

DREAM

Udsatte boligområder i Danmark

En historisk analyse af socioøkonomisk
segregering, flyttemønstre, indkomstforløb og
effekt af bosætning i udsatte boligområder

Rapport udarbejdet af DREAM-gruppen
for Boligøkonomisk Videncenter

af Andreas Østergaard Iversen,
Jonas Zangenberg Hansen,
Marianne Frank Hansen og
Peter Stephensen

Juni 2019

Forord

I rapporten identificeres udsatte boligområder i Danmark gennem perioden 1986-2017 og beboernes karakteristika og adfærdsmønstre analyseres.

Udpegelsen af udsatte boligområder udføres ved anvendelse af en til formålet udviklet metode. Metoden er automatiseret, hvorfor den enkelt kan anvendes til identifikation over længere tidsperioder, da subjektive vurderinger overflødiggøres. Afgrænsningen af de udsatte områder er baseret på de socioøkonomiske karakteristika, der kendetegner beboerne i enhver boligs nabolag. En boligs nabolag defineres ved de 1.000 personer, der er bosat tættest på den betragtede bolig. For hvert nabolag beregnes gennemsnitlige værdier for indkomst, arbejdsmarkedstilknytning, uddannelsesniveau og kriminalitetsomfang. Disse fire karakteristika sammenvejes i ét socioøkonomisk indeks, som for hver enkelt bolig afslører, hvor udsat et nabolag boligen er beliggende i. Opstår der sammenklumpning af boliger, der er udpeget som beliggende i de mest socioøkonomisk udsatte nabolag, er der tale om et udsat boligområde. Med udsatte områder menes herved de områder, hvor andelen af beboere med lav socioøkonomisk status er højest.

Socioøkonomiske karakteristika for befolkningen bosat i de udsatte områder beskrives for perioden 1986-2017 og for seks udvalgte kommuner ses der nærmere på, hvorledes afgrænsningen af de udsatte områder har ændret sig over tid.

Flytninger kan over tid være med til at fastholde eller ændre strukturen i et boligområde, hvorfor rapporten analyserer flyttemønstre i de identificerede udsatte boligområder.

En persons socioøkonomiske formåen kan sammenfattes i udviklingen i den pågældendes indkomst gennem livet. Det analyseres, hvorvidt indkomstforløbene for personer med ophold i et udsat område adskiller sig systematisk fra den øvrige befolknings.

Endeligt undersøges det, hvorledes det påvirker flygtnings erhvervsdeltagelse og kriminalitetsomfang at bo i udsatte boligområder. Det afsøges, hvorvidt konklusionerne kan generaliseres til en større befolkningsgruppe.

Undersøgelsen er bestilt af Boligøkonomisk Videncenter med finansiering fra Realdania. Undersøgelsen er udført af DREAM-gruppen ved Andreas Østergaard Iversen, Jonas Zangenberg Hansen, Marianne Frank Hansen og Peter Stephensen. Forskningsassistent Maria Andreasen har bistået med data- og beregningsopgaver.

DREAM har specificeret alle forudsætninger og beregningsmetoder anvendt i rapporten. I tabeller kan afrunding medføre, at tallene ikke summer til totalen.

København, juni 2019

Udsatte boligområder i Danmark

En historisk analyse af socioøkonomisk segregering, flyttemønstre, indkomstforløb og effekt af bosætning i udsatte boligområder

Indholdsfortegnelse

1. Indledning	3
2. Udsatte boligområder i Danmark	5
2.1 Socioøkonomisk Nabolag-Indeks.....	6
2.2 Geografiske afgrænsninger af udsatte boligområder i Danmark	12
2.3 Udsatte boligområders beboere i Danmark	14
2.4 Udsatte boligområder i seks udvalgte kommuner.....	17
Udsatte boligområder i København.....	18
Udsatte boligområder i Århus.....	22
Udsatte boligområder i Odense.....	26
Udsatte boligområder i Esbjerg	30
Udsatte boligområder i Brøndby og Ishøj.....	33
3. Flyttemønstre i varigt udsatte almene boligområder	37
3.1 Afgrænsning til varigt udsatte almene boligområder	38
3.2 Boligmobilitet.....	40
3.3 Flyttemønstre i varigt udsatte almene boligområder.....	43
3.4 Indkomstudvikling omkring flyttetidspunkt.....	54
3.5 Flyttemønstre og segregeringsgrad	60
4. Individuelle indkomstforløb	63
4.1 Indkomstforløb og baggrundskarakteristika	65
4.2 Gruppering af individuelle indkomstforløb.....	79
4.3 Identifikation af signifikante relationer mellem baggrundskarakteristika og indkomstforløb.....	100
5. Effekt af bosætning i udsatte boligområder: Et quasi-randomiseret eksperiment	106
5.1 Det naturlige eksperiment: Geografisk fordeling af flygtninge	107
5.2 Datagrundlag og balancering af data	108
5.3 Effekt af placering i udsat boligområde.....	115
5.4 Generalisering af resultatet af indsatsanalysen til andre befolkningsgrupper	122
6. Clusted Nearest Neighbor algoritmen	131
6.1 Nabolag og socioøkonomiske variable	131
6.2 SNI via principalkomponentanalyse	132
6.3 Hierarkisk clustering analyse	136
7. Konklusion	137
8. Referencer	140

A.	SNI for udvalgte kommuner	143
B.	Kullback–Leibler divergensindekset	148
C.	Et eksempel på hierarkisk clustering processen	149
D.	Gruppering af indkomstforløb og analyse af individuelle indkomstforløb.....	152
E.	Statistisk analyse af gruppetilhørsforhold.....	155
F.	Balancering af data ved propensity score matching	156
G.	Rubins kausale model.....	158
H.	Betinget effekt på erhvervsfrekvens af placering i udsat område.....	159
I.	Anvendelse af propensity scores til fastlæggelse af gennemsnitlig indsats effekt for målgruppen	164
J.	Undersøgelsens datagrundlag	165

1. Indledning

Formålet med denne undersøgelse er at kortlægge hvor i Danmark, der er udsatte boligområder og at undersøge, hvad der karakteriserer disse områders beboere.

Analysen indledes med at identificere udsatte boligområder i Danmark i perioden 1986-2017. Der er udviklet en metode, der gør det muligt på en konsistent og automatisk måde at udpege udsatte boligområder over længere tidsperioder. Områderne udpeges på basis af socioøkonomiske karakteristika, og etnicitet er bevidst fravalgt som identificerende parameter.

Hovedparten af de udsatte boligområder identificeres til at være større almene bebyggelser, der enten er beliggende centralt i de større byer eller i forstæderne til disse byområder. Områderne har over den betragtede periode været tilnærmelsesvis stabile i både udbredelse og placering. Væsentlige undtagelser er dog København og Odense. Placeringen af udsatte boligområder har i København gennemgået en betydelig forandring, hvilket blandt andet forklares af store byfornyelsesprojekter. En anden undtagelse er det store almene boligområde Vollsmose i Odense, som i midt 80'erne ikke bliver identificeret som et udsat boligområde.

På basis af udsatte nabolag søges graden af socioøkonomisk segregering i Danmark belyst. Det demonstreres, at beboerne i de udsatte nabolag i Danmark lever langt mere segregeret end den socioøkonomiske mellemklasse. Der argumenteres for, at den socioøkonomiske segregering i de udsatte boligområder, over den betragtede periode, har været faldende i København, mens den har været tilnærmelsesvis stabil i resten af landet.

Kendetegnende for beboere i de udsatte områder er, at de har en lavere indkomst, som udgør cirka to tredjedele af indkomsten for personer bosat i ikke-udsatte områder. Desuden har beboerne i udsatte områder lavere arbejdsmarkedstilknytning, ringere uddannelsesniveau og er yngre end befolkningen generelt. Det trækker også mod en lavere gennemsnitlig indkomst. Andelen af beboere i udsatte boligområder, der har været dømt for forseelser udover færdselsloven, er omkring tre-fire gange så stor som i befolkningen som helhed. Andelen af ikke-vestlige indvandrere og efterkommere i udsatte boligområder har været stigende over perioden og tilvæksten overgår markant den generelle stigning på landsplan.

Desuden betragtes til- og fraflytningsmønstre i de udsatte boligområder. Flytninger kan over tid være med til at fastholde eller ændre strukturen i et boligområde. Det findes, at tilflyttere i gennemsnit er dårligere stillet end områdets eksisterende beboere, hvorfor flyttemønstret kan være medvirkende til, at området forbliver udsat. Desuden ses der en indkomstudvikling umiddelbart efter til- og fraflytning til de udsatte områder. Der ses en tendens til, at tilflyttere relativt hurtigt konvergerer mod indkomstniveauet for eksisterende beboere i området. Ved fraflytning ses fastboende og fraflyttere at have samme gennemsnitlige indkomst. Dog ses en tilbøjelighed til, at fraflyttere i årene efter at have forladt området har stærkere indkomstfremgang end de beboere, som bliver boende i området. Dette kan indikere, at det er relativt ressourcerstærke beboere, som fraflytter udsatte boligområder.

En persons socioøkonomiske formåen kan sammenfattes i udviklingen i den pågældendes indkomst gennem livet. Gennem en såkaldt sekvensanalyse afklares det, hvorvidt det er muligt at gruppere indkomstforløb, der minder om hinanden og efterfølgende identificere hvilke baggrundskarakteristika, herunder bopæls erfaring fra et udsat område, der kendetegner personerne i de enkelte grupper. Kan det eksempelvis påvises, at personer bosat i et udsat område systematisk kategoriseres i en særskilt gruppe, kan det sluttes, at individer med ophold i et udsat område har indkomstforløb, der afviger fra den øvrige befolknings. Er personer fra et udsat område omvendt placeret i flere forskellige

indkomstgrupper, kan det indikere, at områdestatus ikke i sig selv er karakteriserende for en persons indkomstforløb. Undersøgelsens resultater peger mod, at det er socioøkonomiske forhold som uddannelse og arbejdsmarkedstilknytning, der kan relateres til forskelle i indkomstforløb frem for bosætning i et udsat område. Det kan ikke entydigt påvises, at indkomstforløbene for personer med ophold i et udsat område adskiller sig systematisk fra den øvrige befolknings. Eftersom der er tale om en gennemsnitsbetragtning for befolkningen som helhed, kan det endvidere ikke udelukkes, at effekten af bosættelse i et udsat område er heterogen og dermed varierer på tværs af undergrupper af befolkningen afgrænset ved eksempelvis oprindelse, uddannelse eller arbejdsmarkedstilknytning.

Endeligt undersøges det, hvorledes det påvirker mennesker at bo i udsatte boligområder. En sådan undersøgelse er i udgangspunktet udsat for problemstillinger omkring selvselektion. Har man en lav indkomst, fordi man bor i et udsat område, eller er det omvendt? I denne delanalyse afhjælpes selvselektionsproblemet med et såkaldt naturligt eksperiment. I perioden 1986-1998 stod Dansk Flygtningehjælp for at fordele nyankomne flygtninge ud over Danmark. Denne fordelingspolitik var tilnærmelsesvis tilfældig og skabte grundlag for et naturligt eksperiment i perioden. Det er muligt at påvise betydelige effekter af bosætning i et udsat område på arbejdsmarkedstilknytning, mens der ingen overordnede effekter findes på kriminalitetsomfanget. Det undersøges afslutningsvist om resultaterne af denne analyse kan generaliseres til en større befolkningsgruppe. Det viser sig, at indsats-effekten også kan påvises for alle flygtninge og familiesammenførte indvandret op til i dag, mens analysen ikke kan generaliseres til befolkningen bestående af alle ikke-vestlige indvandrere.

Udsatte boligområder og socioøkonomisk segregeringsniveau præsenteres i afsnit 2, mens en teknisk beskrivelse af den nyudviklede algoritme til udpegelse af udsatte boligområder kan ses i afsnit 6. Flyttemønstre beskrives i afsnit 3, mens indkomstforløb analyseres i afsnit 4. I afsnit 5 afsøges forholdet mellem bosætning i et udsat område og erhvervsdeltagelsen og kriminalitetsomfanget på kort, mellemlangt og langt sigt. Rapporten afsluttes med en konklusion i afsnit 7.

2. Udsatte boligområder i Danmark

Et centralt element i nærværende analyse er at udpege udsatte boligområder. Der er udviklet en metode, der gør det muligt på en konsistent og automatisk måde at udpege udsatte boligområder over længere tidsperioder. I denne analyse betragtes perioden 1986 til 2017. Områderne udpeges på basis af socioøkonomiske karakteristika. Etnicitet er således ikke en parameter.

Analysen viser, at der har været en vis stabilitet i udbredelsen og placeringen af udsatte boligområder i Danmark. Dog med væsentlige udtagelser. København har gennemgået en betydelig forandring i løbet af de betragtede 32 år. De udsatte boligområder har haft en tendens til at flytte bort fra centrum, herunder Vesterbro og Indre Nørrebro, ud til områder omkring Husum-Tingbjerg, Brønshøj og Utterslev. At Indre Nørrebro og Vesterbro har ændret status forklares blandt andet af store byfornyelsesprojekter op igennem 90'erne og start 00'erne. En anden undtagelse er det store almene boligområde Vollsmose i Odense, som i midt 80'erne ikke bliver identificeret som et udsat boligområde.

Et interessant biprodukt af den udviklede metode er, at vi kan udtale os om udviklingen i den socioøkonomiske segregering i Danmark. Der argumenteres for, at den socioøkonomiske segregering har været faldende i København, men har været nogenlunde stabil i resten af landet.

Med udsatte områder menes de områder, hvor andelen af beboere med lav socioøkonomisk status er højest. Vi måler socioøkonomisk status ud fra indkomst, uddannelsesniveau, arbejdsmarkedstilknytning og kriminalitet. Som nævnt anvendes herkomst ikke som parameter.

Transport-, Bygnings- og Boligministeriet (TBB) har i 2008 – med baggrund i Programbestyrelsen mod ghettoiserings¹ anbefalede områder – udviklet en metode til at udpege udsatte almene boligområder, jf. Transport-, Bygnings- og Boligministeriet (2018). Metoden præsenteret i den nærværende analyse adskiller sig på flere punkter. For det første er etnicitet, som nævnt, ikke et kriterie, når segregerede boligområder identificeres. For det andet er vores metode udelukkende baseret på dataanalyse således, at der altid anvendes de samme kriterier. Som følge heraf er der ikke behov for at udføre områdespecifikke undersøgelser som for eksempel fysiske besøg af boligområder.

Metoden tager udgangspunkt i definitionen af *nebolag*. For en given bolig er boligens individuelle nabolag defineret ved de nærmeste 1.000 personer således, at enhver bolig har sit eget nabolag. Bemærk specielt, at personer i den samme opgang vil have samme nabolag. Vi kan give en socioøkonomisk karakterisering af dette nabolag ved at se på disse 1.000 personers socioøkonomiske karakteristika. Bemærk, at vi herved *ikke* giver en karakterisering af de personer, der bor i ejendommen, men derimod det nabolag ejendommen er beliggende i.

Det enkelte nabolag karakteriseres ud fra følgende socioøkonomiske variable:

- Uddannelses- og arbejdsmarkedstilknytning, som er defineret som andelen af beboere i et nabolag, der er i den arbejdsdygtige alder (18-64 år) uden tilknytning til uddannelsessystemet eller arbejdsmarkedet.

¹ Programbestyrelsen mod ghettoisering blev nedsat af den daværende VK-regering i 2004 og kom med sine anbefalinger i november 2008.

- Kriminalitet, som er defineret som andelen af beboere i et nabolag, der er mindst 15 år gamle, som er dømt for overtrædelse af straffeloven, våbenloven eller loven om euforiserende stoffer indenfor det seneste kalenderår.
- Uddannelsesniveau, der er defineret som andelen af beboere i et nabolag i alderen 30-59 år, der har grundskole som højst fuldførte uddannelsesniveau².
- Indkomst, her defineret som nabolagets gennemsnitlige indkomst for beboere i aldersintervallet 15-64 år eksklusiv uddannelsessøgende sat i forhold til den gennemsnitlige indkomst for samme gruppe i hele landet.

Det vurderes, at kombinationen af disse fire socioøkonomiske variable er dækkende og beskrivende for især udsatte boligområder. Det er vigtigt at understrege, at der i begrebet 'udsatte boligområder' ikke ligger en vurdering af områderne eller disse områders beboere. Begrebet er udelukkende ment som et fagudtryk, der er defineret ud fra objektive socioøkonomiske egenskaber. Ud fra de fire variable defineres et såkaldt Socioøkonomisk Nabolag-Indeks (SNI, defineres nedenfor i afsnit 2.1), som gør det muligt at give en socioøkonomisk rangering af alle ejendomme i Danmark baseret på karakteristika for beboernes nabolag.

Vi vælger de fem procent, der bor i ejendomme med lavest socioøkonomisk status. Dette er muligt, fordi vi har defineret SNI. Disse mennesker bor potentielt i et udsat område. Vi ønsker imidlertid kun at betragte sammenhængende områder over en vis størrelse. Vi udfører derfor en såkaldt clusteranalyse, der beskrives detaljeret i afsnit 6.3. Denne clusteranalyse fordeler de betragtede ejendomme i sammenhængende områder. Til sidst ser vi bort fra alle områder, hvor der bor færre end 1.000 personer. De tilbageværende områder benævnes udsatte områder, jf. afsnit 2.2.

2.1 Socioøkonomisk Nabolag-Indeks

Som første led i udpegningen af udsatte boligområder ønsker vi at rangere alle ejendomme i Danmark efter de socioøkonomiske karakteristika for det nabolag, den enkelte ejendom ligger i (dvs. de socioøkonomiske karakteristika for de nærmeste 1.000 personer).

Som nævnt ovenfor måler vi socioøkonomisk status ud fra fire variable: arbejdsmarkedstilknytning, indkomst, uddannelsesniveau og kriminalitet. Vi samler disse fire variable i et indeks: det Socioøkonomiske Nabolag-Indeks (SNI). Dette er muligt via en såkaldt principal komponentanalyse. Denne metode udtrækker essensen af de fire variable ved at analysere den indbyrdes sammenhæng. Konstruktionen af SNI og de unikke nabolag er defineret og beskrevet i afsnit 6.1 og 6.2.

Det Socioøkonomiske Nabolag-Indeks er konstrueret således, at udsatte nabolag har lave værdier, mens de bedre stillede nabolag har høje værdier. Med andre ord vil nabolag have en lav SNI-værdi, hvis nabolaget har en stor andel af beboere, der er dømte, udenfor arbejdsmarkedet, har lav gennemsnitlig indkomst eller hvis nabolaget har en stor andel med blot en grundskoleuddannelse. Den socioøkonomiske mellemklasse har en SNI-værdi tæt på nul.

Segregering betegner en overrepræsentation eller koncentration af forskellige persongrupper i bestemte områder. SNI kan ses som et relativt socioøkonomisk segregeringsindeks. De

² Uddannelsesniveau måles typisk ud fra begrebet "højst fuldførte uddannelse". Det indebærer en rangordning af uddannelser på en skala, hvor grundskolen rangerer nederst og en Ph.d. grad øverst. Den af en persons fuldførte uddannelser, som rangerer højest, er udslagsgivende for personens "niveau".

individer, der bor i nabolag med høj negativ SNI-værdi, er de individer, der har den højeste sandsynlighed for lokalt at møde personer med lav gennemsnitlig indkomst, lavt uddannelsesniveau, ringe arbejdsmarkedstilknytning og en høj andel af dømte. Disse individer lever således relativt mest segregeret.

Figur 2.1 viser den geografiske placering og fordelingen af det Socioøkonomiske Nabolag-Indeks i 2017. I kortet er indekset opdelt i 20 grupper (ventiler) hver med fem procent af befolkningen. Farvekodningen går fra rød til mørk lilla, hvor farven rød er de mest udsatte nabolag og mørk lilla er de bedst stillede nabolag.

I Figur 2.2 fremhæves den geografiske placering af de udsatte (1. ventil), mellemklassen (defineret til 10. ventil) og de bedst stillede (20. ventil) socioøkonomiske nabolag. Det er vigtigt at understrege, at det er det samme antal personer (cirka 283.000 personer), der er afbilledet på hvert af de tre kort.

Den socioøkonomiske mellemklasse har nabolag med en indekxsværdi tæt på nul og er per konstruktion tæt på landsgennemsnittet. Disse nabolag er hverken socioøkonomisk udsatte eller bedrestillede. Fra Figur 2.1 og Figur 2.2 viser der sig nogle ret klare overordnede tendenser. Den socioøkonomiske mellemklasse bor over hele Danmark, dog med svag repræsentation på Lolland (grønlig nuance i figurerne). Samtidig ses der en tendens til, at de mest udsatte nabolag og de bedst stillede nabolag klumper sig sammen tæt på de store byer (rød og mørk lilla).

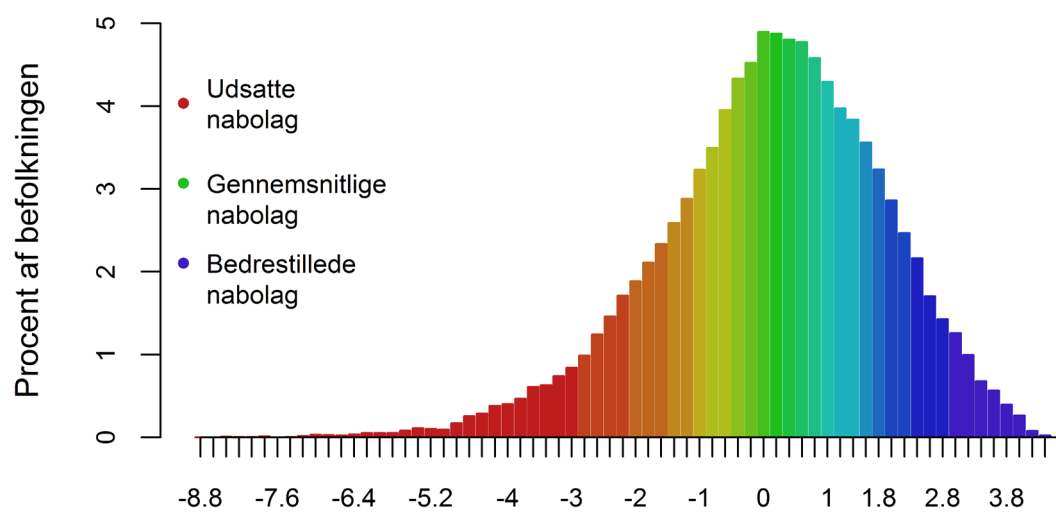
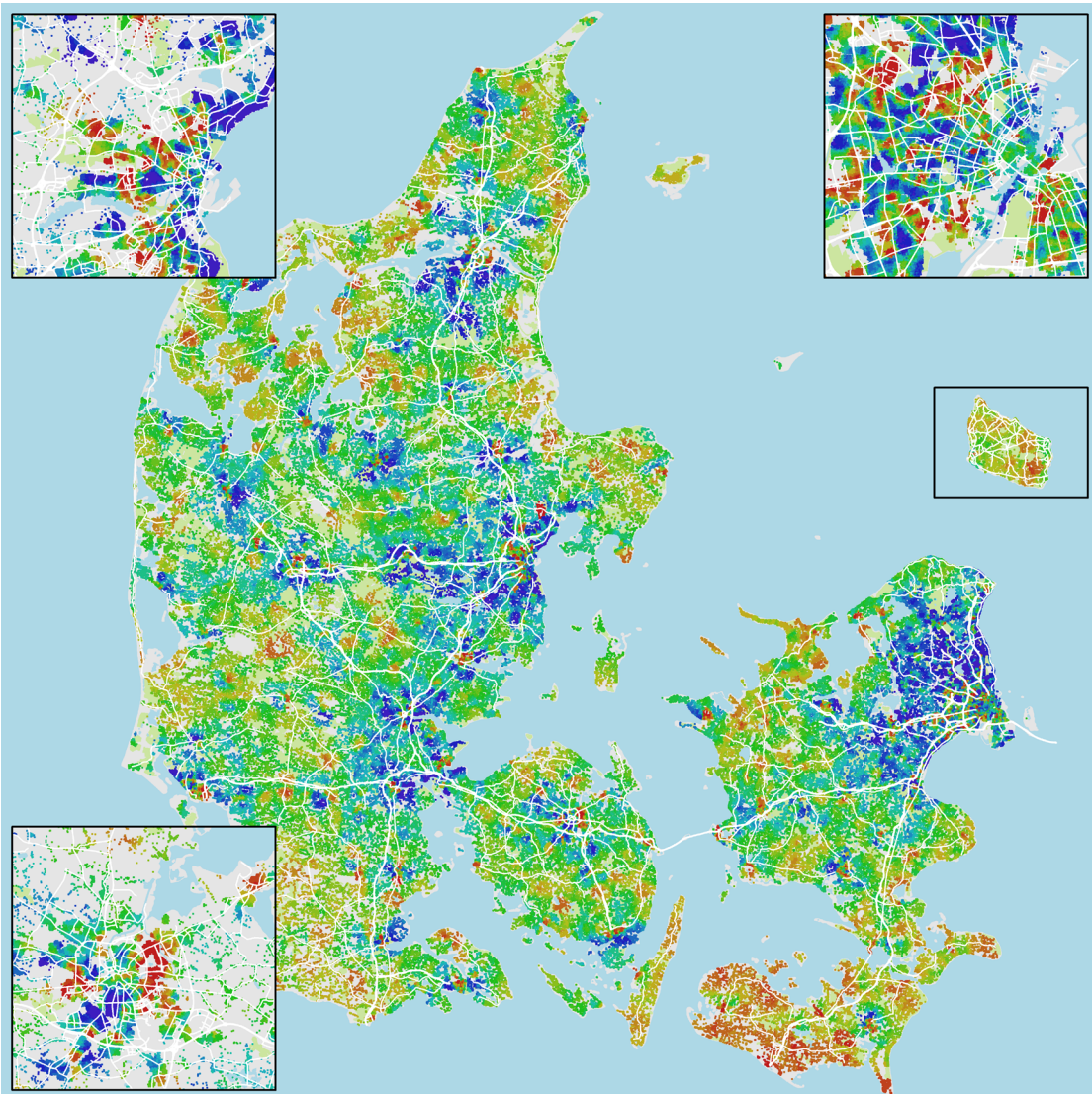
Socioøkonomisk segregering

Segregering betegner en overrepræsentation eller koncentration af forskellige persongrupper i bestemte områder, jf. Ministeriet for By, Bolig og Landdistrikter (2014). Vi interesser os her for socioøkonomisk segregering og vælger derfor at betragte segregering målt ved hjælp af SNI.

Nabolag med ekstreme indekxsværdier er per konstruktion de nabolag, der adskiller sig mest fra landsgennemsnittet med hensyn til de socioøkonomiske karakteristika. Fra Figur 2.2 ses det som nævnt, at disse nabolag klumper sig sammen omkring de større byer. Det tyder derfor på, at både bedrestillede og udsatte nabolag er segregerede i forhold til landsgennemsnitlige nabolag. Der ses en tendens til, at de udsatte nabolag er mest segregerede. De udsatte nabolag er beliggende centralt i de større byer, mens de bedrestillede nabolag er beliggende i oplandet til de større byer. Det ses, at mange små provinsbyer også befinder sig i den udsatte side af fordelingen (den orange nuance). Lolland er helt specielt. Her ses et stort område mørk orange, hvorfor der i denne del af landet er mange udsatte nabolag.

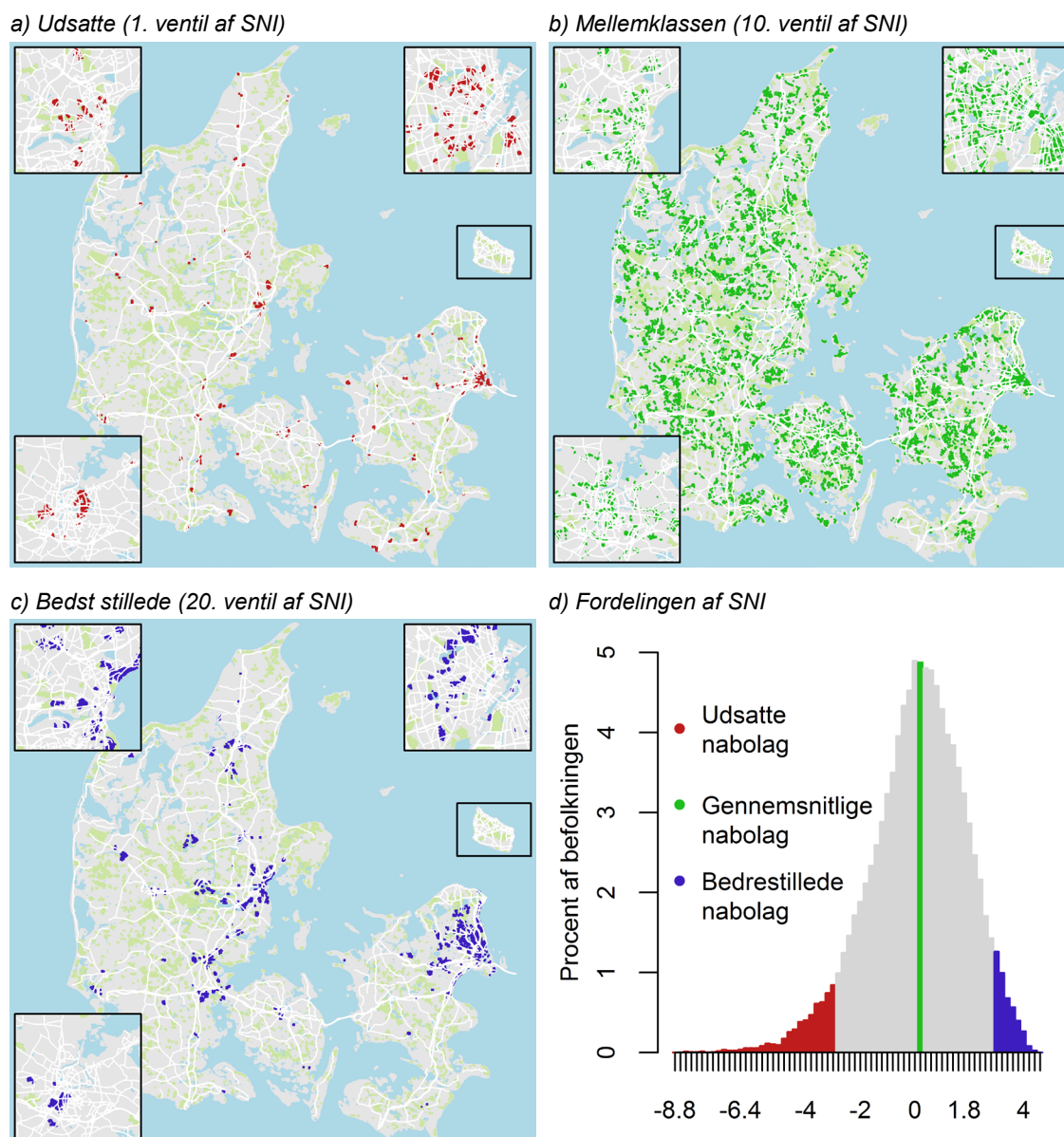
Kort indeholdende SNI for seks udvalgte kommuner i form af København, Århus, Odense, Esbjerg, Brøndby og Ishøj er vist i Appendiks A.

