0.21	0.09	
	4	0.36	0.26	0.09	0.36	0.26	0.10	0.31	0.21	0.09	0.30	0.20	0.10	
	5	0.36	0.26	0.10	0.31	0.21	0.09	0.30	0.20	0.10	0.33	0.21	0.11	
	6	0.31	0.21	0.09	0.30	0.20	0.10	0.33	0.21	0.11	0.32	0.21	0.11	
	7	0.30	0.20	0.10	0.33	0.21	0.12	0.32	0.21	0.11	0.32	0.21	0.11	
	8	0.33	0.21	0.11	0.32	0.21	0.11	0.32	0.21	0.11	0.32	0.21	0.11	
	9	0.32	0.21	0.11	0.32	0.21	0.11	0.32	0.21	0.11	0.32	0.21	0.11	
	10	0.32	0.21	0.11	0.32	0.21	0.11	0.32	0.21	0.11	0.32	0.20	0.11	
	11	0.32	0.21	0.11	0.32	0.21	0.11	0.32	0.20	0.11	0.31	0.20	0.11	
	12	0.32	0.21	0.11	0.32	0.20	0.11	0.31	0.20	0.11	0.32	0.20	0.12	
	13	0.32	0.20	0.11	0.31	0.20	0.11	0.32	0.20	0.12				
	14	0.31	0.20	0.11	0.32	0.20	0.12							
	15	0.32	0.20	0.12										
	R²-metoden													
	Barnets alder (år)	3	0.20	0.13	0.07	0.19	0.11	0.08	0.19	0.10	0.09	0.14	0.07	0.07
4		0.19	0.11	0.08	0.19	0.11	0.08	0.14	0.07	0.07	0.14	0.07	0.07	
5		0.19	0.11	0.09	0.14	0.07	0.07	0.14	0.06	0.07	0.16	0.07	0.08	
6		0.14	0.08	0.07	0.14	0.07	0.07	0.16	0.08	0.08	0.15	0.07	0.08	
7		0.14	0.06	0.07	0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	
8		0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.08	0.08	0.15	0.07	0.08	
9		0.15	0.07	0.08	0.15	0.08	0.08	0.15	0.07	0.08	0.16	0.07	0.08	
10		0.15	0.07	0.08	0.15	0.07	0.08	0.16	0.07	0.08	0.15	0.07	0.08	
11		0.15	0.07	0.08	0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	
12		0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.09	
13		0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.09				
14		0.15	0.07	0.08	0.15	0.06	0.09							
15		0.15	0.07	0.09										

Anm.: F: Det faktiske segregeringsniveau. R: Det randomiserede sammenligningsgrundlag (sammenligningsgrundlag). D: Forskellen mellem det faktiske segregeringsniveau og det randomiserede sammenligningsgrundlag. Indkomsten angiver den disponible indkomst efter skat og overførsler. Den ækvivalerede indkomst er den samlede indkomst divideret med kvadratroden af antal husstandsmedlemmer i det pågældende kalenderår.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Segregering med hensyn til uddannelse

Bilagstabel 5.3 Segregeringen i danske børnehaver med hensyn til forældrenes uddannelsesniveau

		2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Duncans segregationsindeks											
Forældrenes uddannelsesniveau i det pågældende år	F	0.28	0.29	0.30	0.30	0.30	0.30	0.30	0.31	0.31	0.32
	R	0.19	0.20	0.20	0.20	0.21	0.21	0.21	0.22	0.22	0.23
	D	0.08	0.09	0.09	0.10	0.09	0.09	0.09	0.09	0.09	0.09
Forældrenes uddannelsesniveau i kalenderåret, før barnet blev født	F	0.26	0.27	0.27	0.28	0.29	0.30	0.30	0.30	0.30	0.32
	R	0.19	0.18	0.18	0.19	0.20	0.21	0.21	0.21	0.22	0.22
	D	0.07	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
Forældrenes uddannelsesniveau i kalenderåret, hvor barnet fyldte 5 år	F	0.28	0.29	0.30	0.30	0.30	0.30	0.30	0.31	0.31	0.32
	R	0.20	0.20	0.20	0.21	0.21	0.21	0.21	0.22	0.22	0.23
	D	0.08	0.09	0.09	0.10	0.09	0.09	0.09	0.09	0.09	0.09
R²-metoden											
Forældrenes uddannelsesniveau i det pågældende år	D	0.12	0.13	0.13	0.14	0.14	0.14	0.14	0.14	0.14	0.15
	F	0.06	0.06	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.08
	R	0.06	0.06	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
Forældrenes uddannelsesniveau i kalenderåret, før barnet blev født	D	0.10	0.11	0.11	0.12	0.12	0.13	0.13	0.14	0.14	0.15
	F	0.06	0.05	0.05	0.06	0.06	0.07	0.07	0.07	0.07	0.08
	R	0.05	0.05	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.07
Forældrenes uddannelsesniveau i kalenderåret, hvor barnet fyldte 5 år	D	0.12	0.13	0.13	0.14	0.13	0.14	0.14	0.14	0.14	0.15
	F	0.07	0.06	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.08
	R	0.06	0.06	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07