Figur 2.1. Det Socioøkonomiske Nabolag-Indeks, 2017



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 2.2. Placering af udsatte, gennemsnitlige og bedrestillede nabolag defineret ud fra SNI, 2017



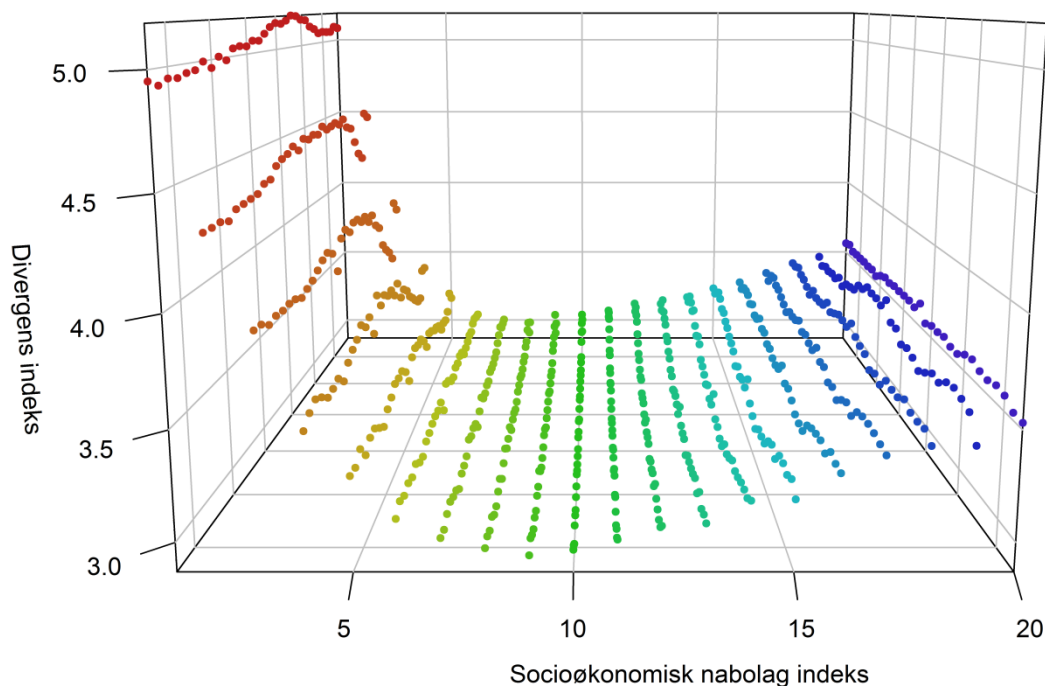
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Ud fra et segregeringssynspunkt er det et problem, at SNI er relativt. Indeksets faktiske værdier kan ikke sammenlignes over tid, og på et givet tidspunkt fås der ikke et mål for det absolutte segregerings-niveau. For at muliggøre dette, opstilles et mål for geografisk koncentration (et såkaldt Kullback–Leibler divergensindeks, jf. Kullback (1987)), der netop kvantificerer, hvor sammenklynget nabolagene er indenfor de forskellige ventilgrupper af det Socioøkonomiske Nabolag-Indeks over tid. Indekset måler for enhver ventilgruppe "afstanden" mellem den faktiske sandsynlighedsfordeling af nabolagsplacering og en sandsynlighedsfordeling, hvor alle nabolag er tilfældigt fordelt udover Danmark. Indekset er nærmere beskrevet i Appendiks B.

Figur 2.3 viser divergensindekset for alle ventiler af det Socioøkonomiske Nabolag-Indeks over perioden 1986 til 2017. 2017 er de nærmeste punkter i figuren, mens 1986 er de punkter, der ligger længst væk. Farvekoden i figuren følger SNI fordelingen i Figur 2.1.

Hvis indekset er nul, er de to sandsynlighedsfordelinger ens. Dette betyder, at nabolagene i den pågældende SNI ventil er spatialt tilfældigt fordelt ud over Danmark, og der er derfor ingen tegn på, at disse nabolag er klynget sammen. Jo større værdi indekset har, desto mere sammenklyngede er de pågældende nabolag, og jo større er graden af segregering.

Figur 2.3. Geografisk koncentration.



Anm.: Punkterne tættest på angiver divergensindekset i 2017, mens punkter længst væk angiver divergensindekset i 1986.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Fra Figur 2.3 ses samme tendens som Figur 2.2 indikerede: De mest udsatte nabolag er mere sammenklyngede relativt til den socioøkonomiske mellemklasse, det vil sige nabolag omkring ventil 10. Det ses også, at de bedst stillede nabolag også har en tendens til at være klynget tæt sammen relativt til den socioøkonomiske mellemklasse. Dette kan tolkes som, at de mest socioøkonomiske udsatte områder og de socioøkonomiske bedst stillede områder er mere segregerede i forhold til de gennemsnitlige områder. Men det ses også, at de mest udsatte områder er mere segregerede end de bedst stillede områder. Over tid ses denne tendens at være tilnærmelsesvis stabil for Danmark som helhed. Bemærk dog, at de udsatte områder ser ud til at være en smule mere segregerede i forhold til den socioøkonomiske mellemklasse i slut 90'erne og start 00'erne.

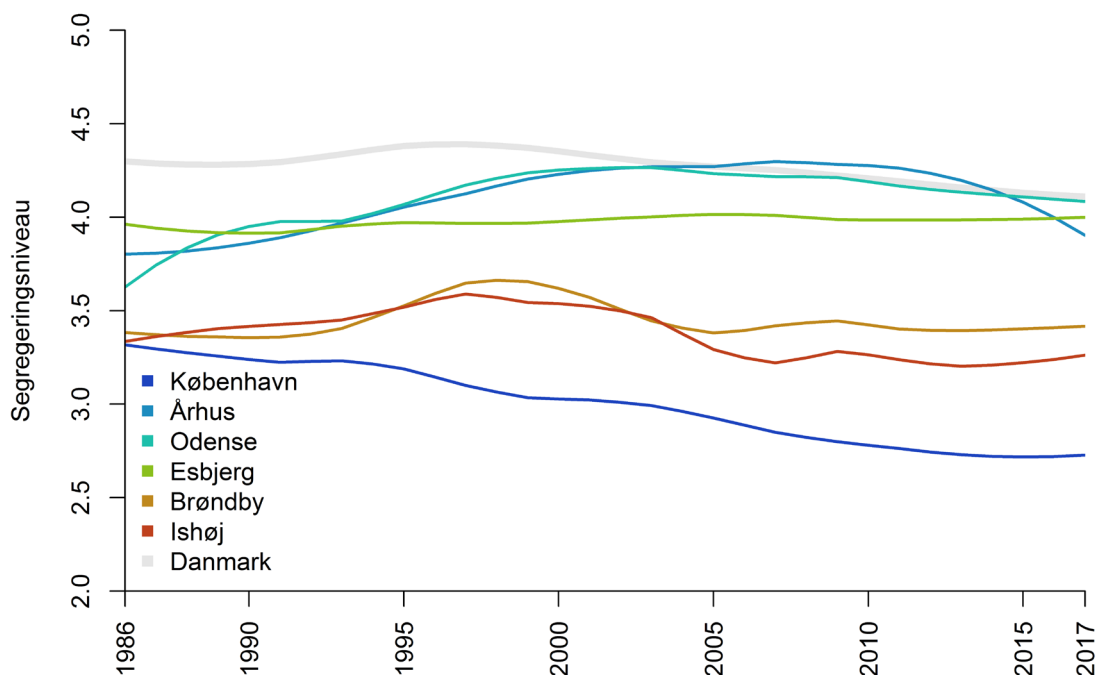
Vi ønsker nu at redegøre for, hvorledes den socioøkonomiske segregering har udviklet sig i de mest udsatte boligområder? Dette gøres ved at måle udviklingen i den geografiske koncentration i den 1. ventil i Figur 2.2, dvs. for de fem procent, der bor i nabolag med et lavt SNI³. På denne måde får vi en fornemmelse af, om det meget karakteristiske billede i Figur 2.1 er stabilt over tid eller ej.

³ Mere præcist: segregeringsmålet er defineret ved SNI i den 1. ventil fratrukket SNI i den 10. ventil (median-værdien af SNI).

Segregeringsmålet er beregnet for hele Danmark og seks udvalgte kommuner i form af København, Århus, Odense, Esbjerg, Brøndby og Ishøj. Figur 2.4 viser segregeringsniveauet over tid. Hvis segregeringsniveauet er nul, er bosættelsesmønsteret for de fem procent mest udsatte nabolag i kommunen det samme som det generelle bosættelsesmønster i kommunen. Der er herved taget højde for kommunernes boligudvikling over tid. Det fremgår, at segregeringsniveauet er større end nul for alle kommunerne, hvorfor personer i de fem procent mest udsatte nabolag i hver kommune lever tættere sammen end personer i gennemsnittet i kommunen. For hver af de seks udvalgte kommuner kan kort indeholdende det Socioøkonomiske Nabolag-Indeks ses i Appendiks A.

Fra Figur 2.4 ses, at segregeringsniveauet i Københavns Kommune er faldende over tid. Det vil sige, at sandsynligheden for at møde én person fra et udsat nabolag et vilkårligt sted i Københavns Kommune er blevet større over tid. For Århus og Odense ses det, at segregeringsniveauet fra 1986 til 2000 har været stigende, hvorefter det i en periode frem til 2013 har været rimeligt konstant. De seneste år har segregeringsniveauet i Århus og Odense været faldende. Udviklingen har for Esbjerg og Vestegnskommunerne Brøndby og Ishøj været rimeligt stabilt over tid. Dog oplevede Brøndby et midlertidigt øget segregeringsniveau i slut 90'erne. Det generelle segregeringsniveau i Danmark har over den betragtede periode været rimeligt stabilt, dog med en tendens til at være svagt faldende.

Figur 2.4. Segregeringsniveauet for Danmark og seks udvalgte kommuner.



Anm.: Segregeringsniveauerne er udglattet med metoden kaldet LOESS.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Det fremgår af Figur 2.4, at segregeringsniveauet i Københavns Kommune har været lavere end i de andre udvalgte kommuner. Det er dog vigtigt at understrege, at bosættelsesmønsteret og især den generelle befolkningstæthed i kommunerne er forskellig, hvorfor segregeringsniveauet mellem kommuner skal sammenlignes med forsigtighed. Århus, Esbjerg og Odense kommuner har indenfor kommunegrænsen et tæt befolket bycentrum, og et relativt tyndt befolket opland, som tilmed er forholdsvis stort. Som det senere fremgår, er der en generel tendens til, at udsatte nabolag i disse kommuner ligger omkring almene

boligområder i udkanten af bycentrum. København har en høj befolkningstæthed helt ud til kommunegrænsen, og de almene boligområder er spredt mere ud over hele kommunen. Hvis man betragter Århus, Esbjerg og Odense by, vil segregeringsniveauet i nogen grad minde mere om København. Det skal yderligere understreges, at man af samme grund ikke direkte kan sammenligne kommunernes segregeringsniveauer med hele Danmarks segregeringsniveau, hvorfor Danmarks segregeringsniveau også er anført med en svag grå nuance i Figur 2.4.

2.2 Geografiske afgrænsninger af udsatte boligområder i Danmark

I afsnit 2.1 defineredes det Socioøkonomiske Nabolag-Indeks, SNI. Ved hjælp af dette kan vi udpege de fem procent af den danske befolkning, der bor i nabolag med lavest socioøkonomisk status. Disse mennesker bor potentielt i et udsat boligområde. Vi ønsker imidlertid, at et udsat boligområde skal være et sammenhængende område med mindst 1.000 beboere. Dette sikres ved først at definere sammenhængende områder (ved at udføre en clusteranalyse), hvorefter vi smider de områder ud, hvor der bor færre end 1.000 beboere. Tilbage har vi de udsatte boligområder. Som vi vil se nedenfor, reducerer dette antallet af beboere fra fem procent til cirka fire procent.

Der er til formålet udviklet et Clusteranalyse-værktøj kaldet CNN-algoritmen (Clusted Nearest Neighbor). Se afsnit 4 for præcis beskrivelse af denne metode.

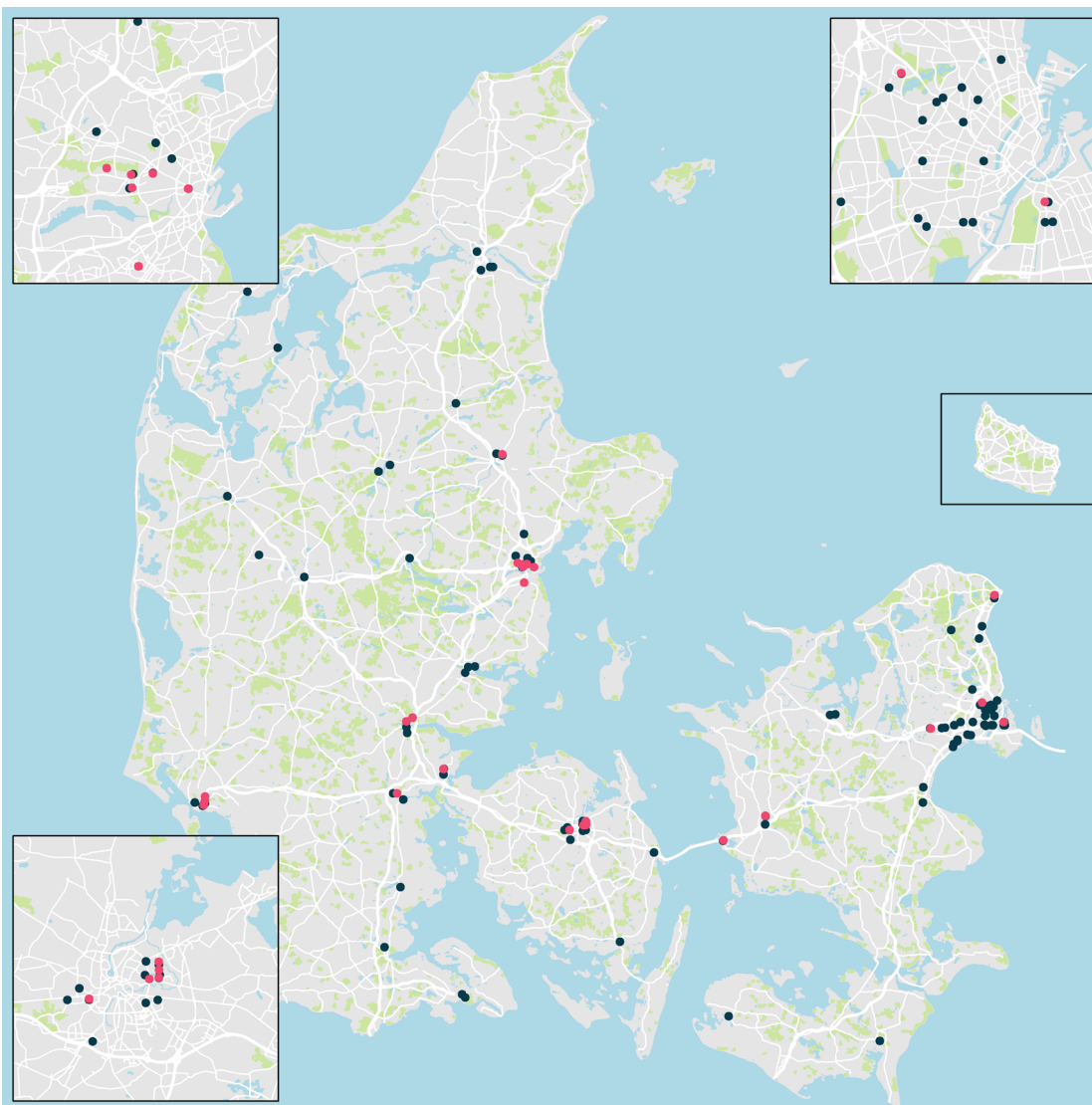
Ved indgangen til år 2017 vurderes 223.000 personer, svarende til cirka 4 procent af befolkningen, at være bosat i et udsat boligområde. Et udsat område defineres som en klynge af ejendomme med de mest udsatte nabolag, hvor fem procent af befolkningen bor. Derudover skal et område bestå af mindst 1.000 beboere.

De udsatte boligområder er fordelt på 38 kommuner. Figur 2.5 viser den geografiske placering af de identificerede udsatte områder ved indgangen til 2017. Det ses, at områderne generelt er placeret omkring de større byer. De røde prikker angiver de boligområder med de mest udsatte nabolag, hvor 1 procent af befolkningen bor.

Transport-, Bygnings- og Boligministeriet (TBB) udgiver årligt en liste over udsatte, almene boligområder. Listen kaldes populært for ghettolisten. TBBs almene boligområder bygger på stort set samme socioøkonomiske variable, som anvendes i denne analyse. Herudover anvender ministeriet også etnicitet som socioøkonomisk karakteristika, idet herkomst udtrykkes ved andelen af ikke-vestlige indvandrere og efterkommere i boligområdet. Der vil i nærværende analyse løbende blive sammenlignet med TBBs liste over udsatte områder. TBBs boligområder har været uændret, siden de blev defineret af Trafik- og Bygningsstyrelsen med baggrund i Programbestyrelsens anbefalede områder. Programbestyrelsen mod ghettoisering blev nedsat af den daværende VK-regering i 2004 og kom med sine anbefalinger i november 2008⁴. TBBs geografiske områder og kriterier for valget af områder er beskrevet i Transport-, Bygnings- og Boligministeriet (2018).

⁴ Bemærk, at TBBs boligområder er defineret ud fra følgende generelle retningslinjer, men at det for hvert enkelt boligområde beror på en konkret vurdering, hvorfor der kan afviges fra de generelle retningslinjer. Boligområderne vil bestå af en eller flere almene afdelinger. Almene boliger er typisk organiseret i afdelinger af boligforeninger. De enkelte afdelinger udgør en selvstændig økonomisk enhed med egen bestyrelse. Områderne skal ifølge almenboligloven være "fysisk sammenhængende", hvorved de almene afdelinger i et område skal grænse op til hinanden. Det enkelte område består af ensartede bebyggelser, hvorfor de f.eks. er opført indenfor en kortere tidsperiode og er af samme type boliger. Et område afgrænses af naturlige barrierer som større veje, jernbanespor, vandløb m.v. eller af områder som f.eks. erhvervsområder, offentlige bygninger, parker, kirkegårde og lignende. Det er vigtigt at understrege, at definitionen af et område vurderes enkeltvist, hvorfor ovenstående retningslinjer ikke

Figur 2.5. Udsatte boligområder i 2017.



Anm.: De sorte prikker markerer udsatte boligområder defineret som en klynge af ejendomme (mindst 1.000 beboere) med de mest udsatte nabolag, hvor fem procent af befolkningen bor. De røde prikker angiver de mest udsatte boligområder, defineret som klynger af ejendomme med de mest udsatte nabolag, hvor 1 procent af befolkningen bor.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Ved indgangen til 2017 bor cirka 1 procent af den danske befolkning i et af de boligområder, som TBBs ghettoliste 2018 klassificerer som en ghetto. Dette svarer nogenlunde til det samme antal personer, som bor i de 1 procent mest udsatte nabolag fundet via CNN-algoritmen (de røde områder i Figur 2.5).

4 procent af befolkningen bor i et udsat boligområde ud fra CNN-algoritmen. Det vil sige, at CNN-algoritmen i 2017 klassificerer 3 procentpoint flere personer bosat i udsatte boligområder end antallet af beboere i boligområder på ministeriets liste. 95 procent af personer, der bor i en TBB ghetto i 2017, er også fundet til at bo i et af udsat område fundet via CNN-algoritmen, mens der trods cirka det samme antal personer kun er 36 procent

nødvendigvis er overholdt for et boligområde. Nogle områder består f.eks. af både højt og lavt byggeri. For andre er der set bort fra, at afdelingerne i et område ikke helt grænser op til hinanden pga. en naturlig barriere.

overlap mellem personerne i TBB ghettoer og de 1 procent mest udsatte boligområder fundet via CNN-algoritmen.

2.3 Udsatte boligområders beboere i Danmark

I de udsatte boligområder, der identificeres af CNN-algoritmen, bor størstedelen af beboerne i almene boliger, mens et mindretal på godt 16 procent af beboerne i udsatte boligområder er bosat i øvrige boligtyper. Det fremgår af Tabel 2.1, hvor karakteristika på beboerne i udsatte boligområder dels sammenlignes med den samlede befolkning og dels med personer bosat i ikke-udsatte områder.

Tabel 2.1. Karakteristika for beboere i udsatte boligområder, primo 2017.

	Total befolkning	Udsatte boligområder		Ikke-udsatte boligområder	
		Almene boliger	Øvrige boliger	Almene boliger	Øvrige boliger
<i>Antal personer, tusinder</i>					
Antal i alt	5.749	187	36	794	4.732
Voksne beboere	4.378	131	31	626	3.589
<i>Aldersfordeling på beboere</i>					
Hjemmeboende børn	24%	30%	14%	21%	24%
Voksne yngre end 30 år	12%	16%	30%	15%	11%
Voksne 30-59 år	39%	38%	36%	35%	40%
Voksne 60 år eller ældre	25%	16%	20%	29%	25%
<i>Indkomst for beboere, tusinde 2016-kroner</i>					
Gns. indkomst, mindst 18 år	321	211	230	238	340
Gns. indkomst, 30-59 år	406	246	298	288	431
<i>Arbejdsmarkedstilknytning for beboere på mindst 18 år</i>					
I beskæftigelse	53%	34%	39%	35%	59%
Studerende	10%	13%	17%	12%	9%
Midlertidigt udenfor	5%	16%	9%	10%	4%
Førtidspension	4%	11%	6%	10%	3%
Tilbagetrækning	22%	16%	17%	28%	22%
Øvrige	6%	10%	12%	6%	5%
<i>Øvrige karakteristika</i>					
Ikke-vestlig herkomst	8%	49%	19%	19%	5%
Dømte, ældre end 15 år	0,9%	2,4%	3,8%	1,4%	0,7%
Grundskole, 30-59 år	26%	66%	48%	46%	21%

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Beboere i udsatte områder ses i gennemsnit at have en indkomst, som udgør cirka to tredjedele af indkomsten for personer bosat i ikke-udsatte områder. Den gennemsnitlige indkomst i udsatte områder er ligeledes lavere end blandt beboere i almene boliger i ikke-udsatte områder.

En lavere beskæftigelsesandel blandt voksne beboere i udsatte boligområder trækker mod en lavere gennemsnitlig indkomst end blandt personer bosat i ikke-udsatte områder. Det trækker også mod en lavere gennemsnitlig indkomst, at beboerne i udsatte områder er yngre end den øvrige befolkning, da indkomsten for midaldrende generelt er højere end for yngre personer. Ligeledes har den gennemsnitlige beboer i et udsat område opnået et lavere uddannelsesniveau end i befolkningen generelt. Da indkomsten generelt stiger, jo højere uddannelsesniveau man opnår, trækker dette ligeledes mod en lavere gennemsnitlig indkomst blandt beboerne i udsatte områder end i befolkningen som helhed.

I almene udsatte boligområder ses 2,4 procent af beboerne at være dømt for forseelser udover færdselsloven, hvilket er omkring tre gange større end som i befolkningen som helhed. I de almene udsatte områder ses andelen af personer med ikke-vestlig herkomst at være 49 procent, hvilket er mere end dobbelt så højt som blandt beboerne i almene boliger beliggende i ikke-udsatte områder.

Figur 2.6 viser den historiske udvikling fra 1986 til 2017 for herkomst, det Socioøkonomiske Nabolag-Indeks og for de socioøkonomiske segregeringsvariable uddannelses- og arbejdsmarkedstilknytning, kriminalitet, uddannelsesniveau og indkomst. Bemærk, at TBBs områder er konstante over tid, mens CNN-algoritmens områder varierer over tid.

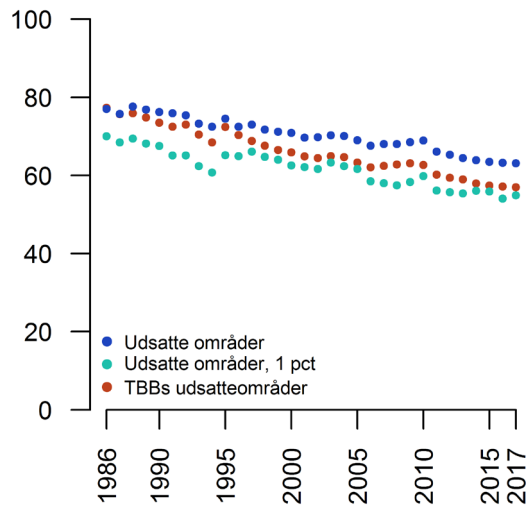
Fra Figur 2.6 ses en generel rangering af de forskellige definitioner af udsatte områder. De udsatte områder fundet via CNN-algoritmen minder mest om landsgennemsnittet, hvilket er oplagt fordi disse områder er større og herved naturligt er mere gennemsnitlige. De 1 procent mest udsatte områder og TBBs ghettoområder minder mere om hinanden. Der er dog en tendens til, at de mest udsatte områder fundet via CNN-algoritmen er mest forskellige fra landsgennemsnittet.

Det fremgår, at andelen i de udsatte områder uden tilknytning til uddannelses- og arbejdsmarkedet og andelen af dømte tilnærmelsesvis er stabile over tid. Det ses, at der for de udsatte områder er en faldende tendens i indkomsten i forhold til landsgennemsnittet over tid. Andelen af personer i udsatte boligområder med grundskole som højst fuldført uddannelsesniveau er i perioden rimeligt stabilt, mens der er en klar faldende tendens i landsgennemsnittet. De udsatte områder er således blevet dårligere uddannet relativt til resten af landet.

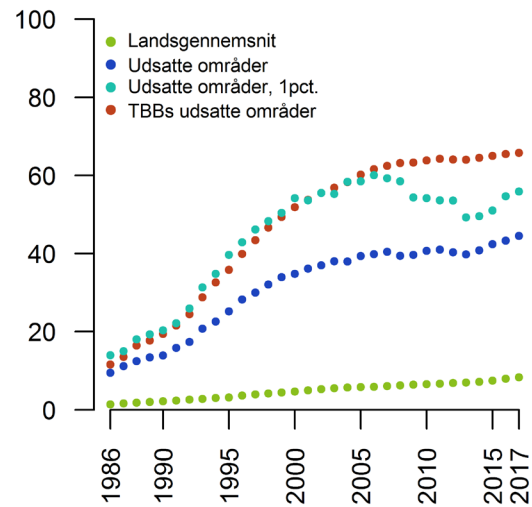
Fra Figur 2.6 ses desuden, at andelen af ikke-vestlige indvandrere og efterkommere er stigende over tid. Derudover ses det, at TBBs områder generelt har en større andel af ikke-vestlige indvandrere og efterkommere end de udsatte områder fundet via CNN-algoritmen.

Figur 2.6. Karakteristika for befolkningen i udsatte områder i perioden 1986-2017.

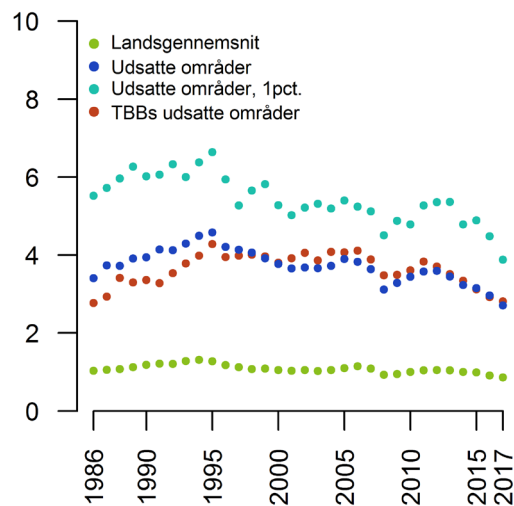
a) Indkomst ift. landsgennemsnittet, procent



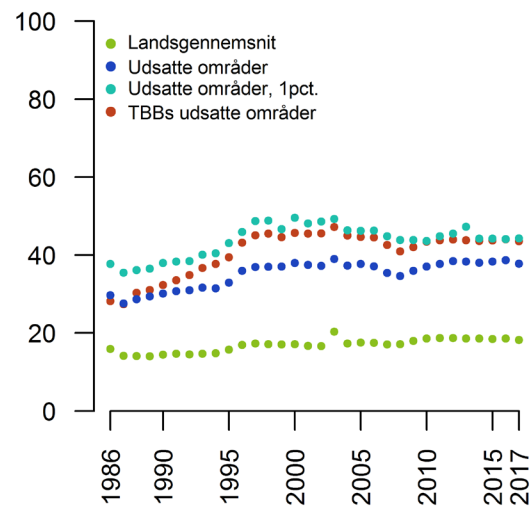
b) Ikke-vestlig herkomst, procent



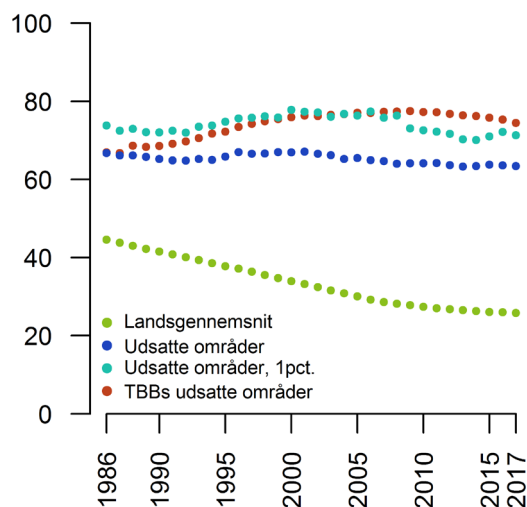
c) Kriminalitet, andel dømte i procent



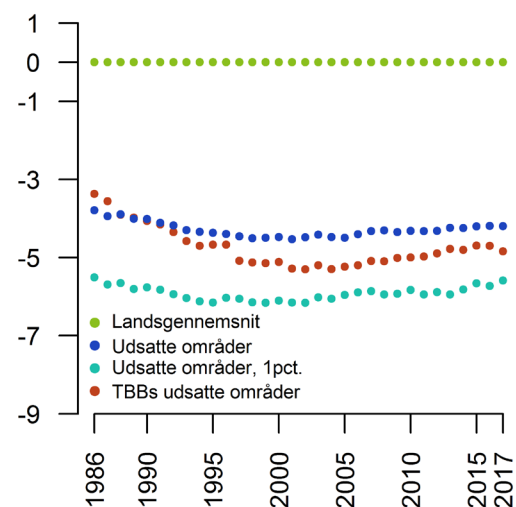
d) Ingen udd. og arb.markeds-tilknytning, procent



e) Grundskole som højst fuldført niveau, procent



f) Socioøkonomisk Nabolag-Indeks



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

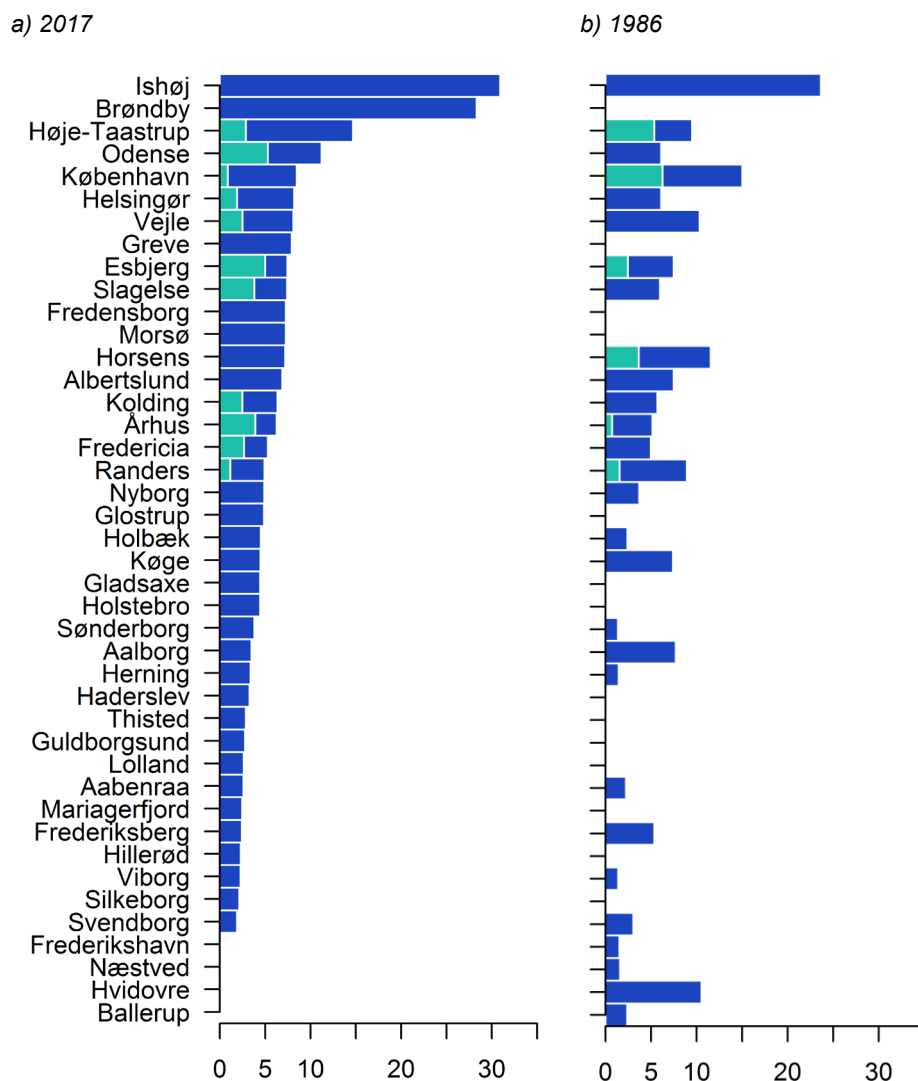
2.4 Udsatte boligområder i seks udvalgte kommuner

Ved indgangen til 2017 har kommunerne Ishøj og Brøndby den største andel af indbyggerne, som bor i et udsat område. I hver af disse to kommuner bor mere end 30 procent af indbyggerne i et udsat boligområde. Herefter følger en anden Vestegnskommune i form af Høje Taastrup, hvor andelen af indbyggere i udsatte boligområder er på cirka 15 procent, jf. Figur 2.7a.

I 1986 havde Ishøj Kommune også den største andel af indbyggere bosat i udsatte boligområder, jf. Figur 2.7b. Brøndby Kommune havde på det tidspunkt ingen udsatte områder. Områderne opstod dog allerede i 1987, jf. Figur 2.16d.

Det ses, at Københavns Kommune i 1986 havde en større andel af indbyggere i udsatte områder end det er tilfældet i 2017. Særligt boede en stor andel af befolkningen i de 1 procent mest udsatte områder i 1986 i Københavns Kommune. I 1986 var der fire kommuner, der havde beboere i udsatte områder, men som i 2017 ikke længere har udsatte områder. Dette omfatter Frederikshavn, Næstved, Hvidovre og Ballerup kommuner.

Figur 2.7. Andelen af kommunes indbyggere i et udsat boligområde.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

I de følgende afsnit fokuseres på seks kommuner: København, Odense, Århus, Esbjerg, Brøndby og Ishøj kommuner. Her vises den geografiske placering af udsatte områder i 2017, 2007, 1997 og 1987. Der vises ligeledes beboernes karakteristika i 2017 for de udsatte boligområder og endelig vises den historiske udvikling siden 1986 i herkomst, det Socioøkonomiske Nabolag-Indeks, indkomst, uddannelsesniveau, uddannelses- og arbejdsmarkedstilknytning og kriminalitet for områderne.

Udsatte boligområder i København

Geografiske afgrænsninger af udsatte boligområder i København

Figur 2.8a til d viser udsatte områder i København for henholdsvis 2017, 2007, 1997 og 1987. På kortene er udsatte områder identificeret via CNN-algoritmen angivet med mørkeblå og de 1 procent mest udsatte områder er angivet med lyseblå. Til sammenligning er TBBs ghettoområder fra 2018 indtegnet med rød. Bemærk, at TBBs områder er konstante over tid, hvorfor TBBs områder på kortet fra 1987, 1997 og 2007 er identiske med områderne på kortet omhandlende 2017.

Den overordnede tendens er, at en del af områderne har været stabile over tid. Det gælder områderne på Nørrebro, Københavns Nordvestkvarter, indre Amager og på Sydhavnen. Der ses dog også en tendens til, at udsatte områder på indre Nørrebro og især på Vesterbro er forsvundet over tid.

TBBs syv udsatte områder i Københavnsområdet i 2017 er også identificeret af CNN-algoritmen. CNN-områderne er generelt større end TBBs områder. Dette gælder især Hørgården på Amager og Degnegården mv. i Nordvestkvarteret. Det ses, at Aldersrogade og Mjølnerparken bliver til ét udsat område. De to meget udsatte områder, der er identificeret i 2017, er også på TBBs ghettoliste.

CNN-algoritmen identificerer foruden TBBs udsatte områder i 2017 også andre udsatte områder. Her tænkes på området omkring Kirsebærhaven og Vigerslev Vænge/Hornemanns Vænge i Valby og området fra P. Knudsens Gade til Scandiagade i Sydhavnen.

Som nævnt er udsatte områder i København ikke alle de samme som i 1987, jf. Figur 2.8. I 1987 var områderne på Ydre Nørrebro, f.eks. Aldersrogade og Mjølnerparken, ikke identificeret som udsatte områder⁵. Disse områder er dog identificeret i 1997.

I 1987 var der et stort udsat område omkring Sønder Boulevard på Vestrebro og et udsat område på Indre Nørrebro. Det ses, at store dele af disse områder også på dette tidspunkt er identificeret til at tilhøre de 1 procent mest udsatte boligområder.

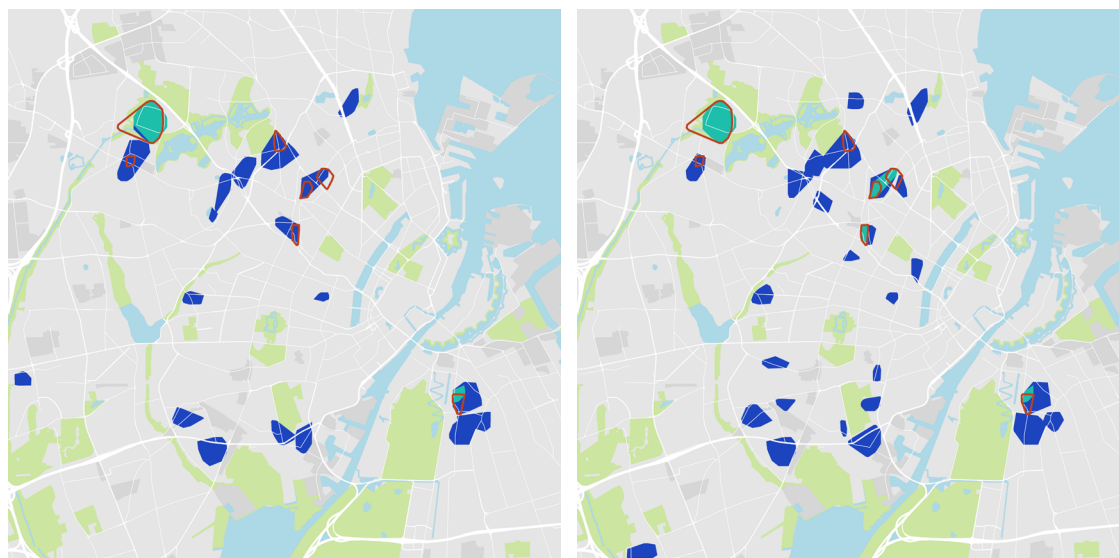
CNN-algoritmen finder, at det udsatte område på Vesterbro forsvinder op igennem 90'erne. Dette forklares blandt andet af store byfornyelsesprojekter, der netop har fundet sted i disse områder i perioden fra 1991 frem til 2005.

⁵ Mjølnerparken var først klar til indflytning i 1987.

Figur 2.8. Geografisk placering af udsatte boligområder i København.

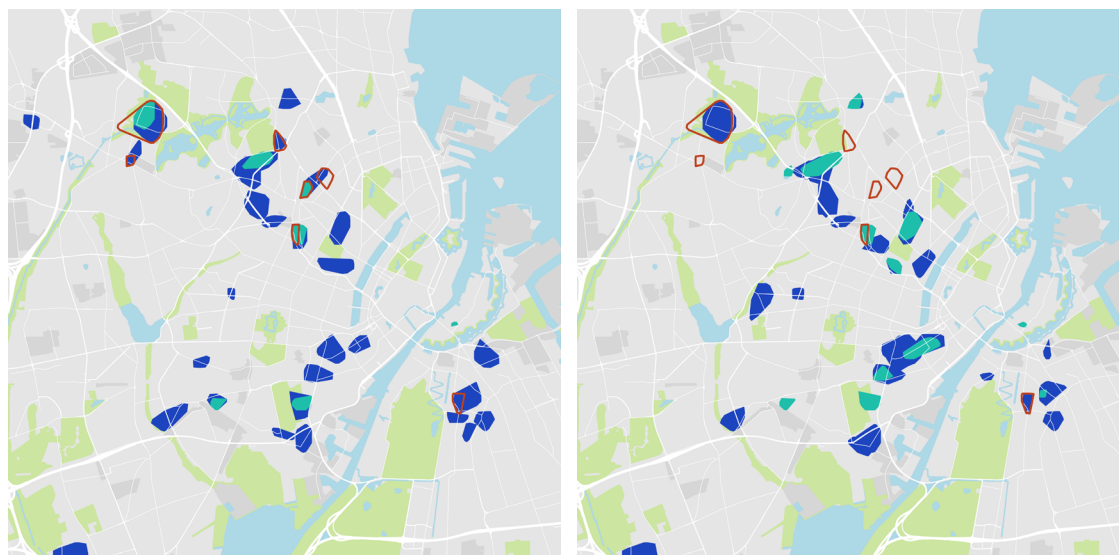
a) Udsatte områder i 2017

b) Udsatte områder i 2007



c) Udsatte områder i 1997

d) Udsatte områder i 1987



Anm: De mørkeblå plamager angiver udsatte områder, de lyseblå plamager angiver de 1 procent mest udsatte områder. De røde hylstre angiver TBBs udsatte områder december 2018.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Københavnere i udsatte boligområder

København er den kommune, hvor de fleste beboere vurderes at være bosat i udsatte områder. Det fremgår af Tabel 2.2, der sammenligner karakteristika for beboere i udsatte områder i Københavns Kommune med kommunens samlede befolkning og personer bosat i ikke-udsatte boligområder.

Ved begyndelsen af 2017 vurderes knap 51.000 af kommunens beboere at bo i udsatte boligområder, hvoraf 43.000 bor i almene boliger. Andelen af kommunens samlede befolkning, som bor i udsatte boligområder, er således cirka 8 procent.

Tabel 2.2. Karakteristika for beboere i udsatte boligområder i Københavns Kommune, primo 2017.

	Total befolkning	Udsatte boligområder		Ikke-udsatte boligområder	
		Almene boliger	Øvrige boliger	Almene boliger	Øvrige boliger
<i>Antal personer, tusinder</i>					
Antal i alt	602	43	8	72	476
Voksne beboere	482	31	7	57	387
<i>Aldersfordeling på beboere</i>					
Hjemmeboende børn	20%	28%	10%	25%	19%
Voksne yngre end 30 år	24%	16%	34%	17%	26%
Voksne 30-59 år	42%	40%	42%	39%	43%
Voksne 60 år eller ældre	14%	16%	13%	20%	13%
<i>Indkomst for beboere, tusinde 2016-kroner</i>					
Gns. indkomst, mindst 18 år	313	216	222	239	334
Gns. indkomst, 30-59 år	399	250	273	286	431
<i>Arbejdsmarkedstilknytning for beboere på mindst 18 år</i>					
I beskæftigelse	53%	37%	43%	38%	57%
Studerende	18%	13%	17%	14%	18%
Midlertidigt udenfor	7%	17%	11%	14%	4%
Førtidspension	2%	6%	3%	7%	1%
Tilbagetrækning	12%	15%	10%	19%	10%
Øvrige	9%	11%	16%	8%	9%
<i>Øvrige karakteristika</i>					
Ikke-vestlig herkomst	15%	51%	25%	33%	9%
Dømte, ældre end 15 år	1,2%	2,7%	5,0%	2,3%	0,9%
Grundskole, 30-59 år	29%	63%	49%	49%	22%

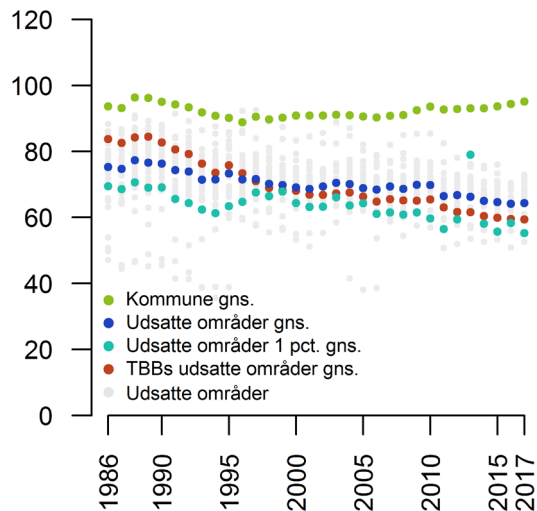
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 2.9 viser den historiske udvikling for København fra 1986 til 2017 for herkomst, det Socioøkonomiske Nabolag-Indeks og for de bagvedliggende socioøkonomiske variable: uddannelses- og arbejdsmarkedstilknytning, kriminalitet, uddannelsesniveau og indkomst. Bemærk, at TBBs områder er konstante over tid, mens CNN-algoritmens områder varierer over tid.

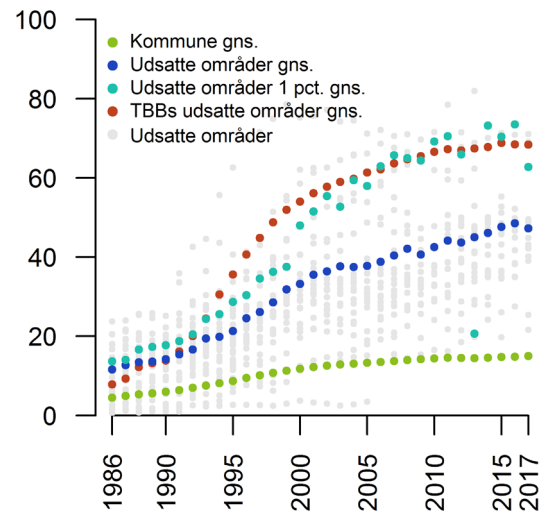
Fra Figur 2.9 ses den samme generelle rangering af de forskellige definitioner af udsatte områder i forhold til kommunegennemsnittet som beskrevet i afsnit 2.3, hvor udviklingen i hele landet beskrives. De udsatte områder fundet via CNN-algoritmen minder mest om landsgennemsnittet. De 1 procent mest udsatte områder og TBBs ghettoområder minder om hinanden. Der er dog en tendens til, at de mest udsatte områder fundet via CNN-algoritmen adskiller sig mest fra kommunegennemsnittet. Dette er især tydeligt de første år af den betragtede periode, hvilket forklares af, at de udsatte områder fundet via algoritmen kan variere over tid.

Figur 2.9 Karakteristika for beboere i udsatte områder i København i perioden 1986-2017.

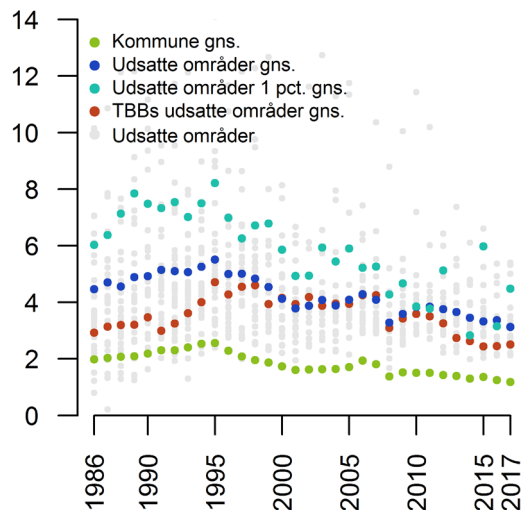
a) Indkomst ift. landsgennemsnittet, procent



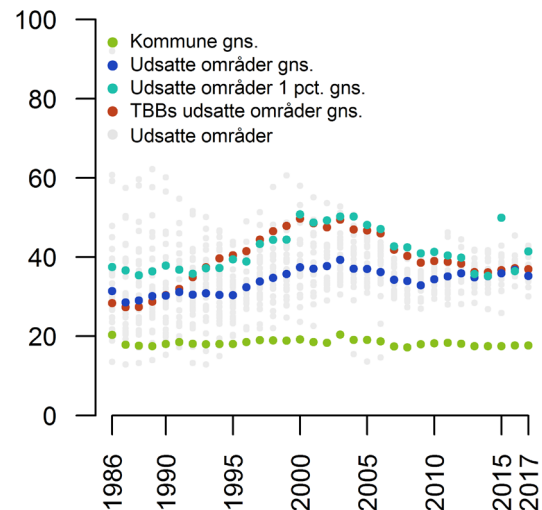
b) Ikke-vestlig herkomst, procent



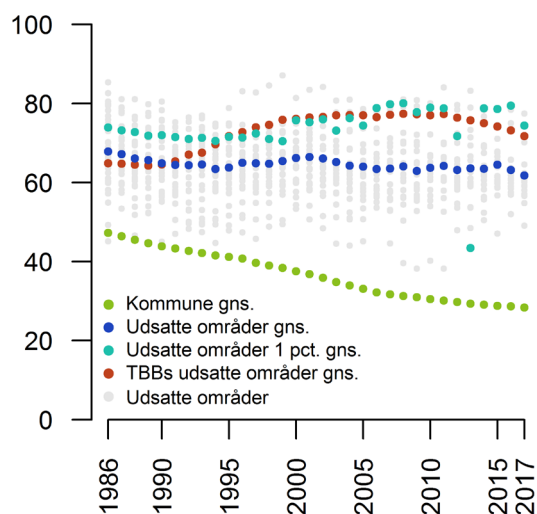
c) Kriminalitet, andel dømte i procent



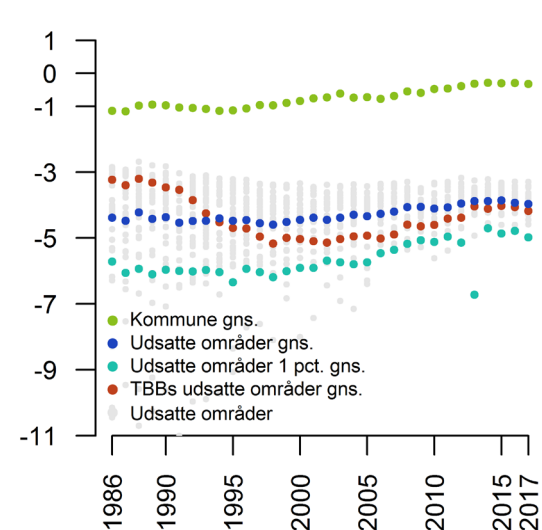
d) Uden udd. og arb.markeds-tilknytning, procent



e) Grundskole som højst fuldført niveau, procent



f) Socioøkonomisk Nabolag-Indeks



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Det Socioøkonomiske Nabolag-Indeks for Københavns Kommune er svagt stigende over tid. Det betyder, at kommunen i gennemsnit er mindre socioøkonomisk udsat i 2017 i forhold til landsgennemsnittet, end det var tilfældet i 1986. Trods dette ses en rimelig stabil udvikling i SNI for de udsatte områder fundet via algoritmen. De udsatte områder er således blevet relativt mere socioøkonomisk belastede i forhold til resten af kommunen. Det fremgår, at andelen i de udsatte områder uden tilknytning til uddannelses- og arbejdsmarkedet samt andelen af dømte er tilnærmelsesvis stabile over tid. Det ses, at der for de udsatte områder er en faldende tendens i indkomsten i forhold til kommunegennemsnittet over tid, og at kommunegennemsnittet er svagt stigende. Andelen af personer i udsatte boligområder med grundskole som højst fuldført uddannelsesniveau er i perioden rimeligt stabil, mens der er en klar faldende tendens i gennemsnittet for Københavns Kommune. De udsatte områder er således blevet dårligere uddannet relativt til resten af kommunen.

Fra Figur 2.9 ses, at andelen af ikke-vestlige indvandrere og efterkommere er stigende over tid. Derudover ses det, at TBBs områder og de 1 procent mest udsatte områder fundet via algoritmen generelt har en større andel af ikke-vestlige indvandrere og efterkommere end de større udsatte områder fundet via CNN-algoritmen.

Udsatte boligområder i Århus

Geografiske afgrænsninger af udsatte boligområder i Århus

Figur 2.10a til d viser udsatte områder i Århus for henholdsvis 2017, 2007, 1997 og 1987. I 2017 ses det, at de udsatte områder især findes i Århus V. Det fremgår, at de udsatte områder er lokaliseret længere væk fra centrum, end det for eksempel er tilfældet i København. Den geografiske placering af udsatte områder i Århus forekommer også mere stabile over tid.

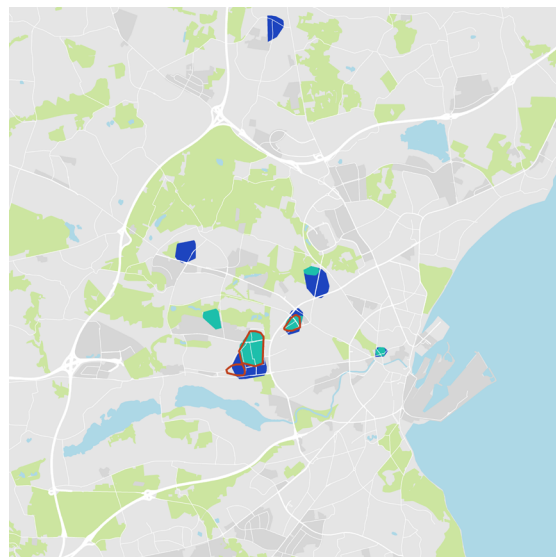
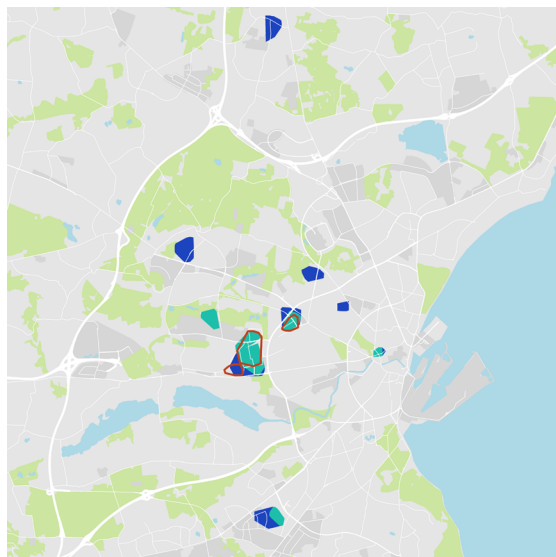
TBBs ghettoliste har tre udsatte områder i Århus: Bispehaven, Skovgårdsparken og Gellerupparken/Toveshøj. Disse tre områder er alle identificeret af CNN-algoritmen. Fra Figur 2.12 ses det, at Gellerupparken og Skovgårdsparken af CNN-algoritmen bliver identificeret som ét område, mens Toveshøj bliver identificeret som et selvstændigt område. Fra Figur 2.10 fremgår det, at Bispehaven og Gellerupparken/Toveshøj begge er identificeret i 2007, 1997 og 1987 som et udsat boligområde.

CNN-algoritmen identificerer foruden TBBs udsatte områder også to udsatte områder på Paludan-Müllers vej. Ét område omkring boligforeningen Ringgården ved Fuglesangs Allé og ét område omkring Lokalcenteret Toftegården. Området omkring Lokalcenteret Toftegården findes også tilbage i tid. Algoritmen finder også et udsat område ved Langkærparken i Tilst tilbage til 90'erne. I det centrale Århus udpeges der et, over tid stabilt, udsat område omkring Vesterbro Torv. Endelig findes der i 2017 et udsat boligområde syd for Århus på Søndervangs Allé i Viby.

Figur 2.10. Geografisk placering af udsatte boligområder i Århus.