Anm.: F: Det faktiske segregeringsniveau. R: Det randomiserede sammenligningsgrundlag (sammenligningsgrundlag). D: Forskellen mellem det faktiske segregeringsniveau og det randomiserede sammenligningsgrundlag. Uddannelsesniveauet er målt i antal år.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Bilagstabel 5.4 Segregeringen igennem barndommen med hensyn til forældrenes uddannelsesniveau i pågældende kalenderår for kohorterne 2002 til 2005

		Fødselskohorte												
		2002			2003			2004			2005			
		F	R	D	F	R	D	F	R	D	F	R	D	
Duncans segregationsindeks														
Barnets alder (år)	3	0.33	0.27	0.06	0.34	0.26	0.08	0.34	0.26	0.08	0.29	0.21	0.08	
	4	0.34	0.26	0.08	0.34	0.26	0.08	0.29	0.21	0.08	0.29	0.20	0.09	
	5	0.34	0.26	0.08	0.29	0.21	0.08	0.29	0.20	0.09	0.30	0.20	0.09	
	6	0.29	0.21	0.08	0.29	0.20	0.09	0.30	0.20	0.09	0.29	0.20	0.09	
	7	0.29	0.20	0.09	0.30	0.20	0.09	0.29	0.20	0.09	0.29	0.20	0.09	
	8	0.30	0.20	0.09	0.29	0.20	0.09	0.29	0.20	0.09	0.29	0.20	0.09	
	9	0.29	0.20	0.09	0.29	0.20	0.09	0.29	0.20	0.09	0.30	0.20	0.09	
	10	0.29	0.20	0.09	0.29	0.20	0.09	0.30	0.20	0.10	0.30	0.20	0.10	
	11	0.29	0.20	0.09	0.30	0.20	0.09	0.30	0.20	0.10	0.30	0.20	0.10	
	12	0.30	0.20	0.10	0.30	0.20	0.10	0.30	0.19	0.10	0.31	0.19	0.11	
	13	0.30	0.20	0.10	0.30	0.20	0.10	0.31	0.20	0.11				
	14	0.30	0.19	0.10	0.31	0.19	0.11							
	15	0.31	0.20	0.11										
	R²-metoden													
	Barnets alder (år)	3	0.20	0.13	0.07	0.19	0.11	0.08	0.19	0.11	0.08	0.19	0.11	0.08
4		0.19	0.11	0.08	0.19	0.11	0.09	0.14	0.07	0.07	0.14	0.07	0.07	
5		0.19	0.11	0.08	0.14	0.07	0.07	0.14	0.06	0.07	0.14	0.06	0.07	
6		0.14	0.07	0.07	0.14	0.06	0.07	0.16	0.07	0.08	0.16	0.07	0.08	
7		0.14	0.06	0.07	0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	
8		0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.08	0.08	0.15	0.08	0.08	
9		0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	
10		0.15	0.07	0.08	0.15	0.07	0.08	0.16	0.07	0.09	0.16	0.07	0.09	
11		0.15	0.07	0.08	0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	
12		0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	
13		0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.09				
14		0.15	0.07	0.08	0.15	0.07	0.09							
15		0.15	0.06	0.09										

Anm.: F: Det faktiske segregeringsniveau. R: Det randomiserede sammenligningsgrundlag (sammenligningsgrundlag). D: Forskellen mellem det faktiske segregeringsniveau og det randomiserede sammenligningsgrundlag. Uddannelsesniveaet er målt i antal år.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Segregering med hensyn til forældrenes oprindelse

Bilagstabel 5.5 Segregeringen i danske børnehaver med hensyn til forældrenes oprindelse

		2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Duncans segregationsindeks											
Efterkommere og indvandrere fra ikke-vestlige lande	F	0.62	0.62	0.61	0.61	0.60	0.60	0.58	0.57	0.55	0.54
	R	0.42	0.39	0.39	0.39	0.38	0.37	0.35	0.34	0.32	0.30
	D	0.20	0.23	0.23	0.22	0.23	0.23	0.23	0.23	0.23	0.24
Efterkommere eller indvandrere (uanset land)	F	0.58	0.58	0.57	0.56	0.55	0.54	0.52	0.51	0.49	0.48
	R	0.40	0.37	0.37	0.36	0.35	0.34	0.32	0.31	0.29	0.27
	D	0.18	0.21	0.21	0.20	0.20	0.20	0.20	0.20	0.21	0.21
R²-metoden											
Efterkommere og indvandrere fra ikke-vestlige lande	F	0.29	0.29	0.29	0.28	0.26	0.25	0.23	0.23	0.21	0.21
	R	0.09	0.08	0.07	0.07	0.07	0.06	0.06	0.05	0.05	0.05
	D	0.21	0.21	0.21	0.20	0.19	0.19	0.18	0.17	0.16	0.16
Efterkommere eller indvandrere (uanset land)	F	0.28	0.27	0.27	0.26	0.24	0.23	0.22	0.21	0.20	0.19
	R	0.09	0.08	0.08	0.07	0.07	0.06	0.06	0.05	0.05	0.05
	D	0.19	0.20	0.20	0.19	0.18	0.17	0.16	0.16	0.15	0.15