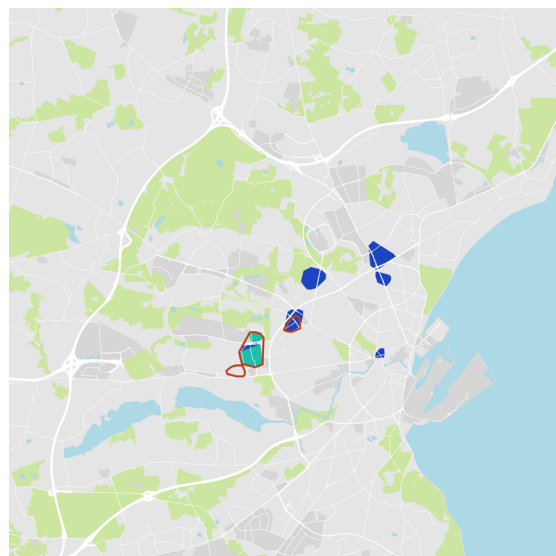
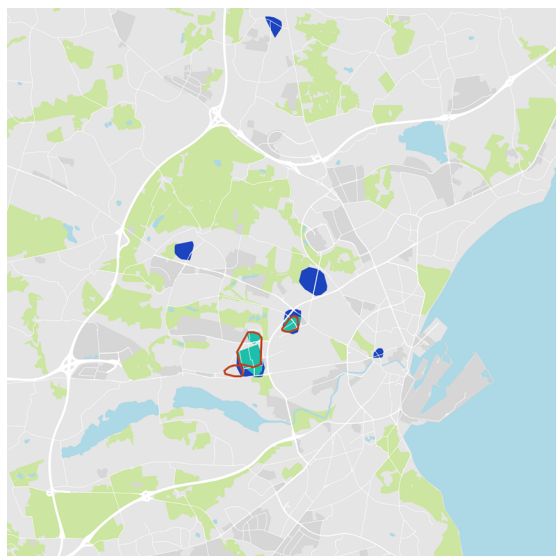
a) Udsatte områder i 2017

b) Udsatte områder i 2007



c) Udsatte områder i 1997

d) Udsatte områder i 1987



Anm.: De mørkeblå plamager angiver udsatte områder, de lyseblå plamager angiver de 1 procent mest udsatte områder. De røde hylstre angiver TBBs udsatte områder december 2018.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Århusianere i udsatte boligområder

I Århus Kommune vurderes godt 21.000 beboere at være bosat i udsatte boligområder, jf. Tabel 2.43. Langt størstedelen af disse bor til leje i almene boliger. Mens antallet af beboere i udsatte områder er landets næststørste efter København, ligger kommunens andel af det samlede antal beboere, som bor i et udsat område, forholdsvis lavt på godt 6 procent.

Beboerne i Århus Kommune er relativt unge, hvilket afspejler sig i en høj andel af personer under 30 år og en høj andel af studerende. Blandt beboerne i almene udsatte områder udgør studerende knap 19 procent af beboerne, hvilket er betydeligt mere end på landsplan. Øvrige beboer karakteristika i udsatte områder i 2017 for Århus Kommune er opsummeret i Tabel 2.3.

Tabel 2.3. Karakteristika for beboere i udsatte boligområder i Århus Kommune, primo 2017

	Total befolkning	Udsatte boligområder		Ikke-udsatte boligområder	
		Almene boliger	Øvrige boliger	Almene boliger	Øvrige boliger
<i>Antal personer, tusinder</i>					
Antal personer (tusind)	336	18	3	67	247
Antal voksne (tusind)	262	12	3	53	194
<i>Aldersfordeling på beboere</i>					
Hjemmeboende børn	22%	33%	8%	21%	22%
Voksne yngre end 30 år	23%	21%	60%	24%	23%
Voksne 30-59 år	36%	35%	21%	35%	36%
Voksne 60 år eller ældre	19%	11%	12%	19%	20%
<i>Indkomst for beboere, tusinde 2016-kroner</i>					
Gns. indkomst, mindst 18 år	309	196	160	227	341
Gns. indkomst, 30-59 år	416	239	77	284	465
<i>Arbejdsmarkedstilknytning for beboere på mindst 18 år</i>					
I beskæftigelse	48%	30%	23%	34%	54%
Studerende	20%	18%	39%	22%	19%
Midlertidigt udenfor	5%	15%	7%	10%	3%
Førtidspension	5%	17%	3%	11%	2%
Tilbagetrækning	16%	10%	10%	17%	16%
Øvrige	6%	10%	18%	7%	5%
<i>Øvrige karakteristika</i>					
Ikke-vestlig herkomst	11%	61%	20%	21%	5%
Dømte, ældre end 15 år	0,9%	2,8%	4,5%	1,4%	0,5%
Grundskole, 30-59 år	22%	69%	43%	39%	14%

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

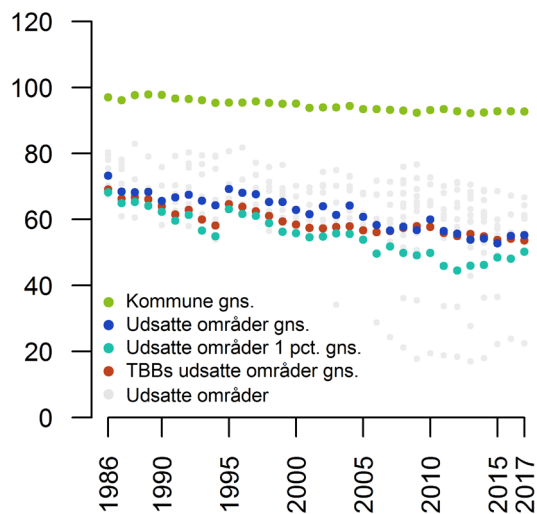
Figur 2.11 viser den historiske udvikling i Århus fra 1986 til 2017 for herkomst, det Socioøkonomiske Nabolag-Indeks og de bagvedliggende fire socioøkonomiske variable.

For de fire socioøkonomiske variable ses det, at de er tilnærmelsesvis stabile tilbage til 1995, og at der ikke er den store forskel mellem TBBs områder og områder fundet via CNN-algoritmen. Fra 1986 til 1995 ses det, at områderne med hensyn til andelen uden uddannelses- og arbejdsmarkedstilknytning og andelen med grundskole som højst fuldført uddannelsesniveau bevæger sig mere og mere væk fra kommunegennemsnittet.

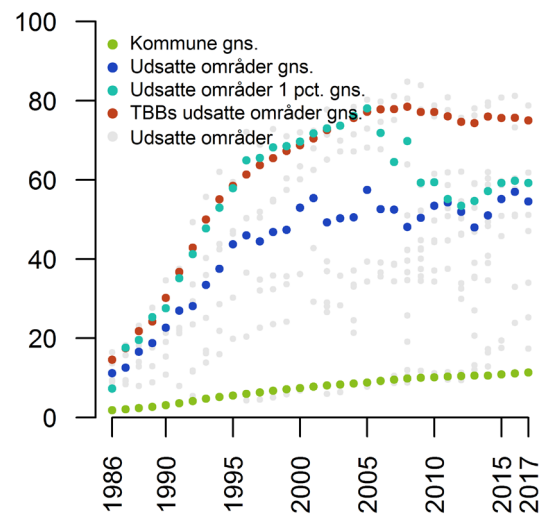
Fra Figur 2.11 ses det, at andelen af ikke-vestlige indvandrere og efterkommere er stigende over tid, og at TBB områderne har en højere andel af ikke-vestlige indvandrere og efterkommere især i de seneste år.

Figur 2.11. Karakteristika for beboere i udsatte områder i Århus i perioden 1986-2017.

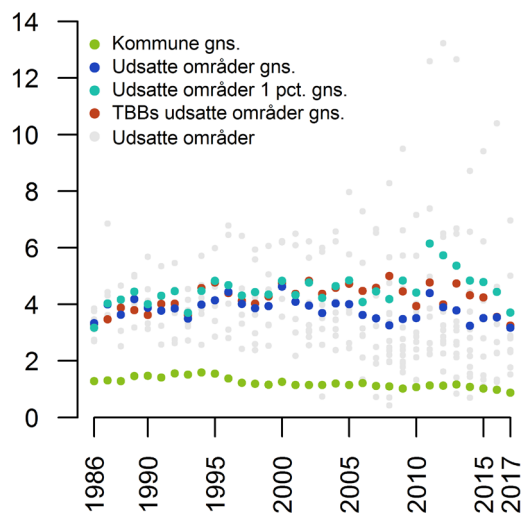
a) Indkomst iff. landsgennemsnittet, procent



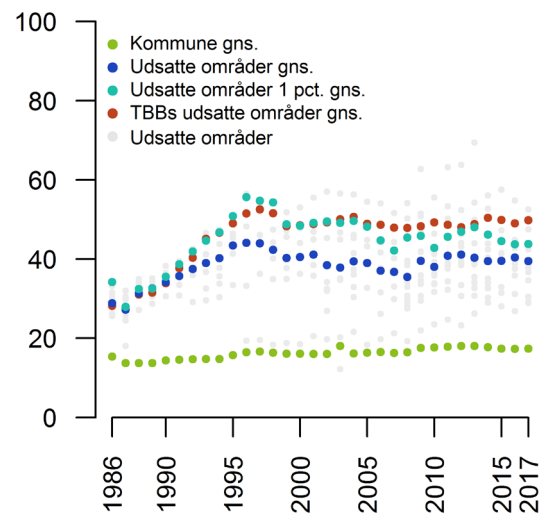
b) Ikke-vestlig herkomst, procent



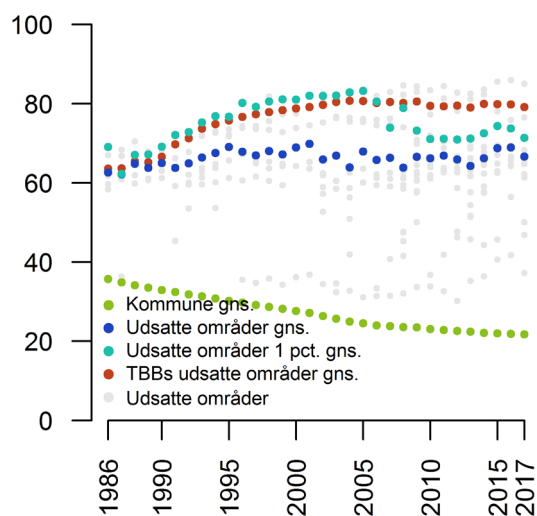
c) Kriminalitet, andel dømte i procent



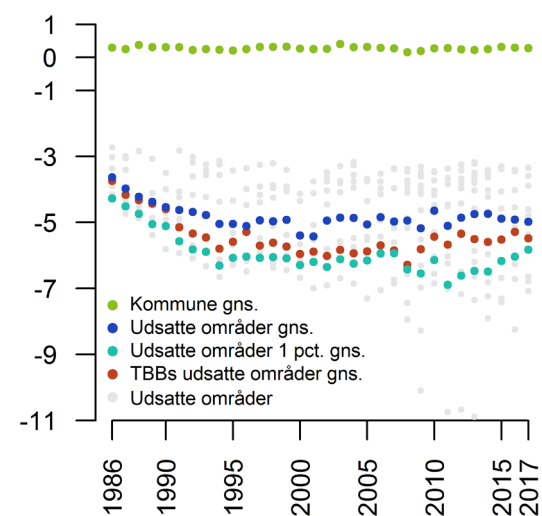
d) Uden udd. og arb.markeds-tilknytning, procent



e) Grundskole som højst fuldført niveau, procent



f) Socioøkonomisk Nabolag-Indeks



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Udsatte boligområder i Odense

Geografiske afgrænsninger af udsatte boligområder i Odense

Figur 2.12a til d viser udsatte områder i Odense for henholdsvis 2017, 2007, 1997 og 1987. Det ses, at de udsatte områder særligt findes i Odense NØ, Odense M og Odense V. TBBs tre udsatte områder i Odense – Solbakken mv., Korsløkkeparken Øst og Vollsmose – er også identificeret af CNN-algoritmen.

Fra Figur 2.12 ses det, at Vollsmose bliver opdelt i mindre områder. Mod nord langs Kertemindevej identificeres Slåenhaven/Tjørnehaven og Hybenhaven som to udsatte boligområder. Disse områder består mest af rækkehuse. Ét østligt område langs Vollsmose Allé bestående af Birkeparken, Bøgeparken og Egeparken og endelig to sydlige områder langs Åsumvej, hvor det ene består af Fyrreparken og Lærkeparken, og det andet består af Granparken. Det østlige og sydlige område af Vollsmose består af etagebyggerier. De fem områder dækker tilsammen et område, der er lidt større end selve Vollsmoseområdet, idet f.eks. boliger fra Biskorpvej på den anden side af Vollsmose Allé medtages i det udsatte område Hybenhaven. Algoritmen vurderer derved, at personer, der bor i dette område, også oplever at bo i et segregeret boligområde. Det fremgår, at store dele af Vollsmose stabilt identificeres til at tilhøre de 1 procent mest udsatte områder.

I 1987 identificeres kun den nordvestlige del af Vollsmose omkring Slåenhaven og Tjørnehaven som et udsat boligområde, jf. Figur 2.12d. Bemærk, at dette område ikke længere identificeres som et udsat område eller en del af Vollsmose-ghettoen i den nyeste ghettoliste fra TBB.

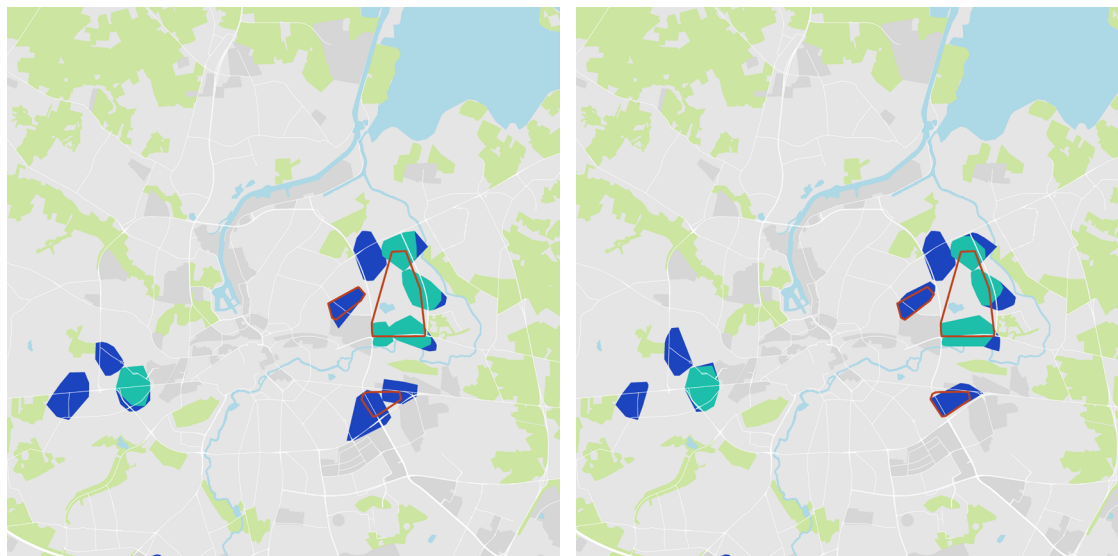
Fra Figur 2.12a ses det, at området omkring Korsløkkeparken Øst og Ejerslykke bliver opdelt i to områder: Ét område omkring Nyborgvej og ét område omkring Rødegårdsvej.

CNN-algoritmen identificerer foruden TBBs udsatte områder også udsatte områder i Odense V. Ét område omkring det almene boligområde ved Bystævnevej, og ét område omkring Rosskovsvej, som blandt andet indeholder et kollegie/botilbud målrettet borgere med svær sindslidelse. Desuden identificeres et område omkring Rasmus Rask Kollegiet.

Figur 2.12. Geografisk placering af udsatte boligområder i Odense.

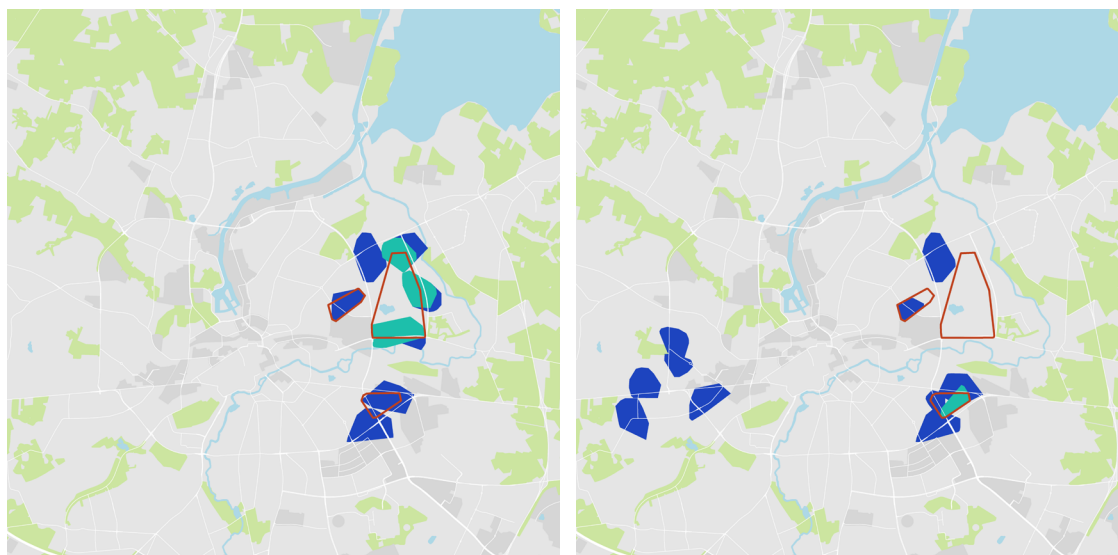
a) Udsatte områder i 2017

b) Udsatte områder i 2007



c) Udsatte områder i 1997

d) Udsatte områder i 1987



Anm: De mørkeblå plamager angiver udsatte områder, de lyseblå plamager angiver de 1 procent mest udsatte områder. De røde hylstre angiver TBBs udsatte områder december 2018.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Odenseanere i udsatte boligområder

I Odense vurderes cirka 23.000 af kommunens i alt 201.000 beboere at bo i udsatte boligområder, jf. Tabel 2.4. I Odense Kommune er andelen af beboere i udsatte boligområder på knap 11 procent, hvilket i 2017 overgår niveauet i København og kun overgås af en række kommuner på den københavnske vestegn.

Størstedelen af de udsatte områder i Odense Kommune er almene boliger. I disse områder er andelen af beboere i beskæftigelse noget lavere end i udsatte almene boligområder i de øvrige større byer. Dette afspejles i, at den gennemsnitlige indkomst blandt beboerne i almene udsatte områder er relativt lav sammenlignet med den gennemsnitlige indkomst for den tilsvarende gruppe i København og Århus.

Tabel 2.4. Karakteristika for beboere i udsatte boligområder i Odense Kommune, primo 2017.

	Total befolkning	Udsatte boligområder		Ikke-udsatte boligområder	
		Almene boliger	Øvrige boliger	Almene boliger	Øvrige boliger
<i>Antal personer, tusinder</i>					
Antal i alt	201	19	4	25	153
Voksne beboere	156	13	3	21	118
<i>Aldersfordeling på beboere</i>					
Hjemmeboende børn	22%	29%	14%	17%	23%
Voksne yngre end 30 år	19%	21%	30%	23%	18%
Voksne 30-59 år	36%	35%	35%	35%	37%
Voksne 60 år eller ældre	22%	15%	26%	26%	22%
<i>Indkomst for beboere, tusinde 2016-kroner</i>					
Gns. indkomst, mindst 18 år	286	187	227	215	311
Gns. indkomst, 30-59 år	377	227	291	265	414
<i>Arbejdsmarkedstilknytning for beboere på mindst 18 år</i>					
I beskæftigelse	46%	26%	35%	29%	52%
Studerende	17%	15%	18%	20%	17%
Midlertidigt udenfor	7%	20%	11%	11%	4%
Førtidspension	5%	14%	6%	11%	3%
Tilbagetrækning	20%	14%	21%	23%	20%
Øvrige	5%	11%	9%	6%	5%
<i>Øvrige karakteristika</i>					
Ikke-vestlig herkomst	11%	54%	19%	16%	5%
Dømte, ældre end 15 år	1,0%	2,8%	4,9%	1,4%	0,7%
Grundskole, 30-59 år	25%	68%	41%	45%	17%

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

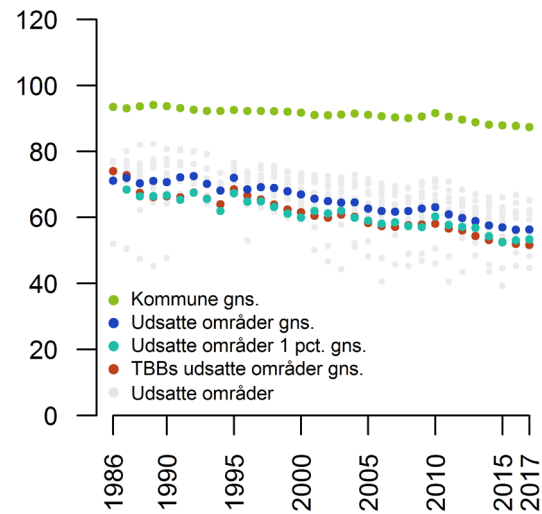
Figur 2.13 viser den historiske udvikling fra 1986 til 2017 for herkomst, det Socioøkonomiske Nabolag-Indeks og de bagvedliggende socioøkonomiske variable: uddannelses- og arbejdsmarkedstilknytning, kriminalitet, uddannelsesniveau og indkomst.

For de fire socioøkonomiske variable ses det, at de er tilnærmelsesvis stabile over tid, og at der ikke er den store forskel mellem TBBs områder og områder fundet via CNN-algoritmen. Dog med en faldende tendens i gennemsnitsindkomsten for de udsatte områder over tid.

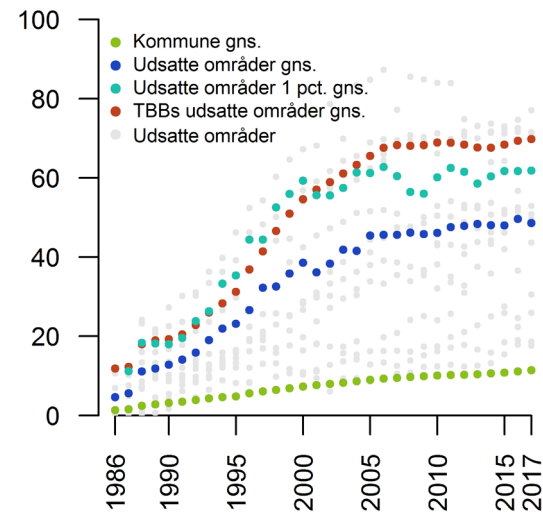
Fra Figur 2.13 ses det, at andelen af ikke-vestlige indvandrere og efterkommere generelt er stigende over tid.

Figur 2.13. Karakteristika for beboere i udsatte områder i Odense i perioden 1986-2017

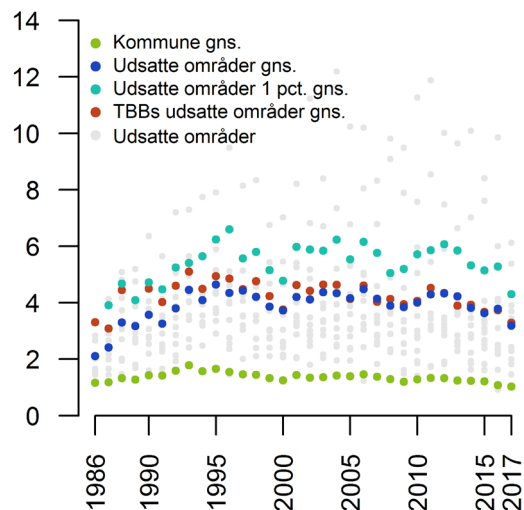
a) Indkomst iff. landsgennemsnittet, procent



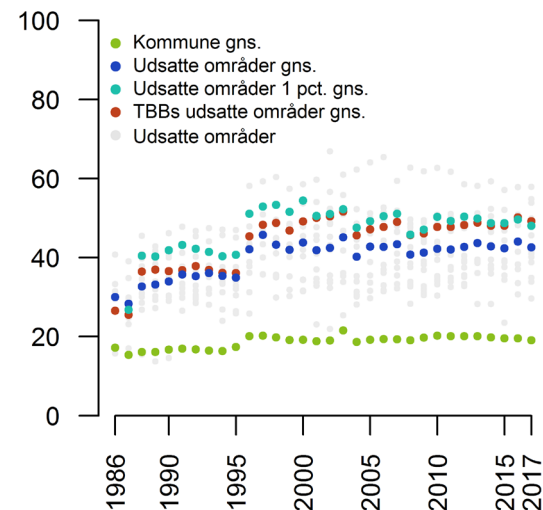
b) Ikke-vestlig herkomst, procent



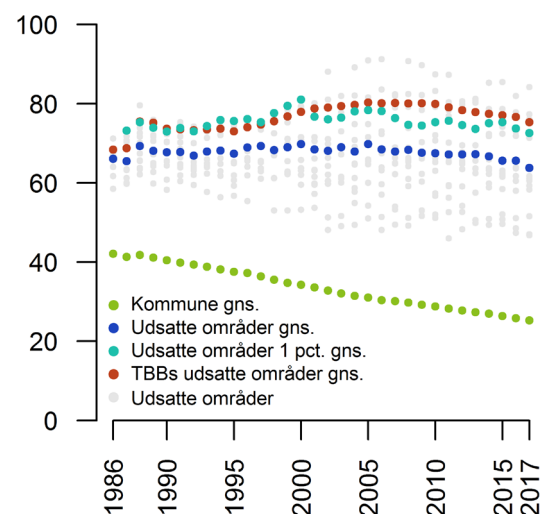
c) Kriminalitet, andel dømte i procent



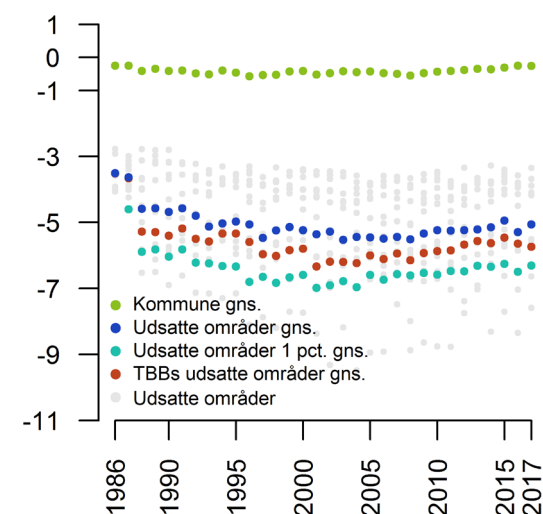
d) Uden udd. og arb.markeds-tilknytning, procent



e) Grundskole som højst fuldført niveau, procent



f) Socioøkonomisk Nabolag-Indeks



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Udsatte boligområder i Esbjerg

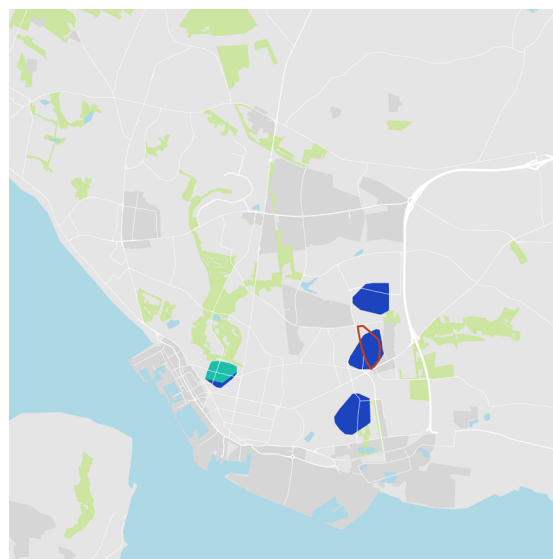
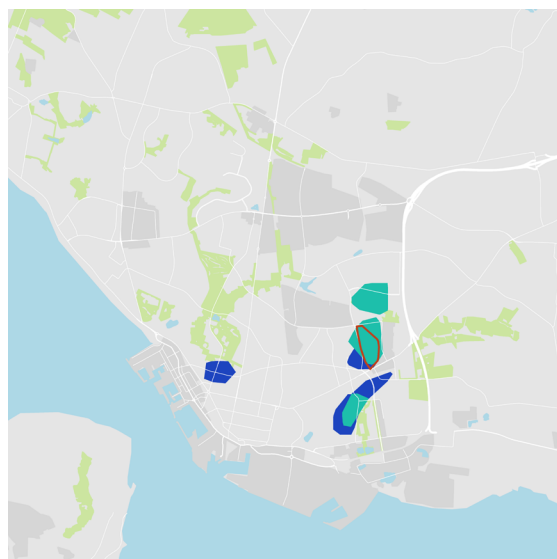
Geografiske afgrænsninger af udsatte boligområder i Esbjerg

Figur 2.14a til d viser udsatte boligområder i Esbjerg for henholdsvis 2017, 2007, 1997 og 1987. Stengårdsvej, som er udpeget til ghettoområde af TBB, identificeres ligeledes af CNN-algoritmen i både 2017 og tilbage i tid. Området er i 2017 også et af de 1 procent mest udsatte områder. Derudover identificeres udsatte område omkring Hedelundgårdparken (området hed Kvaglund før 2013), Præstebakken/Syrenparken og et område omkring Grådybet. Den geografiske placering af udsatte områder i Esbjerg har generelt været stabil over tid.

Figur 2.14. Geografisk placering af udsatte boligområder i Esbjerg.

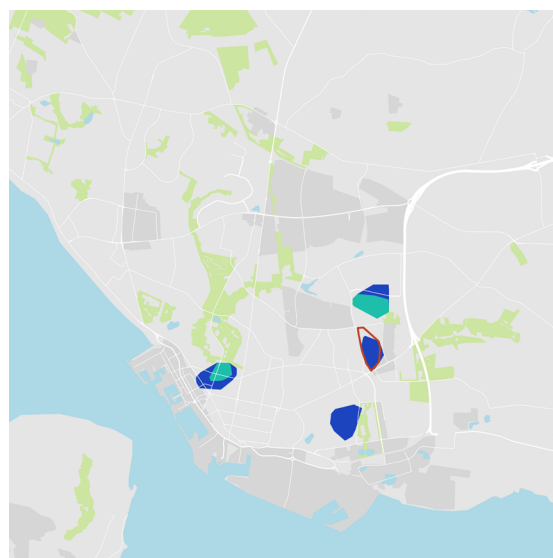
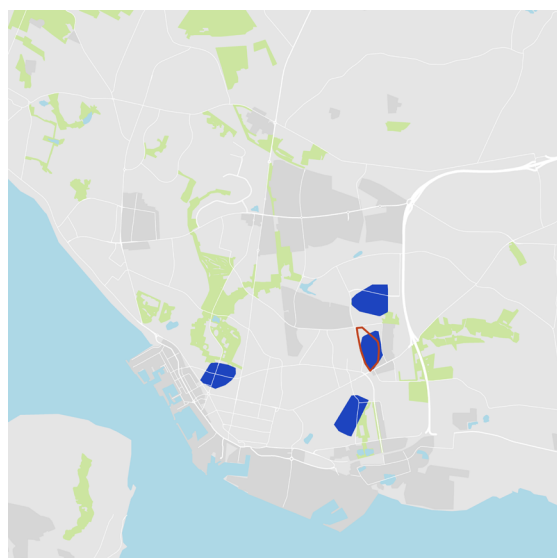
a) Udsatte områder i 2017

b) Udsatte områder i 2007



c) Udsatte områder i 1997

d) Udsatte områder i 1987



Anm: De mørkeblå plamager angiver udsatte områder, mens de lyseblå plamager angiver de 1 procent mest udsatte områder. De røde hylstre angiver TBBs udsatte områder december 2018.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Esbjergensere i udsatte boligområder

I Esbjerg Kommune vurderes cirka 8.000 beboere at bo i udsatte boligområder. Dette svarer til knap 7 procent af kommunens samlede befolkning i 2017. I forhold til andelen af kommunens samlede antal beboere, som bor i et udsat boligområde, placerer Esbjerg Kommune sig herved cirka en tredjedel nede af listen med de i alt 38 kommuner med udsatte boligområder.

Karakteristika for beboere i udsatte områder i Esbjerg Kommune er sammenfattet i Tabel 2.5. I tabellen sammenlignes beboerne i udsatte områder med kommunens samlede gennemsnit og personer bosat i ikke-udsatte boligområder.

Tabel 2.5. Karakteristika for beboere i udsatte boligområder i Esbjerg Kommune, primo 2017.

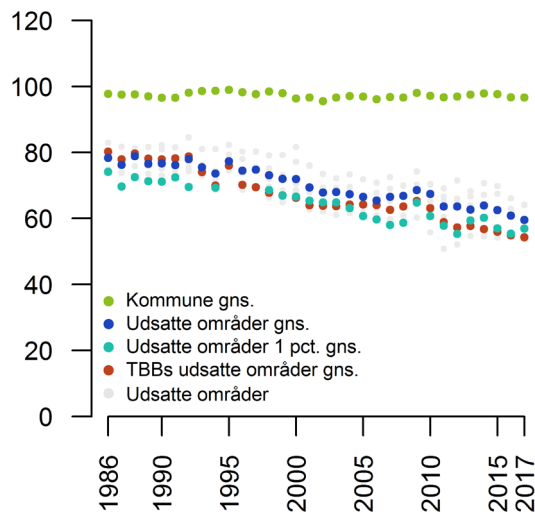
	Total befolkning	Udsatte boligområder		Ikke-udsatte boligområder	
		Almene boliger	Øvrige boliger	Almene boliger	Øvrige boliger
<i>Antal personer, tusinder</i>					
Antal i alt	116	7	1	18	90
Voksne beboere	88	5	1	14	68
<i>Aldersfordeling på beboere</i>					
Hjemmeboende børn	24%	27%	13%	19%	25%
Voksne yngre end 30 år	12%	21%	16%	17%	10%
Voksne 30-59 år	38%	35%	43%	35%	39%
Voksne 60 år eller ældre	26%	18%	28%	29%	26%
<i>Indkomst for beboere, tusinde 2016-kroner</i>					
Gns. indkomst, mindst 18 år	309	195	235	229	336
Gns. indkomst, 30-59 år	387	233	258	279	419
<i>Arbejdsmarkedstilknytning for beboere på mindst 18 år</i>					
I beskæftigelse	52%	28%	30%	33%	58%
Studerende	9%	16%	8%	13%	8%
Midlertidigt udenfor	5%	18%	20%	10%	3%
Førtidspension	5%	11%	12%	11%	3%
Tilbagetrækning	23%	17%	24%	28%	23%
Øvrige	5%	10%	7%	5%	4%
<i>Øvrige karakteristika</i>					
Ikke-vestlig herkomst	7%	39%	11%	10%	3%
Dømte, ældre end 15 år	1,0%	2,6%	5,3%	1,5%	0,7%
Grundskole, 30-59 år	27%	69%	55%	46%	20%

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

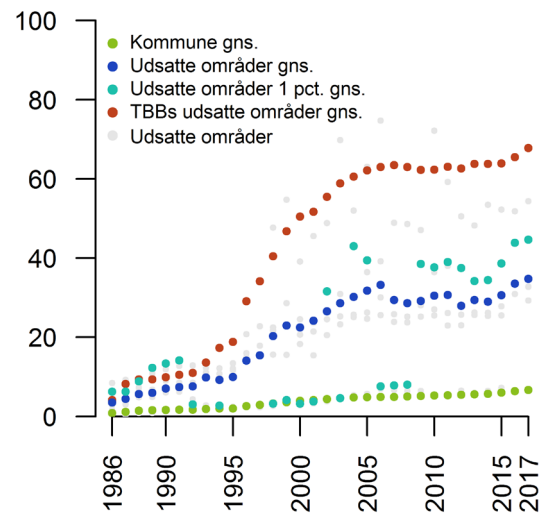
Figur 2.15 viser den historiske udvikling i Esbjerg fra 1986 til 2017 for herkomst, det Socioøkonomiske Nabolag-Indeks og de bagvedlæggende socioøkonomiske variable uddannelses- og arbejdsmarkedstilknytning, kriminalitet, uddannelsesniveaue og indkomst.

Figur 2.15. Karakteristika for beboere i udsatte områder i Esbjerg i perioden 1986 til 2017.

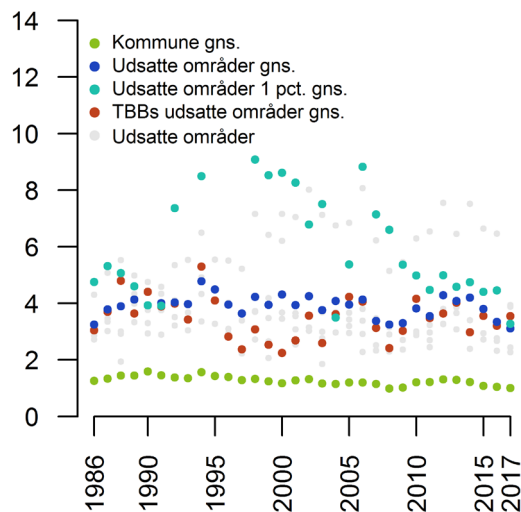
a) Indkomst ift. landsgennemsnittet, procent



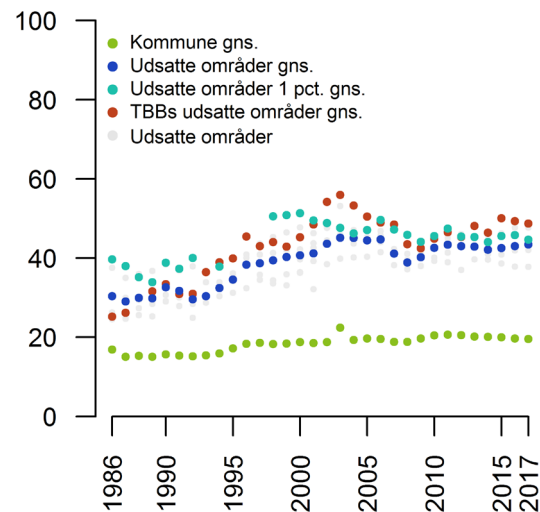
b) Ikke-vestlig herkomst, procent



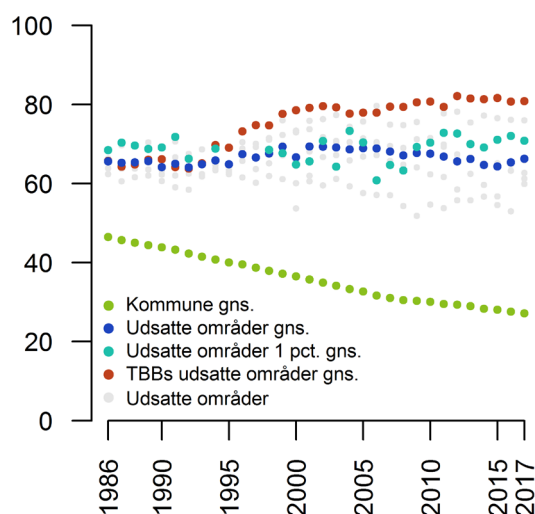
c) Kriminalitet, andel dømte i procent



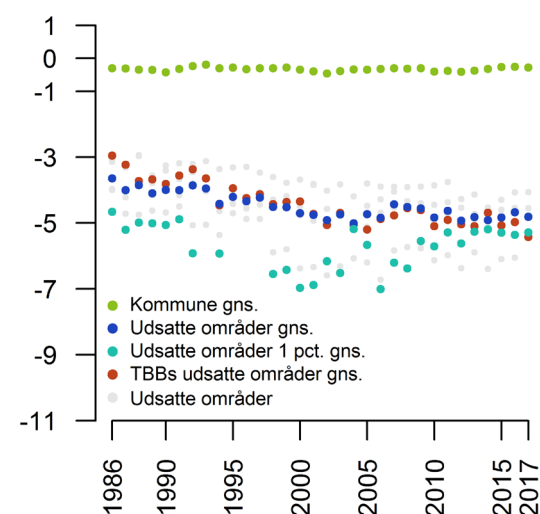
d) Uden udd. og arb.markeds-tilknytning, procent



e) Grundskole som højst fuldført niveau, procent



f) Socioøkonomisk Nabolag-Indeks



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

De geografisk udsatte områder i Esbjerg har været rimeligt stabile over tid, og der er ikke den store forskel ift. TBBs områder. For de fire socioøkonomiske variable ses det, at de er tilnærmelsesvis stabile, mens andelen af ikke-vestlige indvandrere og efterkommere er stigende over tid. Bemærk, at de 1 procent mest udsatte områder svinger meget, hvilket er et udtryk for, at visse områder ligger og svinger mellem at tilhøre denne gruppe eller ej.

Udsatte boligområder i Brøndby og Ishøj

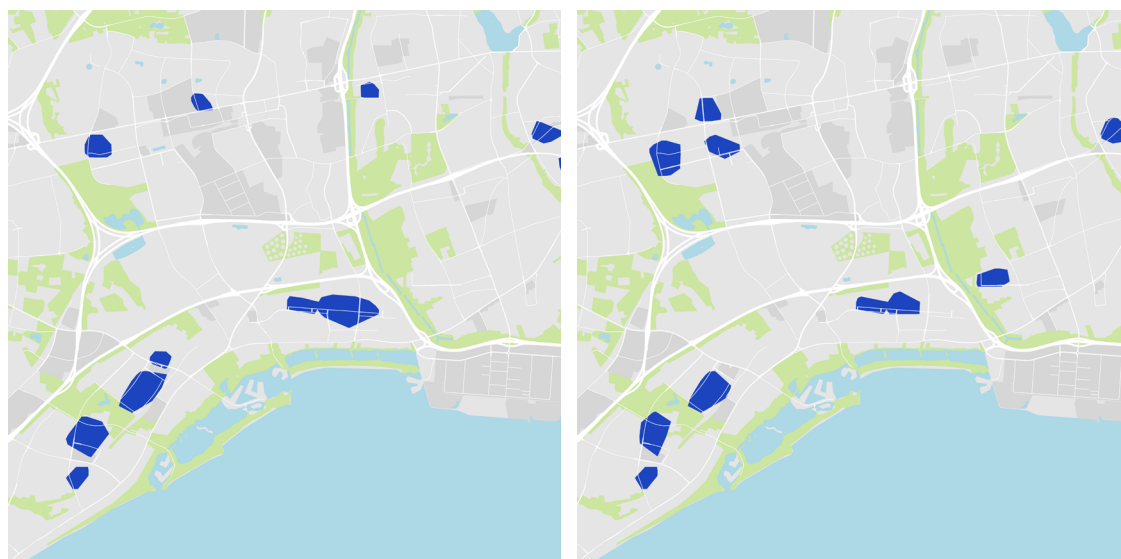
Geografiske afgrænsninger af udsatte boligområder i Brøndby og Ishøj

I det følgende beskrives den geografiske placering og beboere i udsatte områder i Brøndby og Ishøj under ét, idet kommunerne geografisk er placeret op ad hinanden, ligesom kommunernes indbyggere har tilnærmelsesvis samme karakteristika.

Figur 2.16. Geografisk placering af udsatte boligområder i Brøndby og Ishøj.

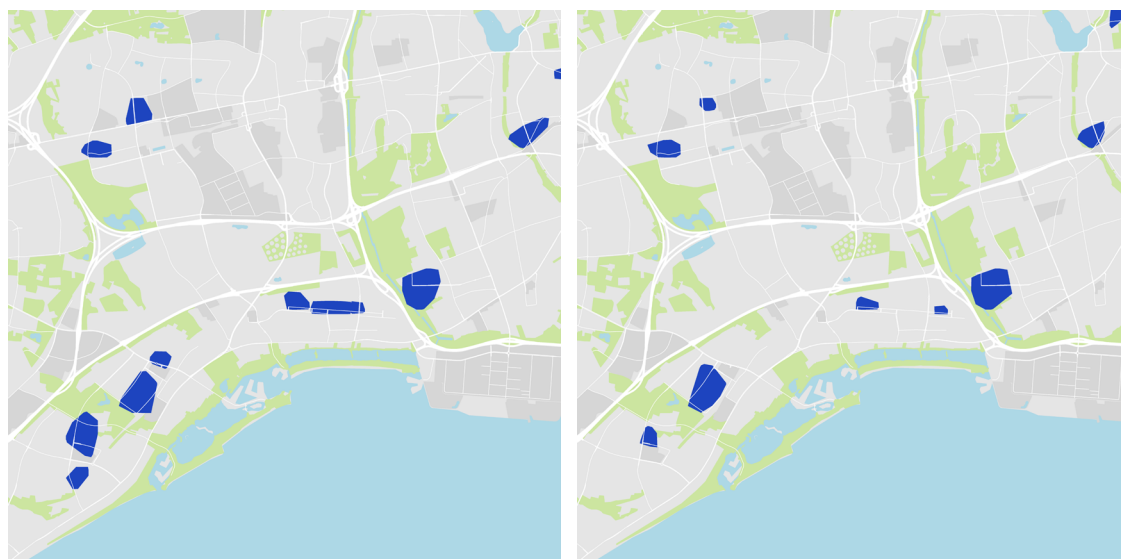
a) Udsatte områder i 2017

b) Udsatte områder i 2007



c) Udsatte områder i 1997

d) Udsatte områder i 1987



Anm: De mørkeblå plamager angiver udsatte områder, de lyseblå plamager angiver de 1 procent mest udsatte områder.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 2.16a til d viser udsatte boligområder i Brøndby og Ishøj for henholdsvis 2017, 2007, 1997 og 1987. Det kan bemærkes, at TBB ikke finder nogle ghettoområder i disse to kommuner, og CNN-algoritmen heller ikke finder nogle af de 1 procent mest udsatte områder i kommunerne.

I Brøndby Kommune identificeres et udsat område ved Brøndby Strand Parkerne. I 2017 identificeres Hallingsparken, Kisumparken, Bjerrelund, Gurrelund, Albjergparken og Tybjergparken som udsatte. I 1987 er det kun området omkring Hallingsparken og Tybjergparken, der findes til at være et udsat boligområde. I 2017 identificeres Brøndby Nord Vej også til et udsat boligområde. I Ishøj Strand identificeres Vildtbanegård, Gadekæret og Vejleåparken som udsatte. Disse områder har været rimeligt stabile over tid.

Figur 2.16 viser, at der også er udsatte områder andre steder på Vestegnen. Fra kortet fremgår, at der ligeledes er udsatte steder i Albertslund, Hundige og Avedøre.

Tablet 2.6. Karakteristika for beboere i udsatte boligområder i Brøndby og Ishøj kommuner, primo 2017.

	Total befolkning	Udsatte boligområder		Ikke-udsatte boligområder	
		Almene boliger	Øvrige boliger	Almene boliger	Øvrige boliger
<i>Antal personer, tusinder</i>					
Antal i alt	58	16	1	16	25
Voksne beboere	43	11	1	12	19
<i>Aldersfordeling på beboere</i>					
Hjemmeboende børn	26%	32%	21%	23%	25%
Voksne yngre end 30 år	11%	13%	15%	12%	8%
Voksne 30-59 år	40%	39%	38%	39%	41%
Voksne 60 år eller ældre	23%	16%	26%	26%	26%
<i>Indkomst for beboere, tusinde 2016-kroner</i>					
Gns. indkomst, mindst 18 år	286	228	255	264	334
Gns. indkomst, 30-59 år	346	269	306	313	413
<i>Arbejdsmarkedstilknytning for beboere på mindst 18 år</i>					
I beskæftigelse	51%	43%	47%	47%	57%
Studerende	9%	12%	7%	8%	7%
Midlertidigt udenfor	7%	12%	8%	9%	3%
Førtidspension	5%	8%	3%	6%	3%
Tilbagetrækning	22%	15%	23%	25%	23%
Øvrige	7%	10%	12%	5%	6%
<i>Øvrige karakteristika</i>					
Ikke-vestlig herkomst	29%	54%	40%	17%	20%
Dømte, ældre end 15 år	1,5%	2,3%	4,0%	1,4%	1,1%
Grundskole, 30-59 år	44%	65%	59%	44%	30%

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

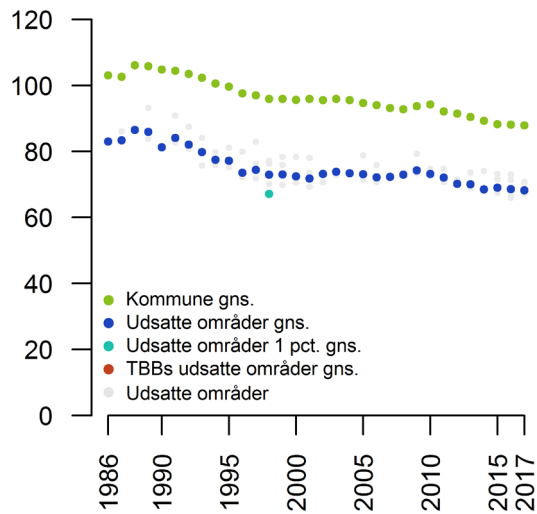
Brøndbyensere og Ishøjere i udsatte boligområder

I kommunerne Brøndby og Ishøj vurderes cirka 17.000 beboere at bo i udsatte boligområder, jf. Tabel 2.5 på foregående side. Tabellen sammenligner beboerne i udsatte områder med kommunernes samlede gennemsnit og personer bosat i ikke-udsatte boligområder. Dette svarer til, at knap 30 procent af kommunernes samlede befolkning i 2017 boede i et udsat boligområde. Det placerer Brøndby og Ishøj i toppen af listen over kommuner, der har den største andel af indbyggere i et udsat boligområde.

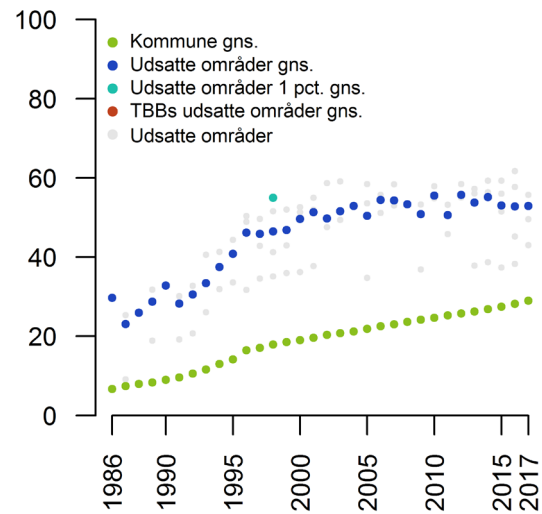
Figur 2.17 viser den historiske udvikling i Brøndby og Ishøj fra 1986 til 2017 for herkomst, det Socioøkonomiske Nabolag-Indeks og de bagvedliggende socioøkonomiske variable uddannelses- og arbejdsmarkedstilknytning, kriminalitet, uddannelsesniveaue og indkomst. Det fremgår, at SNI for kommunerne er faldende, hvorfor kommunerne i gennemsnit er blevet mere socioøkonomisk udfordrede over perioden. De udsatte områder er dog rimeligt stabile over tid. De bagvedliggende socioøkonomiske variable ses at være tilnærmelsesvis stabile, mens andelen af ikke-vestlige indvandrere og efterkommere er stigende over tid.

Figur 2.17. Beboere i udsatte områder i Brøndby og Ishøj i perioden 1986 til 2017.

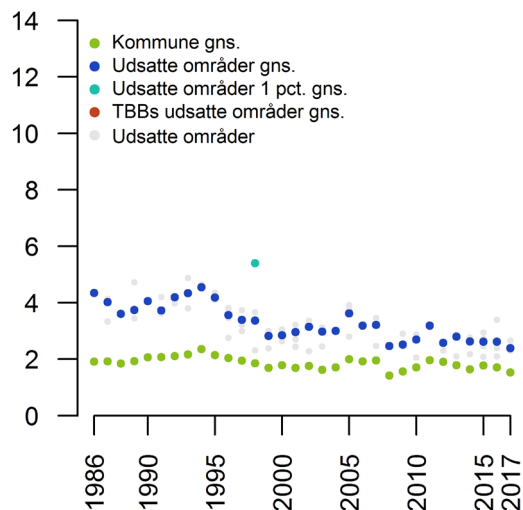
a) Indkomst iff. landsgennemsnittet, procent



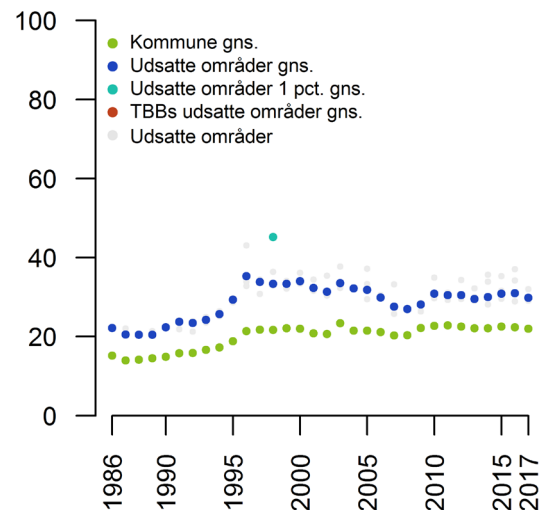
b) Ikke-vestlig herkomst, procent



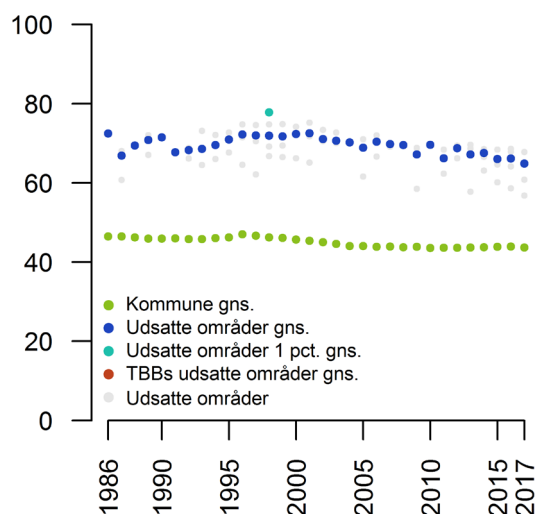
c) Kriminalitet, andel dømte i procent



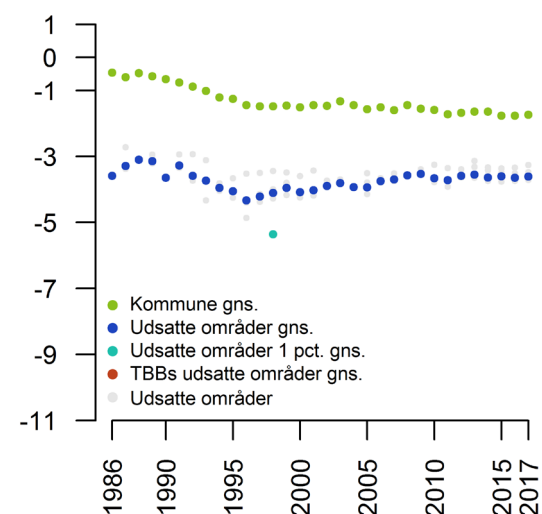
d) Uden udd. og arb.markeds-tilknytning, procent



e) Grundskole som højst fuldført niveau, procent



f) Socioøkonomisk Nabolag-Indeks



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

3. Flyttemønstre i varigt udsatte almene boligområder

I dette kapitel betragtes til- og fraflytningsmønstre i varigt udsatte almene boligområder. Der tages udgangspunkt i de udsatte boligområder udpeget i kapitel 2, dog begrænses datagrundlaget til almene boliger udpeget som værende udsatte i alle de fem seneste år. Afgrænsningen af varigt udsatte almene boligområder beskrives nærmere i delafsnit 3.1.

Flytninger kan over tid være med til at fastholde eller ændre strukturen i et boligområde. Hvis for eksempel tilflyttere er dårligere stillet end områdets eksisterende beboere, kan flyttemønstret være medvirkende til, at området forbliver udsat. Det samme vil være tilfældet, hvis personer, der fraflytter et udsat område, er områdets relativt stærke beboere. Flyttemønstret kan også betyde, at et område blive socioøkonomisk bedre stillet. Et eksempel herpå er udviklingen på Vesterbro i København gennem 1990'erne, jf. afsnit 2.4. Nærværende afsnit undersøger, hvorvidt der observeres et selektivt flyttemønster ved til- og fraflytning fra udsatte boligområder.

Undersøgelsen viser, at udskiftningen af beboere i varigt udsatte almene boligområder er cirka 17 procent årligt. Dette er på samme niveau som i øvrige almene boliger og lidt højere end i andels- og ejerlejligheder. Boligmobilitet er temaet for afsnit 3.2.

Der ses en tendens til, at tilflyttere til udsatte områder i gennemsnit har lavere indkomst og flere har grundskolen som højst fuldførte uddannelse, end det er tilfældet for beboere, som fraflytter området. Flyttemønstret er således medvirkende til at reducere det gennemsnitlige indkomstniveau i de varigt udsatte almene boligområder.

Tabel 3.1 opsummerer kapitlets overordnede resultater. Tabellen viser en række karakteristika for tilflyttere, fraflyttere og personer som bliver boende i varigt udsatte almene boligområder.

Tabel 3.1. Karakteristika for tilflyttere, fraflyttere og fastboende personer i varigt udsatte almene boligområder, gennemsnit 2012-2016.

	Fastboende	Fraflyttere	Tilflyttere	Tilflyttere, ikke indvandret
Indkomst, 30-59 år (2016-kroner)	244.000	243.000	219.000	238.000
Uden tilkn. til arb.marked eller udd., 18-64 år	42 %	33 %	33 %	30 %
Andel grundskole højst fuldført, 30-59 år	69 %	62 %	67 %	63 %
Andel dømt, 15 år eller ældre	2,4 %	4,9 %	4,5 %	5,2 %
Andel af ikke-vestlig herkomst	52 %	40 %	44 %	41 %

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

En stor andel på i gennemsnit 17 procent af tilflytterne til de betragtede udsatte boligområder indvandrer fra udlandet i tilflytningsåret. Disse personer har i gennemsnit ringere arbejdsmarkedstilknytning, er dårligere uddannet og har lavere indkomstniveau end tilflyttere til områderne, som boede i landet primo tilflytningsåret.

Betragtes udelukkende tilflyttere til varigt udsatte almene boligområder, som flytter internt i landet (kolonnen yderst til højre i tabellen ovenfor), ligner disse tilflyttere i høj grad de personer, som fraflytter områderne. Afsnit 3.3 giver en yderligere beskrivelse af, hvad der karakteriserer personer, som flytter til eller fra et varigt udsat alment boligområde. Der fokuseres på arbejdsmarkedstilknytning, uddannelsesniveau, indkomstforhold og herkomst herunder opholdstid i Danmark.

Personer, der tilflytter et varigt udsat alment boligområde, har på tilflytningstidspunktet lavere gennemsnitlig indkomst end områdets eksisterende beboere. Der ses dog en klar tendens til, at tilflytterne over få år nærmer sig indkomstniveauet for områdets øvrige beboere. Dette følger øjensynligt af, at der går noget tid før den relativt store andel af tilflyttere, som er nyindvandret, etablerer sig på arbejdsmarkedet. Indkomstudvikling efter tilflytning til et varigt udsat alment boligområde beskrives i afsnit 3.4.