Anm.: F: Det faktiske segregeringsniveau. R: Det randomiserede sammenligningsgrundlag (sammenligningsgrundlag). D: Forskellen mellem det faktiske segregeringsniveau og det randomiserede sammenligningsgrundlag. De vestlige lande er ifølge Danmarks Statistiks definition samtlige 28 EU-lande samt Andorra, Island, Liechtenstein, Monaco, Norge, San Marino, Schweiz, Vatikanstaten, Canada, USA, Australien og New Zealand.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

Bilagstabel 5.6 Segregeringen gennem barndommen med hensyn til forældrenes oprindelse for kohorterne 2002 til 2005

		Fødselskohorte												
		2002			2003			2004			2005			
		F	R	D	F	R	D	F	R	D	F	R	D	
Duncans segregationsindeks														
Barnets alder (år)	3	0.66	0.51	0.16	0.67	0.47	0.19	0.67	0.48	0.19	0.58	0.41	0.17	
	4	0.67	0.47	0.19	0.67	0.47	0.20	0.58	0.41	0.17	0.57	0.40	0.18	
	5	0.67	0.47	0.20	0.58	0.41	0.17	0.57	0.39	0.18	0.57	0.39	0.18	
	6	0.58	0.41	0.17	0.57	0.39	0.18	0.57	0.39	0.18	0.56	0.39	0.18	
	7	0.57	0.39	0.18	0.57	0.39	0.18	0.56	0.38	0.18	0.56	0.38	0.18	
	8	0.57	0.39	0.18	0.56	0.39	0.18	0.56	0.38	0.18	0.56	0.38	0.18	
	9	0.56	0.39	0.18	0.56	0.38	0.18	0.56	0.38	0.18	0.55	0.37	0.18	
	10	0.56	0.38	0.18	0.56	0.38	0.18	0.55	0.37	0.18	0.53	0.36	0.17	
	11	0.56	0.38	0.18	0.55	0.37	0.18	0.53	0.36	0.17	0.53	0.36	0.17	
	12	0.55	0.37	0.18	0.53	0.36	0.17	0.53	0.36	0.17	0.55	0.36	0.19	
	13	0.53	0.36	0.17	0.53	0.36	0.17	0.55	0.36	0.19				
	14	0.53	0.35	0.17	0.55	0.36	0.19							
	15	0.55	0.35	0.19										
	R2-metoden													
	Barnets alder (år)	3	0.20	0.13	0.07	0.19	0.11	0.08	0.19	0.11	0.08	0.14	0.07	0.07
4		0.19	0.11	0.08	0.19	0.11	0.08	0.14	0.07	0.07	0.14	0.07	0.07	
5		0.19	0.11	0.08	0.14	0.07	0.07	0.14	0.07	0.07	0.16	0.07	0.08	
6		0.14	0.07	0.07	0.14	0.06	0.07	0.16	0.07	0.08	0.15	0.07	0.08	
7		0.14	0.07	0.07	0.16	0.08	0.08	0.15	0.07	0.08	0.15	0.07	0.08	
8		0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	
9		0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	0.16	0.07	0.08	
10		0.15	0.07	0.08	0.15	0.07	0.08	0.16	0.07	0.08	0.15	0.07	0.08	
11		0.15	0.07	0.08	0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	
12		0.16	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.09	
13		0.15	0.07	0.08	0.15	0.07	0.08	0.15	0.07	0.09				
14		0.15	0.07	0.08	0.15	0.07	0.08							
15		0.15	0.07	0.09										

Anm.: F: Det faktiske segregeringsniveau. R: Det randomiserede sammenligningsgrundlag (sammenligningsgrundlag). D: Forskellen mellem det faktiske segregeringsniveau og det randomiserede sammenligningsgrundlag. De vestlige lande er ifølge Danmarks Statistiks definition samtlige 28 EU-lande samt Andorra, Island, Liechtenstein, Monaco, Norge, San Marino, Schweiz, Vatikanstaten, Canada, USA, Australien og New Zealand.

Kilde: Egne beregninger baseret på registerdata fra Danmarks Statistik.

VIDEN
VELFÆRD

DET NATIONALE FORSKNINGS-
OG ANALYSECENTER FOR VELFÆRD