3.1 Afgrænsning til varigt udsatte almene boligområder

Ved analysen af flyttemønstre i udsatte boligområder afgrænses data til udelukkende at omfatte almene boliger beliggende i boligområder, som er udpeget som udsatte i alle de fem seneste dataår, dvs. perioden 2013-2017. Disse områder betegnes varigt udsatte almene boligområder.

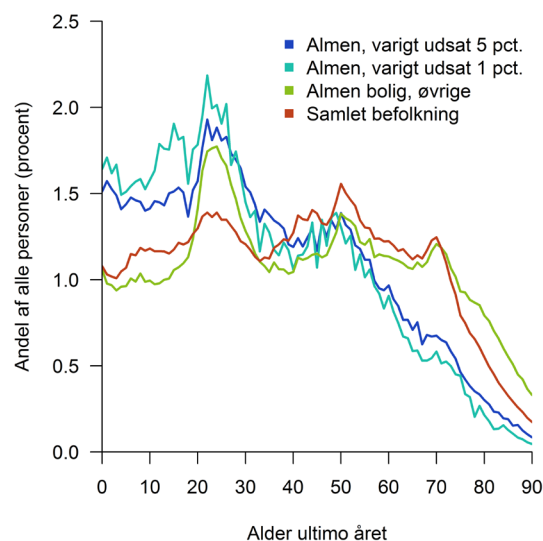
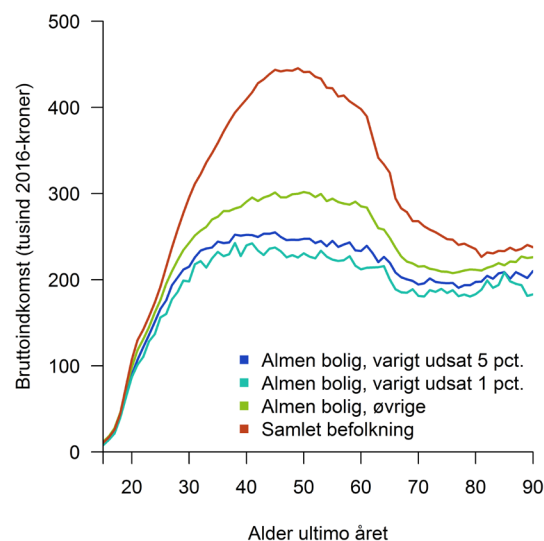
I områder karakteriseret som udsatte bor den overvejende del af befolkningen i almene boliger. Primo 2017 gælder dette knap 85 procent af de i alt 223.000 beboere i de i afsnit 2.2 identificerede udsatte boligområder. Undersøgelsen af flyttebevægelser begrænses til denne hovedgruppe, der vurderes som en relativt homogen gruppe. De resterende beboere i de udpegede udsatte boligområder omfatter hovedsageligt personer bosat i private udlejningsboliger eller i ejerboliger. Beboere i disse to boligformer udgør hver cirka fem procent af det samlede antal beboere i udsatte områder. Beboersammensætningen i udlejnings- og ejerboliger ses at være mere uensartet, end det er tilfældet for almene boliger. Til- og fraflytning til andre boligformer end almene boliger vil derfor være broget, hvorfor boligformerne udelades ved undersøgelsen af flyttebevægelser.

Ved at betragte varigt udsatte områder undgås det, at undersøgelsen påvirkes af flyttebevægelser i områder, der kun kortvarigt karakteriseres som udsatte. I begyndelsen af 2017 vurderes 167.000 beboere at være bosat i varigt udsatte boligområder. Dette udgør tre fjerdedele af de i alt 223.000 personer bosat i udsatte områder. Af beboerne i varigt udsatte områder bor knap 150.000 i almene boliger. Analysen i dette afsnit koncentrerer sig om flyttebevægelser for disse personer. Reduktionen i antallet af beboere i varigt udsatte almene boligområder er relativt ens over hele perioden 2013-2017.

Karakteristika for beboere i varigt udsatte almene boligområder

Beboerne i varigt udsatte almene boligområder er typisk forholdsvis unge, hvorimod relativt få er ældre end 60 år. Alderssammensætningen er medvirkende til, at der bor forholdsvis mange børn i områderne, jf. Figur 3.1a.

Blandt varigt udsatte almene boliger beliggende i de mest udsatte nabolag, hvor fem (en) procent af befolkningen bor, udgør beboere yngre end 40 år 62 (65) procent af beboerne. Til sammenligning er knap halvdelen af beboerne i øvrige almene boliger yngre end 40 år gamle, hvilket nogenlunde svarer til aldersgruppens andel af befolkningen som helhed. I øvrige almene boliger bor især mange unge i de studieaktive aldre, mens der bor relativt få børn. Dette skyldes en højere andel af studie- og ungdomsboliger. For de ældre over 75 år bor der relativt mange i ikke-udsatte almene boliger, da gruppen blandt andet omfatter en del ældreboliger.

Figur 3.1. Karakteristika for beboere i udvalgte boligformer, primo 2017.**a) Alderssammensætning****b) Indkomst efter alder**

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Beboerne i almene boliger har i gennemsnit lavere indkomst end befolkningen generelt, jf. Figur 3.1b. I varigt udsatte områder ses gennemsnitsindkomsten at være betydeligt lavere end i øvrige almene boliger. Blandt beboerne i varigt udsatte almene boliger beliggende i de mest udsatte nabolag, hvor den mest udsatte procent af befolkningen bor, er den gennemsnitlige årsindkomst cirka 15.000 kr. lavere end i almene boliger i de mest udsatte nabolag, hvor fem procent af befolkningen bor.

Adfærden blandt beboerne i varigt udsatte almene boligområder ønskes sammenlignet med en befolkningsgruppe, der i størst muligt omfang er sammenlignelig. Flyttemønstre blandt beboerne i de udvalgte almene områder sammenholdes således med adfærden for beboere i øvrige almene boliger, beboerne i private udlejningslejligheder og ejere af andels- samt ejerlejligheder.

Blandt de tre sidstnævnte grupper udvælges udelukkende lejligheder, da disse vurderes bedst sammenlignelige med boligerne i de varigt udsatte almene boligområder. De varigt udsatte almene boligområder er beliggende i de større byer – eller i forstæderne til de større byer – og typisk i tæt bebyggelse. Som følge heraf er der oftest tale om lejligheder eller rækkehuse. Disse to boligformer udgør henholdsvis 93,6 og 5,3 procent af boligerne i de varigt udsatte almene boligområder, mens disse på landsplan udgør 32 og 13 procent af den samlede boligbestand.

Beboerne i lejligheder udenfor den almene sektor har i gennemsnit højere indkomst end beboere i almene boliger, blandt andet som følge af en højere beskæftigelsesandel og uddannelsesniveau. I almene boligområder ses især en større andel af personer midlertidigt udenfor arbejdsstyrken og flere førtidspensionister. Det fremgår af Tabel 3.2, som sammenligner karakteristika for beboerne i varigt udsatte almene boligområder med beboerne i sammenlignelige boligformer.

Det fremgår af tabellen, at ud fra alderssammensætning og arbejdsmarkedstilknytning ligner beboerne i varigt udsatte almene områder nogenlunde beboerne i øvrige almene boliger. Dog ses beboerne i de udsatte områder i gennemsnit at have lavere indkomst, ligesom det gennemsnitlige uddannelsesniveau er lavere. Yderligere ses der i de varigt udsatte områder en større andel af dømtede.

Tabel 3.2. Karakteristika for beboere i udvalgte boligformer, primo 2017.

	Alle boligtyper			Etageboliger		
	Total befolkning	Almen, varigt udsat	Almen, øvrige	Ejer	Andel	Privat udlej.
<i>Antal personer, tusinder</i>						
Antal i alt	5.749	147	834	234	277	590
Voksne beboere	4.378	102	656	194	229	509
<i>Aldersfordeling på beboere</i>						
Hjemmeboende børn	24%	31%	21%	17%	17%	14%
Voksne yngre end 30 år	12%	16%	15%	14%	23%	37%
Voksne 30-59 år	39%	38%	35%	43%	43%	33%
Voksne 60 år eller ældre	25%	15%	28%	26%	17%	16%
<i>Indkomst for beboere, tusinde 2016-kroner</i>						
Gns. indkomst, mindst 18 år	321	208	237	404	306	260
Gns. indkomst, 30-59 år	406	243	286	484	382	346
<i>Arbejdsmarkedstilknytning for beboere på mindst 18 år</i>						
I beskæftigelse	53%	34%	35%	61%	56%	45%
Studerende	10%	13%	12%	8%	16%	23%
Midlertidigt udenfor	5%	17%	10%	3%	4%	7%
Førtidspension	4%	11%	10%	2%	2%	3%
Tilbagetrækning	22%	14%	28%	20%	15%	14%
Øvrige	6%	11%	6%	6%	7%	8%
<i>Øvrige karakteristika</i>						
Ikke-vestlig herkomst	8%	52%	20%	10%	8%	8%
Dømte, ældre end 15 år	0,9%	2,5%	1,5%	0,4%	0,7%	1,3%
Grundskole, 30-59 år	26%	68%	47%	19%	20%	34%

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

3.2 Boligmobilitet

Ved at betragte andelen af beboerne, som i løbet af et år skifter bopælsadresse, belyses boligmobiliteten i forskellige boligtyper⁶.

Af den samlede befolkning primo året er der omkring 775.000 personer, som skifter bopælsadresse i løbet af år 2016 (inklusive personer, som dør eller udvandrer). Flytteandelen, dvs. andelen af befolkningen primo perioden, som skifter bolig i løbet af året, er 13,6 procent. Andelen er nogenlunde konstant over tid, jf. Figur 3.3 nedenfor. Af det samlede antal flytninger gælder det for omkring 100.000 personer, at disse fraflytter deres bolig ved dødsfald eller som følge af fraflytning til udlandet.

⁶ En flytning mellem to boliger defineres ud fra, om en person skifter folkeregisteradresse i løbet af et kalenderår. Fraflytningsboligen registreres som personens bolig primo året, mens tilflytningsboligen er vedkommendes bolig ultimo året. Ud fra denne definition kan hver person maksimalt foretage én flytning i løbet af perioden.

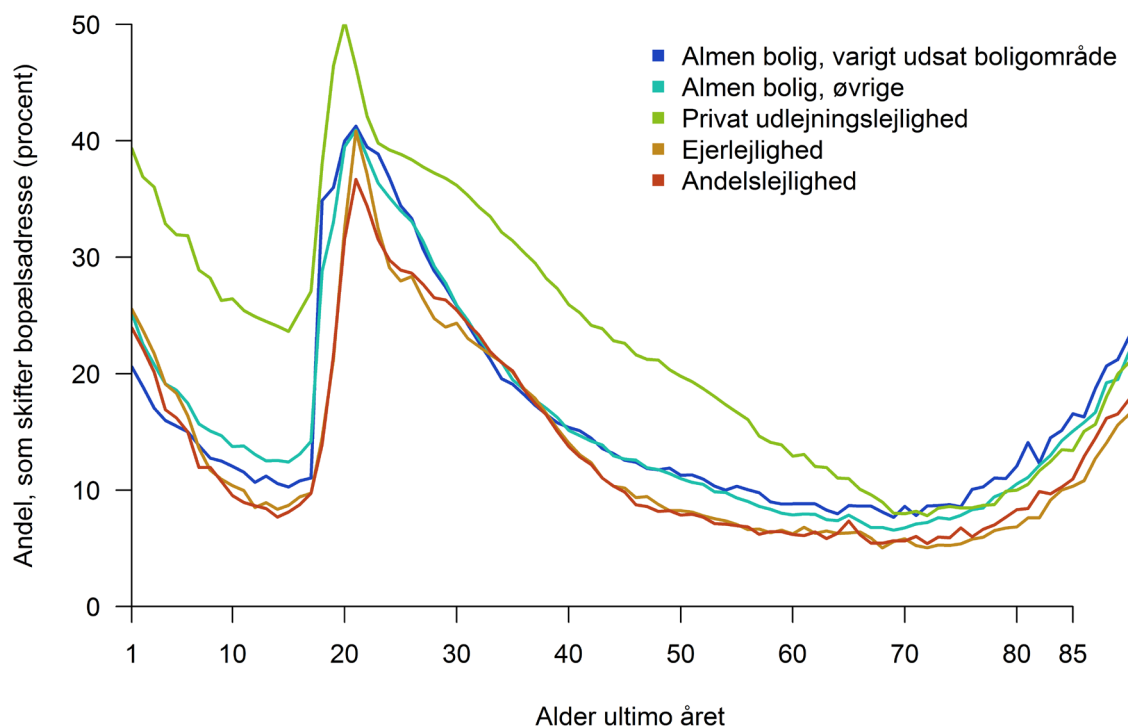
Herudover kommer tilflyttere i form af nyfødte og personer indvandret fra udlandet. Disse to grupper omfatter cirka 140.000 personer i 2016. Samlet set er det således mere end 900.000 personer, der i løbet af året skifter bolig.

Personer i 20'erne skifter oftest bolig ved fraflytning fra forældrene eller ved boligskit i forbindelse med uddannelse, stiftelse af familie mv. Efterhånden som folk bliver ældre, aftager flytteandelen, jf. Figur 3.2.

Målt ud fra flytteandelen er boligmobiliteten blandt beboere i udsatte boligområder nærmest identisk med øvrige almene områder. Blandt almene boliger i varigt udsatte områder ses en større andel, end det er tilfældet i øvrige almene boliger, at fraflytte området ved udvandring til udlandet.

Flytteandelen er generelt højere i lejeboliger end i andels- og ejerlejligheder. Mobiliteten i almene boliger er blot lidt højere end mobiliteten blandt beboere i andels- og ejerlejligheder. For lejere i privatejede udlejningslejligheder ses en betydeligt større mobilitet end i øvrige boligtyper. Dette kan følge af, at private udlejningslejligheder er relativt let tilgængelige, hvorfor dette kan være en persons foretrukne boligvalg, hvis personen søger bolig for en forholdsvis kort periode.

Figur 3.2. Flytteandel efter alder, gennemsnit 2012-2016.



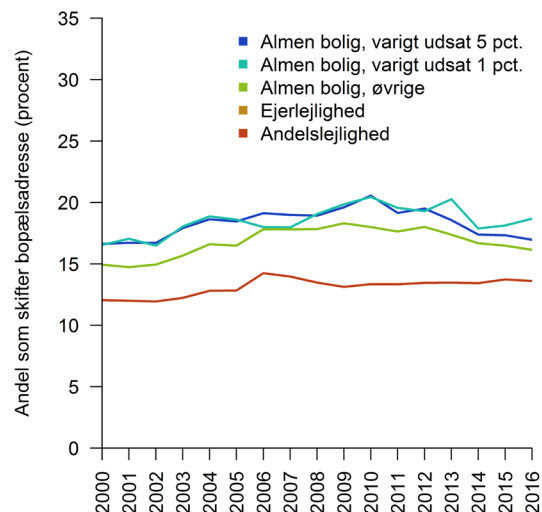
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Unge i 20'erne og 30'erne udgør en overvægt af beboerne i almene boliger, jf. Figur 3.1a. I disse alderstrin observeres en generelt høj flyttetilbøjelighed. Denne kombination gør, at flytteandelen generelt er højere i almene boliger end i øvrige boligformer. Blandt varigt udsatte almene boliger var flytteandelen i 2016 på cirka 17 procent, hvilket er en anelse højere end blandt øvrige almene boliger, hvor 16 procent af beboerne flyttede i løbet af 2016. I hele befolkningen var den gennemsnitlige flyttetilbøjelighed på 13,6 procent. Over tid ses flyttetilbøjeligheder for både den samlede befolkning, og blandt beboerne i almene boligområder, at være ganske konstant, jf. Figur 3.3a. Dog ses mindre udsving i

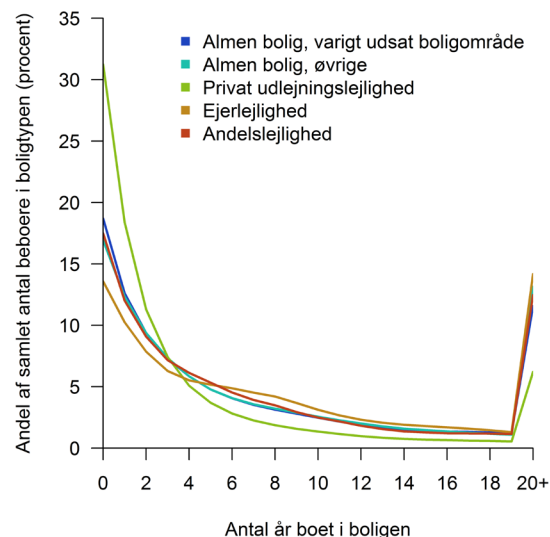
flyttetilbøjeligheden for beboere i almene boliger, i disse tilfælde følger varigt udsatte og øvrige almene boligområder samme tendens.

Figur 3.3. Flytteandel over tid og antal år boet i nuværende bolig.

a) Flytteandel over tid



b) Antal år i nuværende bolig, gns. 2012-2016



Anm.: Hjemmeboende børn er udeladt af figuren til højre.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Beboerudskiftningen i almene boliger beliggende i varigt udsatte almene områder er på niveau med øvrige almene boliger og blot lidt lavere end andels- og ejerlejligheder. Som et supplement hertil betragtes det, hvor lang tid en person i gennemsnit er bosat i et varigt udsat område.

Beboerne i almene boliger bor nogenlunde lige så længe i deres bolig som beboerne i sammenlignelige boligtyper. Det fremgår af Figur 3.3b, som for forskellige boligtyper viser, hvor længe voksne beboere har boet i deres nuværende bolig. Fordelingen af antal år bosat i boligen er stort set ens for øvrige boligformer end private udlejningsboliger, hvorfra flytninger sker betydeligt oftere end blandt de øvrige boligtyper. Dette stemmer overens med de observerede flyttetilbøjeligheder i Figur 3.2. Dette illustrerer, at beboerne i varigt udsatte almene boligområder i gennemsnit bor i deres bolig i samme antal år som beboerne i sammenlignelige boligtyper.

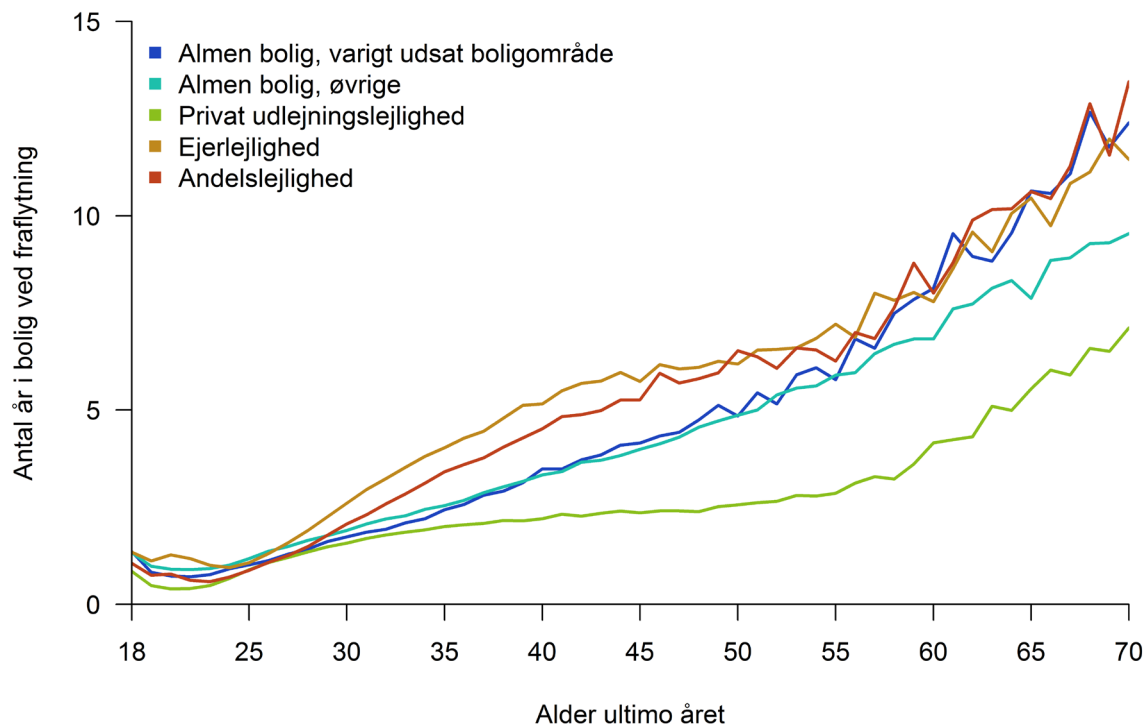
Slutteligt ses der på, hvor lang tid en person, som fraflytter en bolig, i gennemsnit har boet i boligen. Personer i 20'erne, som fraflytter og ikke er opvokset i et varigt alment udsat boligområde, har i gennemsnit boet i området i 2-4 år på fraflytningstidspunktet, jf. Figur 3.4. Dette er på niveau med andre boligformer. Dog ses en tendens til, at personer under 35 år i gennemsnit bor lidt kortere tid i boligen ved fraflytning fra varigt udsatte almene boligområder, end det er tilfældet ved fraflytning fra øvrige almene boligområder.

Fra slutningen af 20'erne ses beboere i både varigt udsatte og ikke-udsatte almene boliger at have boet kortere tid i deres bolig ved fraflytning end beboere i ejer- eller andelslejligheder. Fra omkring 60-årsalderen ligger varigt udsatte almene boliger på niveau med andels- og ejerlejligheder. Ved fraflytning har beboerne i private udlejningslejligheder boet betydeligt kortere tid i deres bolig end de øvrige boligformer.

Man skal være opmærksom på, at Figur 3.4 ikke udtrykker, hvor længe en person i en given alder kan forventes at bo i en given boligtype. At for eksempel 25-årige blot har boet få år i en

bolig ved fraflytning, det betyder ikke, at 25-årige i gennemsnit bor kort tid i deres bolig. Dem, som bor relativt lang tid i boligen, er blot ældre på fraflytningstidspunktet, og er således med til at øge det gennemsnitlige antal år i fraflytningsboligen for ældre personer. Dette er hovedforklaringen bag den i alder stigende tendens observeret i figuren nedenfor.

Figur 3.4. Fraflyttere efter alder og antal år i fraflytningsbolig, gennemsnit 2012-2016.



Anm.: Figuren indeholder ikke personer, som er opvokset i boligområdet. Denne gruppe er defineret til personer, som senest tilflytter området som 13-årige. Figuren indeholder således fraflyttere, som tidligst er tilflyttet boligen som 14-årige.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

3.3 Flyttemønstre i varigt udsatte almene boligområder

I dette afsnit betragtes flyttemønstre til de varigt udsatte almene boligområder beskrevet i afsnit 3.1. Der fokuseres på flytninger til og fra disse områder i løbet af de seneste fem dataår, dvs. perioden 2012-2016.

Undersøgelsen af flyttemønstre baseres således på almene boligområder, der er karakteriseret som udsatte i alle de fem seneste år. Karakteristika for personer, som flytter til og fra disse områder, betragtes tilbage til år 2000. Ved opgørelse af historiske serier er afgrænsningen således områder, der er udsatte i de seneste år, og ikke de områder, som er karakteriseret som udsatte løbende over perioden. Dog er områderne i vidt omfang rimelig konstante i perioden siden årtusindeskiftet.

Antallet af nye beboere i de varigt udsatte almene områder svinger mellem 21.500 og 23.000 i den betragtede femårige periode. Dette omfatter omkring 16.000-17.500 personer, som årligt tilflytter områderne, godt 2.000 nyfødte per år og i størrelsesordenen 3.500-4.000 nyindvandrede personer per år, som bosætter sig i områderne. Nogenlunde samme antal personer forlader områderne, hvoraf cirka 4.000 personer udtræder af befolkningen ved udvandring eller dødsfald.

I den resterende del af dette afsnit betragtes det, hvorvidt flyttebevægelser ind og ud af de varigt udsatte almene områder er med til at fastholde områderne som udsatte, eller er flyttebevægelserne med til at forbedre de udsatte områders socioøkonomiske status. Dette undersøges ved at sammenligne karakteristika på til- og fraflyttere samt fastboende i områderne. Foruden de fire socioøkonomiske segregeringsvariable, som er gennemgående i rapporten, fokuseres der på beboernes alder og herkomst herunder indvandreres opholdstid i Danmark.

Alder

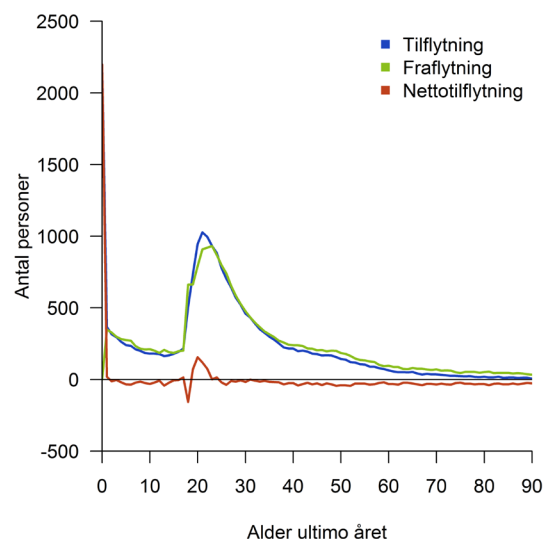
Over et år bliver antallet af fraflyttere fra varigt udsatte almene boligområder modsvaret af et tilsvarende antal tilflyttere, hvorfor nettotilflytningen til varigt udsatte almene boligområder er tæt ved nul. Aldersfordelingen på tilflytterne svarer nogenlunde til aldersfordelingen på de beboere, som fraflytter områderne, jf. Figur 3.5.

I almene boliger beliggende i nabolag, hvor de fem procent mest udsatte bor, sker en nettofraflytning på cirka 425 børn under 18 år. Dette opvejes af, at cirka 2.200 nyfødte tilgår områderne. I første halvdel af 20'erne sker en nettotilflytning på cirka 350 personer. Herefter er nettobevægelsen svagt negativ frem mod slut 30'erne, hvorefter der igen sker en tydelig nettofraflytning. For beboere ældre end 35 år er nettofraflytningen samlet set på cirka 1.850 personer.

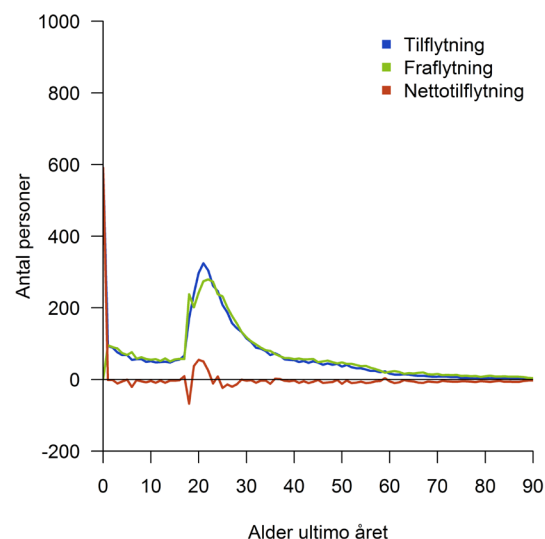
Flyttemønstret i varigt udsatte almene boliger beliggende i nabolag, hvor den mest udsatte procent af befolkningen er boende, ligner flyttebevægelserne i tilsvarende områder, hvor de fem procent mest udsatte bor. Således sker en nettotilflytning af nyfødte og personer i første halvdel af 20'erne, mens der i alle øvrige alderstrin er en svag nettofraflytning.

Figur 3.5. Flytninger i varigt udsatte almene boligområder efter alder, gennemsnit 2012-2016.

a) 5 procent mest udsatte områder



b) 1 procent mest udsatte områder

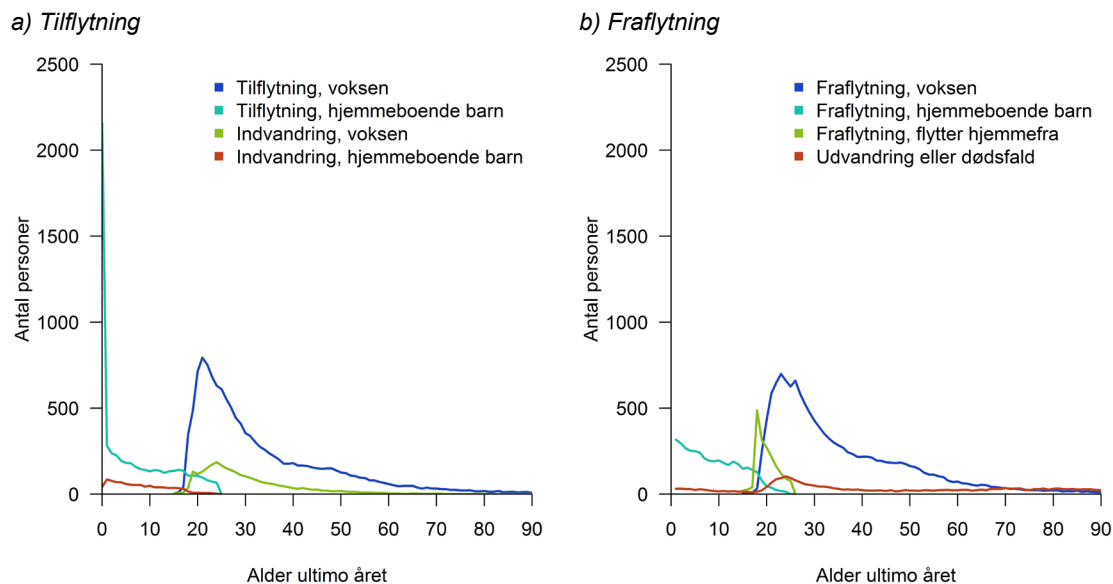


Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Opdeles flyttebevægelserne til og fra varigt udsatte almene områder efter årsag, ses det, at der for børn under 18 år sker lidt større fraflytning end tilflytning. Dette opvejes dog delvis af flere indvandrede børn end udvandrede, jf. Figur 3.6. Over det 10-årige aldersinterval fra 18 til 27 år sker en nettotilflytning af voksne beboere til varigt udsatte almene områder. Denne bevægelse modsvarer til dels af, at en del hjemmeboende børn i samme alderstrin flytter

hjemmefra til et andet boligområde. Desuden ses en betydende nettoindvandring til områderne, idet antallet af indvandrere i alle alderstrin overstiger antallet af udvandrede.

Figur 3.6. Flytninger i varigt udsatte almene boligområder efter alder og flytteårsag, gennemsnit 2012-2016.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Herkomst og opholdstid

Over de seneste årtier er indvandrere og efterkommeres andel af den samlede befolkning steget. Fra at udgøre omkring 7 procent af befolkningen i begyndelsen af 2000'erne stiger andelen jævnt til godt 13 procent af befolkningen i 2018. Særligt personer fra ikke-vestlige lande udgør i dag en større andel af befolkningen, end det tidligere var tilfældet.

Den ændrede herkomstsammensætning i befolkningen har haft afsmittende effekt på herkomstsammensætningen i de varigt udsatte almene boligområder. Andelen af beboere med ikke-vestlig herkomst stiger således fra knap 40 procent omkring årtusindeskiftet til lidt over halvdelen i 2016, jf. Figur 3.7a. Stigningen skyldes, at der særligt først i 2000'erne observeres en noget større andel af personer af ikke-vestlig herkomst blandt tilflytterne end blandt fraflytterne i de varigt udsatte almene boligområder.

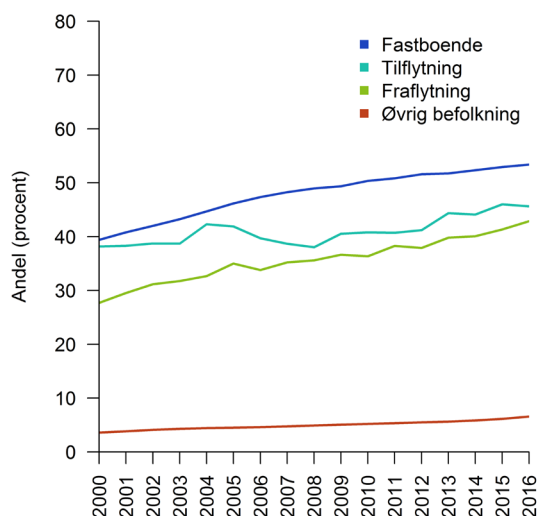
I de yngre alderstrin, hvor en del hjemmeboende børn flytter hjemmefra, er en større andel af fraflyttere end tilflyttere af ikke-vestlige herkomst. For de ældre – og særligt for flyttende i 30'erne – forholder det sig modsat, jf. Figur 3.7b.

Tabel 3.3 viser herkomstsammensætningen blandt personer, der flytter til eller fra et varigt udsat alment boligområde. Som i områderne generelt udgør personer af dansk herkomst den største herkomstgruppe blandt personer, som flytter.

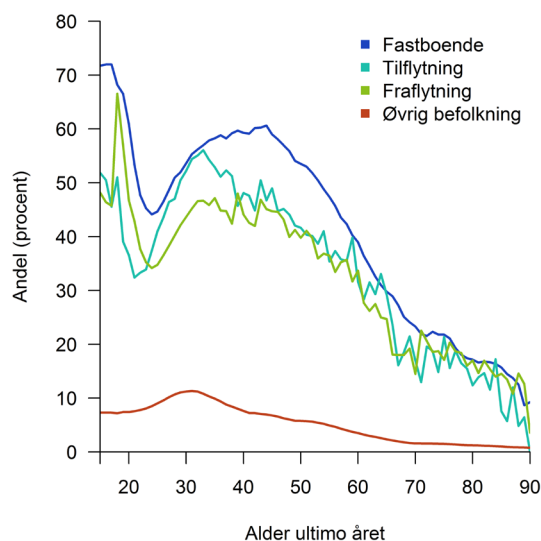
Blandt fraflyttere udgør personer af dansk herkomst en lidt større andel, end det er tilfældet for tilflyttere. Betragtes tilflyttere, som bor i Danmark primo året, ses andelen af tilflyttere af dansk herkomst dog at matche andelen blandt fraflyttere. Herkomstsammensætningen blandt nyindvandrede personer, som tilflytter varigt udsatte almene boligområder, har således været med til at øge andelen af personer af ikke-vestlig herkomst i områderne. Personer af vestlig herkomst ses at udgøre en relativt stor andel af de flyttende både ind og ud af områderne, i forhold til disse gruppers andel af befolkningen som helhed.

Figur 3.7. Andel med ikke-vestlig herkomst for tilflyttere, fraflyttere og fastboende.

a) Varigt udsatte almene områder



b) Gns. andel efter alder, gns. 2012-2016



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Tabel 3.3. Tilflyttere, fraflyttere og fastboende i varigt udsatte almene områder efter herkomst, andel i procent, gennemsnit 2012-2016.

	Dansk opr.	Indvandrere		Efterkommere		Antal Gns. per år
		Ikke-vestlige	Vestlige	Ikke-vestlige	Vestlige	
Samlet befolkning	89	5	4	2	0,4	5.494.000
Fastboende	43	31	4	21	0,7	119.000
Tilflyttere	45	27	10	17	1,1	22.000
- I Danmark primo året	52	23	5	19	1,3	18.300
- Indvandret fra udlandet	10	50	33	7	0,2	3.700
Fraflyttere	51	25	8	15	0,7	22.000

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Andelen af tilflyttere i hver herkomstgruppe er nogenlunde konstant gennem den betragtede periode, hvor antallet af indvandrere fra særligt ikke-vestlige lande øges. Fra 2012 til 2015 sker således en fordobling i det samlede antal indvandrere til Danmark i denne herkomstgruppe. Andelen af det samlede antal tilflytninger til varigt udsatte almene områder, som udgøres af indvandrere fra ikke-vestlige lande, øges dog kun svagt fra 25,5 procent i 2012 til 28,6 procent i 2015.

En relativt stor andel af tilflyttere til varigt udsatte almene boligområder er personer, som indvandrer til Danmark i løbet af tilflytningsåret. Gennem den betragtede femårige periode udgør personer indvandret fra udlandet således knap 17 procent af det samlede antal tilflyttere til områderne. Andelen varierer svagt gennem perioden. I årene 2013-2015 observeres et generelt stigende antal indvandring fra særligt ikke-vestlige lande. Andelen af tilflytninger til varigt udsatte områder, som er indvandret fra udlandet, stiger en smule i disse år fra 15 procent i 2012 til 17-18 procent i 2013-2015. Antallet af indvandrere topes i 2015, hvorefter der ses en aftagende tendens. Andelen af tilflyttere, som indvandrer fra udlandet, falder til 16 procent i 2016.

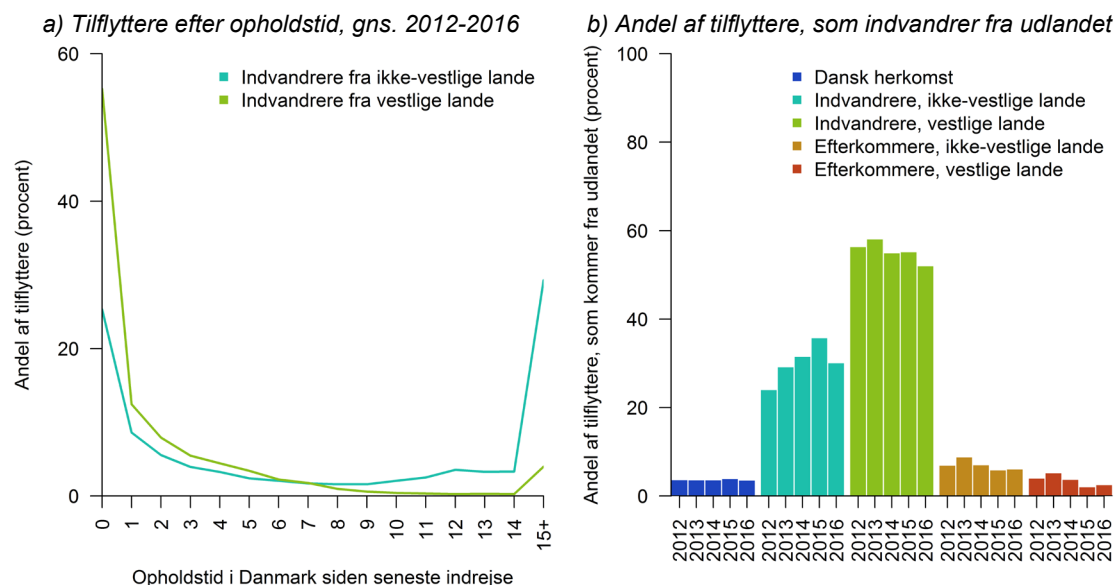
Udviklingen skyldes, at der i årene med generelt høj indvandring er en større andel af tilflytterne fra ikke-vestlige lande, som indvandrer direkte fra udlandet. Det fremgår af Figur 3.8b, der viser, hvor stor en andel af det årlige antal tilflyttere i hver herkomstgruppe, som er indvandret fra udlandet i løbet af tilflytningsåret. For de øvrige herkomstgrupper ses andelen nogenlunde konstant over perioden.

En persons opholdstid i Danmark opgøres som opholdstiden i landet siden personens seneste indrejse. Opholdstid opgøres som hele år, hvor nul angiver, at personen på opgørelses tidspunktet ultimo året har opholdt sig mindre end ét kalenderår i Danmark, værdien én viser, at der er gået mellem ét og to år siden personens seneste indvandring osv.

I perioden 2012-2016 kom der i gennemsnit 25.000 indvandrere fra ikke-vestlige lande årligt, som tager ophold i Danmark. Heraf bosætter 7,5 procent sig i et varigt udsat alment boligområde. Dette betyder, at blandt indvandrere fra ikke-vestlige lande, som tilflytter et varigt udsat alment boligområde, er hver fjerde person nyindvandret, jf. Figur 3.8a. Mere end 40 procent af tilflytterne i denne herkomstgruppe har en opholdstid i Danmark på færre end fem år.

En anden stor gruppe af indvandrere fra ikke-vestlige lande, som tilflytter varigt udsatte boligområder, har opholdt sig i Danmark i mindst 15 år. Disse udgør knap 30 procent af alle tilflyttere i herkomstgruppen.

Figur 3.8. Tilflyttere til varigt udsatte almene områder efter opholdstid og andel, som indvandrer fra udlandet, 2012-2016.



Anm.: Figuren til venstre indeholder personer, som ultimo tilflytningsåret er mindst 18 år gamle. Opholdstid nul angiver, at personen har opholdt sig mindre end ét år i Danmark, og dermed er indvandret i løbet af tilflytningsåret. Opholdstid ét angiver, at personen har været [1, 2] år i landet, opholdstid to [2, 3] år osv. Gruppen 15+ angiver en opholdstid på mindst 15 år.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Ovenstående indikerer, at tilflytningsmønstret til varigt udsatte almene boligområder i nogen grad er påvirket af det generelle indvandringsomfang til Danmark. Tilflytninger direkte fra udlandet ser dog kun ud til at stige forholdsvis svagt i en periode, hvor antallet af indvandrere fra ikke-vestlige lande fordobles på landsplan. Den betragtede femårige periode anses derfor ikke som værende voldsomt forskellig fra det generelle billede. De varigt udsatte almene områder har i en længere årrække været præget af en relativt stor tilflytning af nyindvandrede

personer. Set over hele perioden siden årtusindeskiftet ses mere end hver femte tilflytter i herkomstgruppen indvandrere fra ikke-vestlige lande således at være nyindvandret fra udlandet. Dette er således ikke et fænomen, som udelukkende er koncentreret omkring perioden 2012-2016, hvor indvandringsomfanget var usædvanligt højt i forhold til det generelle billede siden årtusindeskiftet.

Indkomst

Beboere, som fraflytter et varigt udsat alment boligområde i en given alder, har i gennemsnit samme indkomst som områdets resterende beboere af samme alder⁷. Tilflyttere til områderne har derimod i gennemsnit en lavere indkomst end gennemsnitsindkomsten for områdets eksisterende beboere. Dog er tilflyttere sidst i 50'erne eller ældre indkomstmæssigt tæt ved de personer, som allerede er bosat i boligområderne. Det fremgår af Figur 3.9b, som sammenligner indkomsten for personer, der flytter til, fra eller bliver boende i et varigt udsat alment boligområde.

Tendensen til, at tilflyttere til varigt udsatte almene boligområder i gennemsnit har et lavere indkomstniveau end beboere, som fraflytter samme boligområder, er ganske klar gennem hele perioden siden årtusindeskiftet, jf. Figur 3.9a. Flyttemønstret er således medvirkende til at reducere indkomstniveauet i de varigt udsatte almene boligområder.

Personer, som bliver boende i de betragtede boligområder, er i gennemsnit ældre end personer, der flytter. Da indkomsten udviser en stigende tendens til man aldersmæssigt befinder sig i midten af 40'erne, betyder denne forskel i gruppernes alderssammensætning, at fastboende i de erhvervsaktive aldre i gennemsnit har højere indkomst end de, som fraflytter, jf. Figur 3.9a.

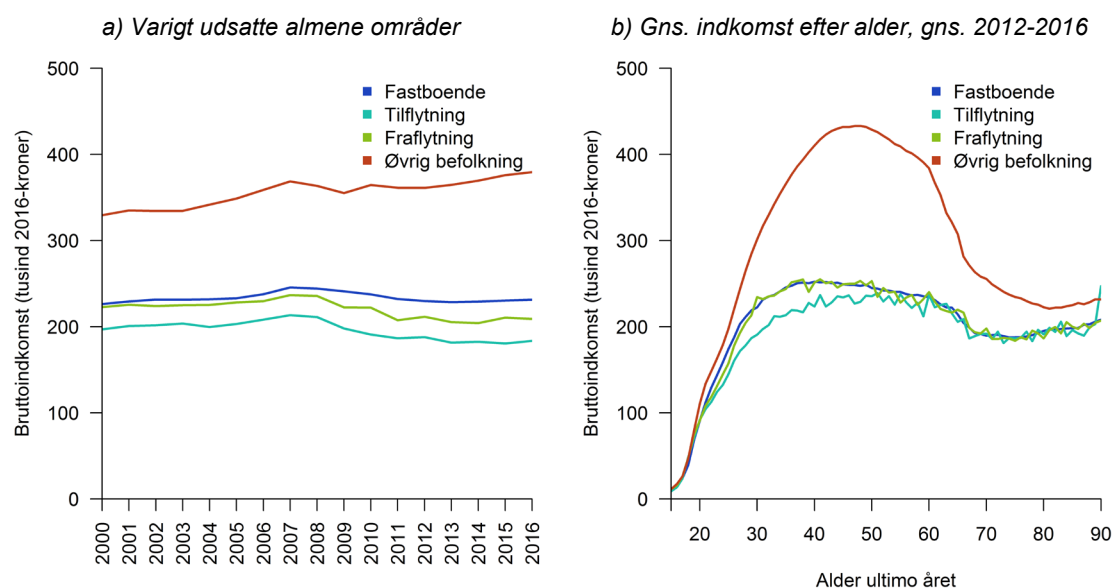
Ser man bort fra de studieaktive aldre, har beboerne i varigt udsatte almene boligområder generelt et betydeligt lavere indkomstniveau end den øvrige befolkning.

At tilflyttere til de betragtede områder i gennemsnit har lavere indkomst end øvrige personer i områderne skyldes en forholdsvis høj tilflytning af personer, som indvandrer fra udlandet. Personer indvandret fra udlandet har typisk et relativt lavt indkomstniveau, hvilket dog afhænger af alder og herkomst.

Indvandrere i 20'erne har lavt indkomstniveau, da mange kommer til landet for at studere. Skelnes på herkomstgrupper ses der tydelig forskel på personer, som genindvandrer og personer, som er helt nye i landet. For efterkommere, som genindvandrer til Danmark, ses et gennemsnitligt indkomstniveau, der er på niveau med personer i samme alder, som allerede opholder sig i landet. Særligt for indvandrere fra ikke-vestlige lande ses et relativt lavt indkomstniveau i indvandringsåret. Gennemsnitsindkomsten for denne gruppe er således lavere end 15.000 kr. per måned, for personer yngre end 30 år også noget lavere.

For at gøre indkomsten for indvandrere sammenlignelig med personer, som har fast bopæl i landet, opregnes den observerede indkomst for alle indvandrere svarende til, at personen havde opholdt sig i landet hele året.

⁷ I kapitlet fokuseres der på personindkomster. Man kan argumentere for, at husstandsindkomst – eller gennemsnitlig indkomst per voksne i husstanden – ville være et mere retvisende indkomstbegreb. Par, hvor den ene er i vellønnet beskæftigelse og partneren er hjemmegående eller på nedsat tid, kan være relativt velstående. I dette tilfælde udtrykker partnerens indkomstniveau, som er relativt lavt, ikke nødvendigvis personens reelle forbrugsmuligheder. At der alligevel fokuseres på personindkomster skyldes, at identificeringen af udsatte boligområder typisk sker på baggrund heraf.

Figur 3.9. Gennemsnitlig indkomst for tilflyttere, fraflyttere og fastboende.

Anm.: Indkomst er opgjort som den gennemsnitlige bruttoindkomst i 2016-kroner. For personer, som kun har opholdt sig i landet en del af året – dvs. er indvandret, udvandret eller døde – er personens indkomst opregnet til en årsindkomst.

I figuren til venstre medregnes skattepligtige i alderen 15-64 år eksklusive uddannelsessøgende. I figuren til højre inkluderes alle personer inklusiv studerende.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

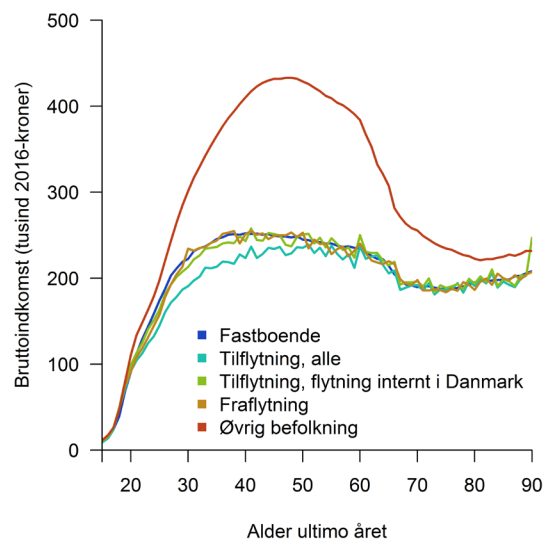
Tilflyttere til varigt udsatte almene boligområder opdeles efter, om personen opholdte sig i landet primo året, eller om vedkommende er indvandret fra udlandet i tilflytningsåret. Det ses af Figur 3.10, at tilflyttere til varigt udsatte almene boligområder, som flytter internt i Danmark, i gennemsnit har tæt ved samme indkomst som personer, der allerede bor i boligområdet. Dette gælder for indvandrere fra både vestlige og ikke-vestlige lande.

For personer af dansk herkomst og for efterkommere udgør indvandring fra udlandet en så beskeden andel af det samlede antal tilflytninger, at nyindvandrede personer ikke i nævneværdigt omfang kan påvirke den gennemsnitlige adfærd. For efterkommere ses nyindvandrede personer desuden at have en gennemsnitlig indkomst, som er relativt tæt ved de personer af samme herkomstgruppe, som i forvejen bor i det varigt udsatte almene boligområde.

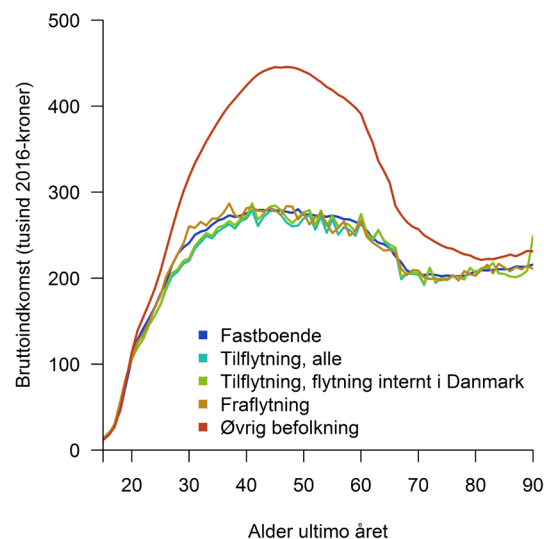
Det kan på denne baggrund konkluderes, at forskellen i den gennemsnitlige indkomst mellem personer, der flytter fra og til et varigt udsat alment område skyldes, at der blandt tilflytterne er en relativt høj andel af personer indvandret fra udlandet.

Figur 3.10. Gennemsnitlig indkomst for tilflyttere, fraflyttere og fastboende efter herkomst, gennemsnit 2012-2016.

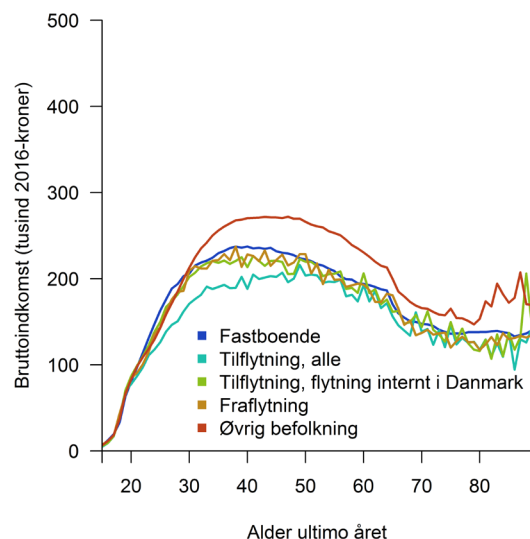
a) Alle herkomstgrupper



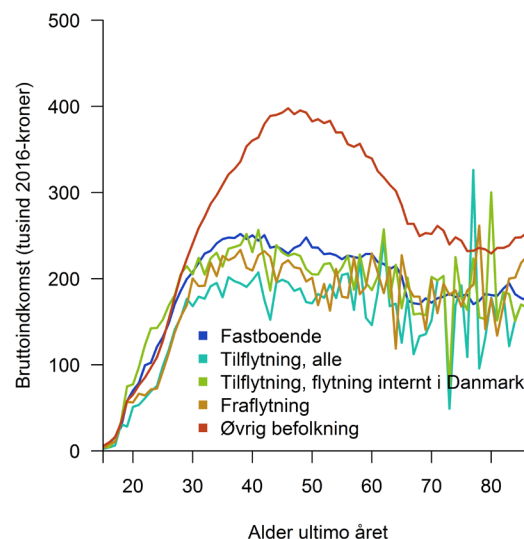
b) Dansk herkomst



c) Indvandrere fra ikke-vestlige lande



d) Indvandrere fra vestlige lande



Anm.: Indkomst er opgjort som den gennemsnitlige bruttoindkomst i 2016-kroner. For personer, som kun har opholdt sig i landet en del af året – dvs. er indvandret, udvandret eller døde – er personens indkomst opregnet til en årsindkomst.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Arbejdsmarkedstilknytning

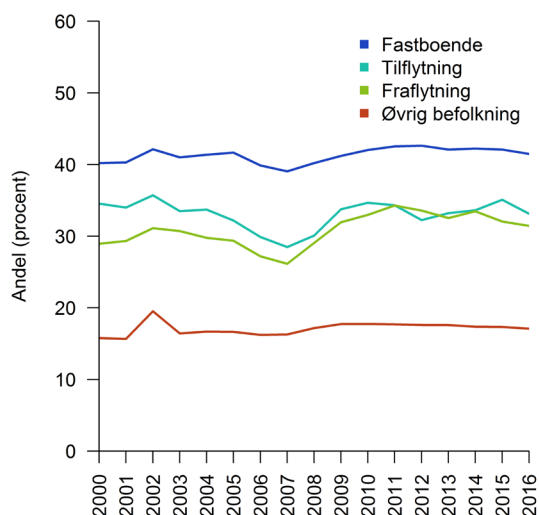
Andelen af beboere i varigt udsatte almene områder, som ikke har tilknytning til hverken arbejdsmarkedet eller er under uddannelse, har været relativt konstant på cirka 40 procent siden årtusindeskiftet, jf. Figur 3.11a. Under højkonjunktoren 2005-2008 ses andelen af beboere uden tilknytning til arbejdsmarked eller uddannelse at aftage, hvorefter andelen stiger under den efterfølgende krise.

Samlet set ses fastboende i områderne at have ringere tilknytning til arbejdsmarkedet eller uddannelse, end det er tilfældet blandt personer, som flytter. Dette skyldes hovedsageligt, at personer, der flytter, er yngre end fastboende. Betragtes andelen uden tilknytning til arbejdsmarked eller uddannelsessystemet efter alder, ses der således ikke den store forskel mellem fastboende og flyttende personer, jf. Figur 3.11b. Dog ses tilflyttere i sidste halvdel af

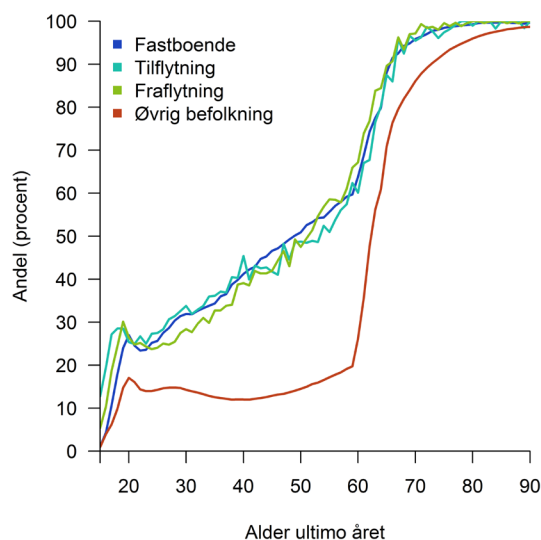
20'erne og i 30'erne at have ringere tilknytning til arbejdsmarkedet og uddannelsessystemet end de personer, som fraflytter områderne. Flyttemønstret er således medvirkende til at fastholde en relativt høj andel af beboerne uden tilknytning til arbejdsmarked eller uddannelse.

Figur 3.11. Andel uden tilknytning til arbejdsmarkedet eller uddannelse for tilflyttere, fraflyttere og fastboende.

a) Varigt udsatte almene områder



b) Gns. andel efter alder, gns. 2012-2016



Anm.: Andelen af personer uden tilknytning til arbejdsmarked eller uddannelse er i figuren til venstre opgjort for personer i alderen 18-64 år.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Set over perioden 2012-2016 er to tredjedele af både til- og fraflyttere i varigt udsatte almene boligområder enten i beskæftigelse, er ledige eller er i gang med en uddannelse, jf. Tabel 3.4. Blandt tilflyttere ses en lidt større andel af studerende og ledige end blandt fraflyttere. Dette modsvarer af, at der sker en større fraflytning af beskæftigede beboere med en relativt høj indkomst⁸, mens beskæftigede tilflyttere typisk har lavere indkomstniveau end de personer, som fraflytter. Af tabellen fremgår det desuden, at der blandt personer på offentlig forsørgelse ses en relativt lav flyttetilbøjelighed.

At tilflyttere, der aldersmæssigt befinder sig i 20'erne eller 30'erne, har lidt svagere tilknytning til arbejdsmarked og uddannelse skyldes hovedsageligt, at der for de yngre blandt disse alderstrin observeres en relativt stor fraflytning af studerende, mens der for beboere i 30'erne ses tendens til øget fraflytning af beskæftigede med relativt højt indkomstniveau.

⁸ Ved inddeling af beskæftigede i indkomstgrupper anvendes medianindkomsten for beskæftigede. Medianindkomsten er det beløb, hvor halvdelen af de beskæftigede har en lavere indkomst og halvdelen en højere indkomst. Medianen er altså udtryk for indkomsten for en beskæftiget, der ligger lige midt i indkomstspektret. Højindkomstgruppen defineres til personer med en indkomst på mere end 40 procent over medianindkomsten for en beskæftiget. Lavindkomst defineres til en indkomst lavere end 60 procent af medianindkomsten for en beskæftiget. Øvrige beskæftigede tilhører mellemindkomstgruppen.

Tabel 3.4. Socioøkonomisk status for tilflyttere, fraflyttere og fastboende i varigt udsatte almene områder, andel i procent, gennemsnit 2012-2016.

	Fastboende	Tilflytning	Fraflytning	Øvrig befolkning
Tilknytning til arbejdsmarked eller uddannelse	58	67	67	83
- Beskæftiget, højindkomst	1,3	0,8	1,4	13
- Beskæftiget, mellemindkomst	29	18	23	47
- Beskæftiget, lavindkomst	8	13	10	7
- Studerende	13	28	27	13
- Ledig	7	7	6	3
Uden tilknytning til arbejdsmarked eller udd.	42	33	33	17
- Midlertidigt udenfor arbejdsstyrken	15	12	11	3
- Førtdispension	16	5	7	6
- Tilbagetrækning	1,0	0,2	0,3	2
- Øvrige	10	16	14	6

Anm.: Tabellen indeholder personer i alderen 18-64 år.

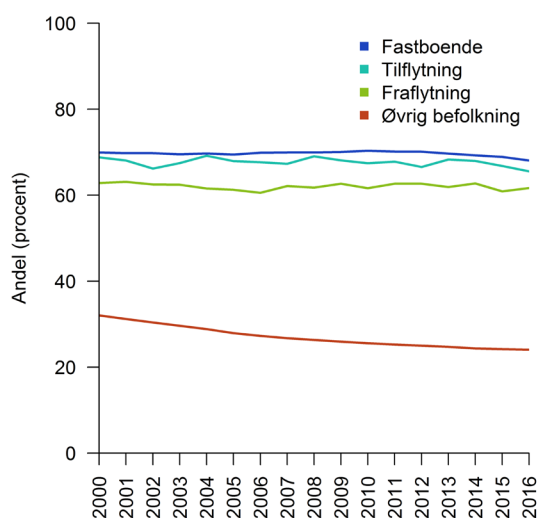
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Uddannelsesniveau

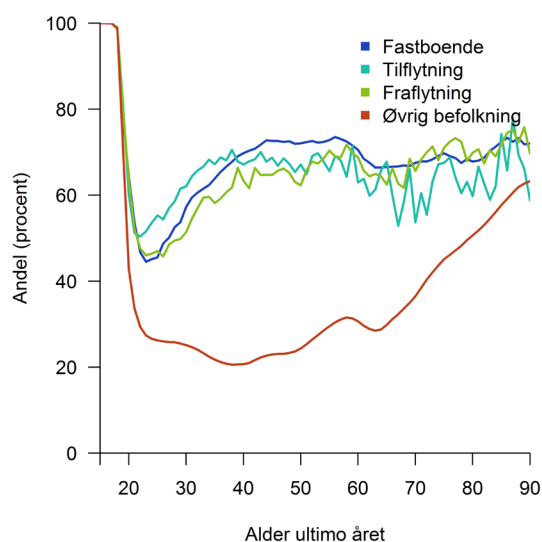
Måles uddannelsesniveaet blandt en gruppe personer ud fra andelen af disse, som ikke har fuldført en uddannelse ud over grundskolen, ses fraflyttere fra varigt udsatte almene boligområder i gennemsnit at have højere uddannelsesniveau end både fastboende i boligområderne og de personer, der tilflytter, jf. Figur 3.12a. Tendensen ses at være ganske klar over hele den betragtede periode. Flyttemønstret medvirker således til at fasholde andelen af ufaglærte i varigt udsatte almene boligområder på et relativt højt niveau.

Figur 3.12. Andel med grundskole som højst fuldførte uddannelse for tilflyttere, fraflyttere og fastboende.

a) Varigt udsatte almene områder



b) Gns. andel efter alder, gns. 2012-2016



Anm.: Andelen af personer, som har grundskolen inklusiv uoplyst uddannelsesniveau som vedkommendes højst fuldførte uddannelse er i figuren til venstre opgjort for personer i alderen 30-59 år.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Fordelt efter alder ses en større andel af tilflyttere i alderstrinnene mellem 25 og 50 år, som ikke har opnået en uddannelse udover grundskolen, end det er tilfældet for personer, som fraflytter områderne, jf. Figur 3.12b. Dette skyldes primært, at en forholdsvis stor andel af tilflytterne er personer, som indvandrer fra udlandet. Relativt mange af disse personer registreres med uoplyst uddannelsesniveau, der traditionelt medregnes blandt ufaglærte. Som følge heraf øges andelen af tilflyttere, som ikke har opnået en uddannelse udover grundskolen betydeligt.

Betragtes udelukkende tilflyttere til varigt udsatte almene boligområder, som primo tilflytningsåret havde folkeregisteradresse i landet, reduceres andelen af ufaglærte blandt tilflytterne med omkring 4 procentpoint til 63,0 procent. Dette er på niveau med andelen blandt personer, som fraflytter områderne, for hvem andelen er 62,0 procent.

I Tabel 3.5 sammenholdes uddannelsesniveaet blandt personer i varigt udsatte almene boligområder med fordelingen blandt øvrige personer. Det fremgår, at der i de udsatte områder er en generel underrepræsentation af erhvervsuddannede samt personer med videregående uddannelser. Desuden er der en væsentligt større andel med uoplyst uddannelsesbaggrund sammenlignet med befolkningen generelt.

I denne undersøgelse er det valgt at rense uddannelsesdata for alle selvrapporterede og skønnede data. Herved anvendes samme fremgangsmåde som Transport-, Bygnings- og Boligministeriet anvender ved opgørelse af udsatte boligområder. I den højre del af Tabel 3.5 vises uddannelsesfordelingen såfremt selvreporterede og imputerede uddannelsesoplysninger anvendes.

Eksklusion af skønnede data øger forskellen i det målte uddannelsesniveau for beboerne i varigt udsatte boligområder og den resterende befolkning. Dette skyldes, at andelen af indvandrede i de varigt udsatte almene boligområder er høj, og at der særligt er usikkerhed om uddannelsesniveauet for denne befolkningsgruppe, hvoraf forholdsvis mange har færdiggjort deres uddannelse udenfor landet.

Betragtes uddannelsesfordelingen, hvor skønnet uddannelsesniveau anvendes, ses der ikke at være den store forskel på uddannelsesniveauet blandt personer, der flytter til og fra et varigt udsat alment boligområde. I de varigt udsatte almene boligområder ses dog fortsat at være relativt få erhvervsuddannede og særligt personer med videregående uddannelser i forhold til fordelingen i den resterende befolkning.

Tabel 3.5. Tilflyttere, fraflyttere og fastboende i varigt udsatte almene områder efter højst fuldførte uddannelse, andel i procent, gennemsnit 2012-2016.

	Imputeret niveau som grundskole				Imputeret niveau anvendes			
	Fastbo	Tilflyt	Fraflyt	Øvrig	Fastbo	Tilflyt	Fraflyt	Øvrig
Uoplyst	-	-	-	-	6	10	9	2
Grundskole	69	67	62	24	42	34	35	18
Gymnasial	3	3	3	5	8	7	7	6
Erhvervsfaglig	18	17	20	36	27	24	26	37
Kort videreg.	2	2	3	6	4	4	4	6
Mellemlang videreg.	6	7	7	18	9	11	10	19
Lang videreg.	2	4	5	11	5	10	8	12

Anm.: Uddannelsesniveaet er opgjort for personer i alderen 30-59 år.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

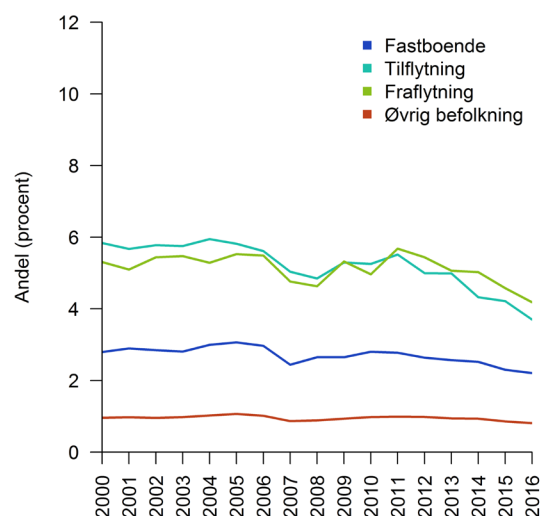
Kriminalitet

Niveauet af kriminalitet i et boligområde måles ud fra, hvor stor en andel af områdets beboere, som indenfor det seneste år er dømt for overtrædelse af straffeloven, våbenloven eller lov om euforiserende stoffer.

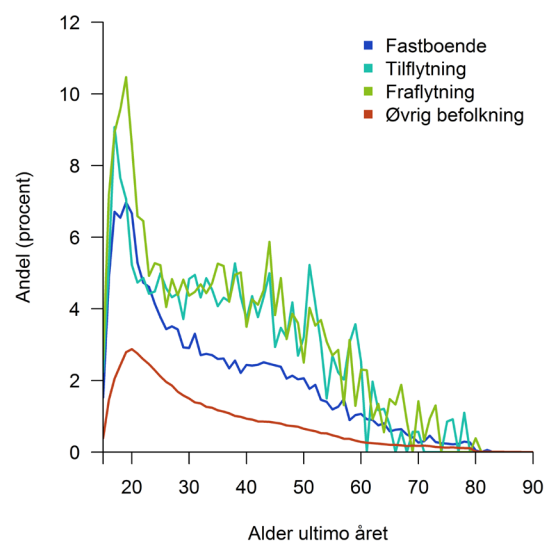
Der ses en klar tendens til, at en større andel af til- og fraflyttere i varigt udsatte almene områder er dømt i forhold til områdets fastboende beboere, jf. Figur 3.13a. Dette skyldes et generelt højere omfang af kriminalitet blandt til- og fraflyttere end blandt fastboende i alle aldre, jf. Figur 3.13b. Over tid ses en tilbøjelighed til, at kriminalitet blandt tilflyttere til områderne aftager i større grad end kriminalitet blandt fraflyttere⁹.

Figur 3.13. Andel af dømte for tilflyttere, fraflyttere og fastboende.

a) Varigt udsatte almene områder



b) Gns. andel efter alder, gns. 2012-2016



Anm.: Dømte omfatter personer dømt for overtrædelse af straffeloven, våbenloven eller lov om euforiserende stoffer indenfor seneste kalenderår. I figuren til venstre omfatter andelen personer på mindst 15 år.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

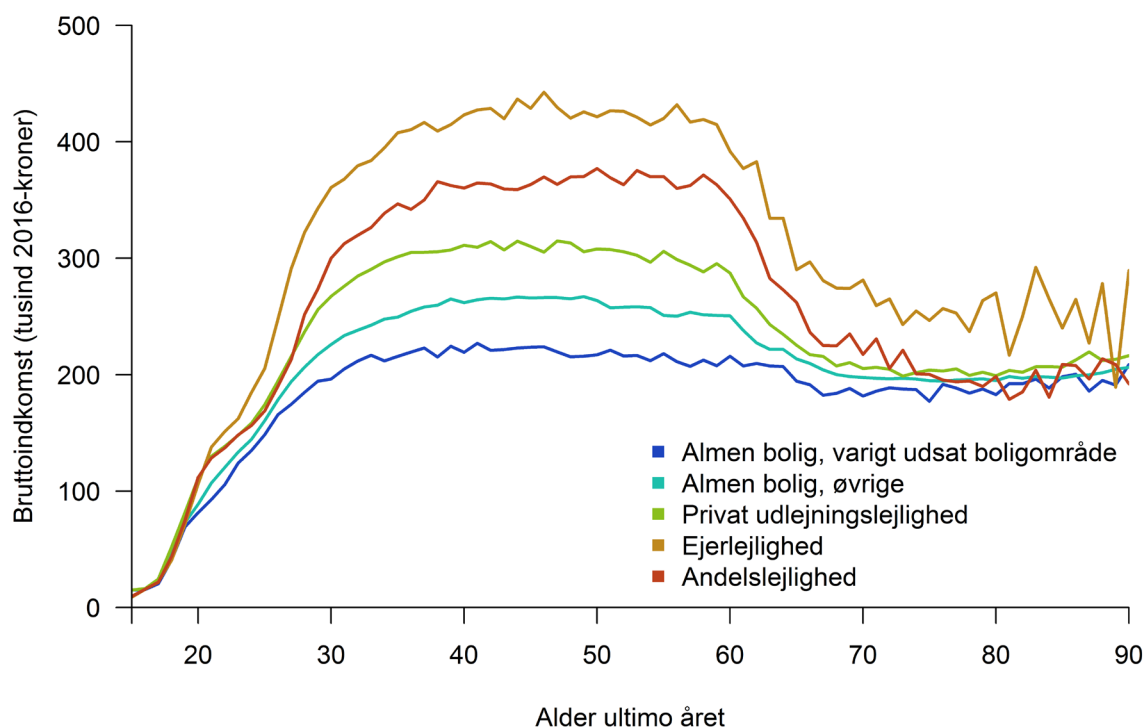
3.4 Indkomstudvikling omkring flyttetidspunkt

Tilflyttere til varigt udsatte almene boligområder har relativt lavt indkomstniveau sammenlignet med tilflyttere til øvrige almene boliger og i særdeleshed i forhold til tilflyttere i andre boligformer, jf. Figur 3.14.

I dette afsnit undersøges det, hvorvidt det relativt lave indkomstniveau blandt tilflyttere til varigt udsatte almene boligområder kan forklares af uforudsete hændelser i perioden frem mod tilflytning. Desuden ses der på indkomstudviklingen i årene efter tilflytning til et varigt udsat alment boligområde. På tilsvarende vis betragtes hændelser i årene frem mod fraflytning fra et varigt udsat alment boligområde, og fraflytternes indkomstudvikling efter fraflytning fra området belyses.

⁹ Det medfører som udgangspunkt ingen ændring af en persons bopælsregistrering, at vedkommende kommer i fængsel. Personen kan selv vælge at fraflytte boligen i forbindelse med eventuel afsoning.

Figur 3.14. Medianindkomst for tilflyttere til forskellige boligtyper, gennemsnit 2012-2016.



Anm.: Indkomst er opgjort som den gennemsnitlige bruttoindkomst i 2016-kroner. For personer, som kun har opholdt sig i landet en del af året – dvs. er indvandret, udvandret eller døde – er personens indkomst opregnet til en årsindkomst.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Uforudsete hændelser i årene op til flytning til et varigt udsat alment boligområde

I det følgende betragtes den andel af flytninger, hvor personen, der flytter, er bosat i Danmark i årene frem mod flyttetidspunktet. Disse udgør knap 96 procent af det samlede antal flytninger. Idet en del tilflytninger til varigt udsatte almene boligområder sker ved indvandring fra udlandet, er andelen i disse områder lidt lavere, hvorfor lige over 91 procent af alle tilflyttere til områderne indgår.

Det undersøges, hvorvidt tilflytterne har oplevet uforudsete hændelser op mod tilflytningen, som kan have påvirket flyttebeslutningen. Som uforudsete hændelser indgår:

- Parsplittelse, dvs. at tilflytteren er gået fra partneren i et regulært parforhold¹⁰.
- Overgang til midlertidig offentlig forsørgelse, der omfatter dagpenge under ledighedsperioder, sygedagpenge, kontanthjælp inklusiv integrationsydelse mv. Det er en forudsætning, at perioden med offentlig forsørgelse er af en vis varighed¹¹.
- Permanent offentlig forsørgelse, som dækker over tilgang til førtidspension.

¹⁰ Med parsplittelse menes der, at personen er fraflyttet sin partner fra et ægtepar (inklusive ægtepar af samme køn) eller et samlevende par, dvs. to personer der ikke er i ægteskab med hinanden, men som har mindst ét fælles barn.

¹¹ En person defineres til at være overgået til midlertidig offentlig forsørgelse, hvis vedkommende overgår til offentlige ydelser, der udgør mindst 20 procent af personens årsindkomst eller hvis personens offentlige overførsler stiger således, at disse udgør hovedparten af personens forsørgelsesgrundlag. Uddannelsesstøtte medregnes ikke ved opgørelsen af midlertidige offentlige ydelser.

Tilflytternes indkomstgrundlag betragtes over en treårig periode, som omfatter tilflytningsåret og de to foregående år.

Over denne treårige periode er det omkring fem procent af den samlede befolkning i den arbejdsdygtige alder (eksklusiv hjemmeboende børn), der oplever parsplittelse, jf. Tabel 3.6. Ikke overraskende er denne persongruppe overrepræsenteret blandt personer, der flytter, idet mindst én og ofte begge parter skifter bolig i forbindelse med parsplit. Samlet set er det 10 procent af de personer, der flytter i løbet af et år, som har oplevet parsplittelse i tilflytningsåret eller indenfor de to foregående år. Blandt tilflyttere til almene boliger ses en svag overrepræsentation af personer, som har oplevet parsplittelse. Dette gælder både til almene boliger i varigt udsatte boligområder og øvrige almene boliger.

Personer, som overgår til midlertidig offentlig forsørgelse, er klart overrepræsenteret blandt tilflyttere til almene boliger. Dette gælder cirka hver fjerde tilflytter mod blot hver femte af flyttende personer generelt. Andelen er lidt højere blandt tilflyttere til varigt udsatte almene boliger end blandt dem, som flytter til øvrige almene boliger. Også personer, der for nyligt er overgået til førtidspension, ses oftere blandt tilflyttere til almene boliger end blandt personer, der flytter til andre boligformer.

Tabel 3.6. Uforudsete hændelser i årene op til flytning, andel i procent, 2012-2016.

	Befolkning	Flyttende personer efter tilflytningsbolig			
		Alle	Almen bolig, varigt udsat	Almen bolig, øvrige	Lejligheder, øvrige
Hændelse indtruffet indenfor 2 år op til flytning					
- Parsplittelse	5	10	11	12	8
- Overgang til dagpenge/kontanthjælp	14	20	28	24	18
- Overgang til førtidspension	1,2	1,2	1,8	2,1	0,6
Indkomst på flyttetidspunkt ift. maksimal indkomst tre foregående år					
- Indkomstnedgang 10-20 procent	12	11	11	11	10
- Indkomstnedgang 21-40 procent	11	12	13	13	12
- Indkomstnedgang 41-60 procent	6	9	10	8	9
- Indkomstnedgang 60+ procent	3	4	4	3	4
- Maks indkomst 3 forudgående år (tusind)	384	278	195	224	251

Anm.: Tabellen indeholder personer på 18-59 år, som ikke er hjemmeboende ved forældrene. Gruppen "Lejligheder, øvrige" indeholder personer, som flytter til en ejer-, andels- eller privat udlejningslejlighed.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

I Tabel 3.6 ses desuden, hvor stor en andel af de personer, der flytter, som har oplevet indkomstnedgang frem mod flyttetidspunktet. Her beregnes indkomstændringen som forholdet mellem personens indkomst i flytteåret sat i forhold til personens maksimale indkomst i de tre foregående år.

En forholdsvis stor andel af befolkningen oplever udsving i indkomsten. Dette hænger sammen med, at udviklingen i forhold til den maksimale indkomst de foregående år betragtes. Lavere arbejdstid i forhold til tidligere, udbetaling af bonus i foregående år, jobskifte mv. kan være forklaringer på en mindre indkomstnedgang. Indkomstnedgang på op til 10 procent vises således ikke i tabellen.

Af tabellen ovenfor fremgår det, at tilflytterne til almene boliger oplever indkomstnedgang af forskellig grad i nogenlunde samme omfang som øvrige personer, der flytter. Det er bemærkelsesværdigt, at tilflytterne til almene boliger – og særligt varigt udsatte almene boliger – har noget lavere indkomst set over den treårige periode frem mod tilflytningen, selv hvis den maksimale indkomst i perioden registreres. Kombineret med, at tilflytterne til varigt udsatte almene boliger i større grad end andre overgår til offentlig forsørgelse, indikerer et permanent lavt indkomstniveau, at tilflytterne til varigt udsatte almene boligområder i nogen grad er personer på kanten af arbejdsmarkedet.

I tabellen ovenfor betragtes alle personer i den arbejdsduelige alder under ét. Betragtes mindre aldersgrupper indeholdende fem alderstrin, er billedet det samme: Der ses ikke nogen nævneværdig forskel på andelen af tilflyttere til varigt udsatte almene boligområder, som har oplevet indkomstnedgang i årene frem mod tilflytning i forhold til tilflyttere til øvrige almene boliger eller generelt. Dog ses der, som i tabellen ovenfor, en tendens til overrepræsentation på 1-3 procentpoint af andelen af personer, der har oplevet kraftig indkomstnedgang på mindst 40 procent blandt tilflytterne til almene boliger.

Indkomstudvikling efter flytning til et varigt udsat alment boligområde

Personer, som tilflytter et varigt udsat alment område, har i gennemsnit lavere indkomst end de beboere, som allerede bor i området, jf. Figur 3.9. Særligt for de yngre tilflyttere observeres en lavere gennemsnitlig indkomst end for fastboende, hvilket hovedsageligt tilskrives, at der blandt tilflytterne er en større andel af nyindvandrede personer, som ved indvandring har et relativt lavt indkomstniveau.

Der ses en tydelig tendens til, at tilflytterne i årene efter tilflytning til et varigt udsat alment boligområde nærmer sig indkomstniveauet for områdets øvrige beboere. Det fremgår af Figur 3.15, hvor den gennemsnitlige indkomst for tilflyttere sættes i forhold til indkomsten for fastboende beboere i området i samme aldersklasse.

Tilflyttere i 20'erne har på flyttetidspunktet en indkomst, der i gennemsnit er 80 procent af indkomsten for fastboende. Det lavere indkomstniveau skyldes blandt andet, at der blandt tilflytterne er flere nyindvandrede og studerende – der i gennemsnit har relativt lav indkomst i forhold til øvrige grupper – end det er tilfældet blandt de fastboende, hvor der til gengæld ses en større andel af beskæftigede. Efter tilflytning ses tilflytterne at have en mere gunstig indkomstudvikling end de personer, som allerede boede i området. Efter to år udgør tilflytternes indkomst således omkring 85 procent af indkomsten blandt de oprindeligt fastboende personer. Andelen fortsætter med at stige, jo længere tid der går siden tilflytning. Efter fem år udgør indkomsten for de oprindelige tilflyttere 90 procent af indkomsten for fastboende.

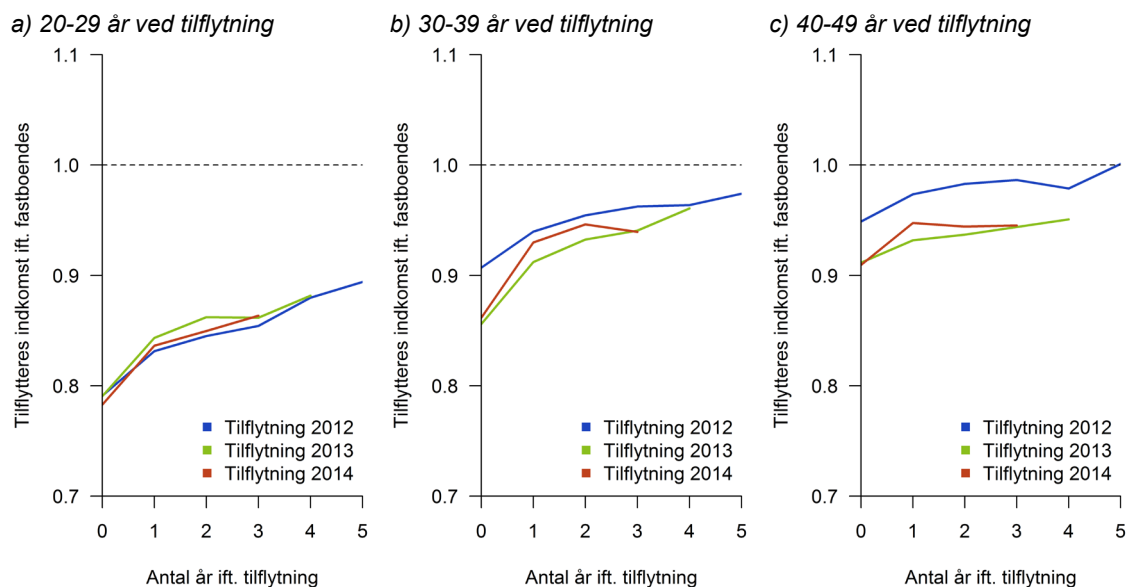
Samme billede tegner sig, hvis tilflyttere i 30'erne eller 40'erne betragtes. Ved tilflytning til et varigt udsat område ses den gennemsnitlige indkomst at være lavere end blandt eksisterende beboere, hvorefter tilflytternes indkomstniveau konvergerer mod indkomsten for personer, der allerede boede i området. Jo ældre tilflyttere, som betragtes, jo tættere er deres indkomst på tilflytningstidspunktet ved den gennemsnitlige indkomst for områdets eksisterende beboere, hvilket ligeledes observeredes i Figur 3.9.

Der kan være flere forklaringer på, at tilflyttere har en mere positiv indkomstudvikling end områdets eksisterende beboere. Dels kan tilflytning til det varigt udsatte boligområde ske i en periode, hvor personen midlertidigt oplever et lavt indkomstniveau, f.eks. som følge af ledighed eller sygdom. Dels udgør indvandringer fra udlandet en relativt stor andel af de personer, som tilflytter områderne. For nogle af disse personer kan der gå noget tid, før de

etablerer sig på arbejdsmarkedet, og deres indkomstniveau derfor er på niveau med områdets øvrige beboere.

Tendensen observeret i figuren nedenfor – at tilflytteres indkomstniveau over tid konvergerer mod indkomstniveauet for de eksisterende beboere – gælder alle herkomstgrupper.

Figur 3.15. Indkomstudvikling i årene efter tilflytning til et varigt udsat alment boligområde.



Anm.: Figuren viser bruttoårsindkomsten for tilflyttere til varigt udsat almene boligområder sat i forhold til indkomsten for fastboende i samme aldersinterval i samme områder. År nul er året, hvor personen tilflytter boligen, år 1 er efterfølgende år, år 2 er to år efter tilflytning osv.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Hændelser, som kan påvirke tilbøjeligheden til flytning fra varigt udsatte almene boligområder

En husholdnings beslutning om at flytte afhænger af en lang række faktorer. Hændelser såsom påbegyndelse eller færdiggørelse af uddannelse, graviditet eller familieforøgelse, ændringer i jobsituation mv. kan påvirke flyttebeslutningen. Også en række fysiske forhold som boligens beliggenhed, størrelse, nærområde mv. kan have indflydelse på, hvorvidt en husholdning vælger at skifte bolig.

I det følgende fokuseres på, hvorvidt der sker en øget fraflytning fra varigt udsatte almene boliger, når beboerne overgår til beskæftigelse¹², færdiggør uddannelse eller oplever indkomstfremgang.

Blandt befolkningen på 18-59 år (som ikke er hjemmeboende ved deres forældre) er det cirka 5 procent, der færdiggør en uddannelse over en toårig periode, jf. Tabel 3.7. Færdiggørelse af uddannelse er langt mere udbredt blandt personer, som flytter. Her er det i gennemsnit 14 procent, som har færdiggjort en uddannelse indenfor de seneste to år. Dette er naturligt, da flytninger typisk sker i 20'erne, hvor der ligeledes observeres stor uddannelsesaktivitet, ligesom en del flytninger er påtvunget ved afslutning af uddannelse, f.eks. fraflytning af deciderede studieboliger.

¹² Overgang til beskæftigelse defineres ved, at en person går fra offentlig forsørgelse til at have erhvervsindkomst, som udgør mere end en tredjedel af personens årsindkomst, eller at personen går til, at erhvervsindkomst udgør hovedparten af personens samlede fortjeneste.

Som det fremgår af tabellen nedenfor ses der ikke tendens til øget fraflytning fra almene boliger ved færdiggørelse af uddannelse eller ved overgang til beskæftigelse. Andelen af fraflytterne fra almene boliger, som har oplevet disse begivenheder, er på niveau med andelen blandt fraflyttere generelt.

Af tabellen fremgår det desuden, hvor stor en andel af de personer, der flytter, som har oplevet indkomstfremgang frem mod flyttetidspunktet. Her beregnes indkomstændringen som forholdet mellem personens indkomst i flytteåret (eller efterfølgende år, hvis indkomsten her er større end i flytteåret) sat i forhold til personens mindste indkomst i de to år, som går forud for flytteåret. Der ses ikke tegn på, at fraflytterne fra almene boligområder oplever højere indkomstfremgang frem mod fraflytningstidspunktet end fraflytterne i øvrige boligformer.

Resultaterne vist i tabellen nedenfor ændres ikke i betydelig grad, hvis der betragtes aldersgrupper bestående af færre alderstrin end dem i den arbejdsdygtige alder. Der ses en tendens til, at en lidt større andel af fraflytterne fra varigt udsatte almene boligområder oplever kraftig indkomstfremgang, men dette gælder blot 1-2 procentpoint flere end fraflyttere fra øvrige boligtyper.

Tabel 3.7. Hændelser i årene op til flytning, andel i procent, 2012-2016.

	Befolkning	Flyttende personer efter fraflytningsbolig			
		Alle	Almen bolig, varigt udsat	Almen bolig, øvrige	Lejligheder, øvrige
Hændelse indtruffet indenfor flytteåret eller foregående år					
- Færdiggørelse af uddannelse	5	14	12	14	17
- Overgang til beskæftigelse	5	6	7	7	6
Indkomst på flyttetidspunkt ift. mindste indkomst to foregående år					
- Indkomstfremgang 10-20 procent	12	12	10	11	11
- Indkomstfremgang 21-40 procent	12	14	12	13	15
- Indkomstfremgang 41-60 procent	6	8	8	8	9
- Indkomstfremgang 60 procent eller mere	19	33	35	33	37

Anm.: Tabellen indeholder personer på 18-59 år som ikke er hjemmeboende ved forældrene. Gruppen "Lejligheder, øvrige" indeholder personer, som fraflytter en ejer-, andels- eller privat udlejningslejlighed.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Indkomstudvikling efter flytning fra et varigt udsat alment boligområde

De foregående undersøgelser viste, at fraflytterne fra varigt udsatte almene boligområder i gennemsnit har samme indkomstniveau som de beboere af samme alder, som forbliver boende i områderne, jf. Figur 3.9.

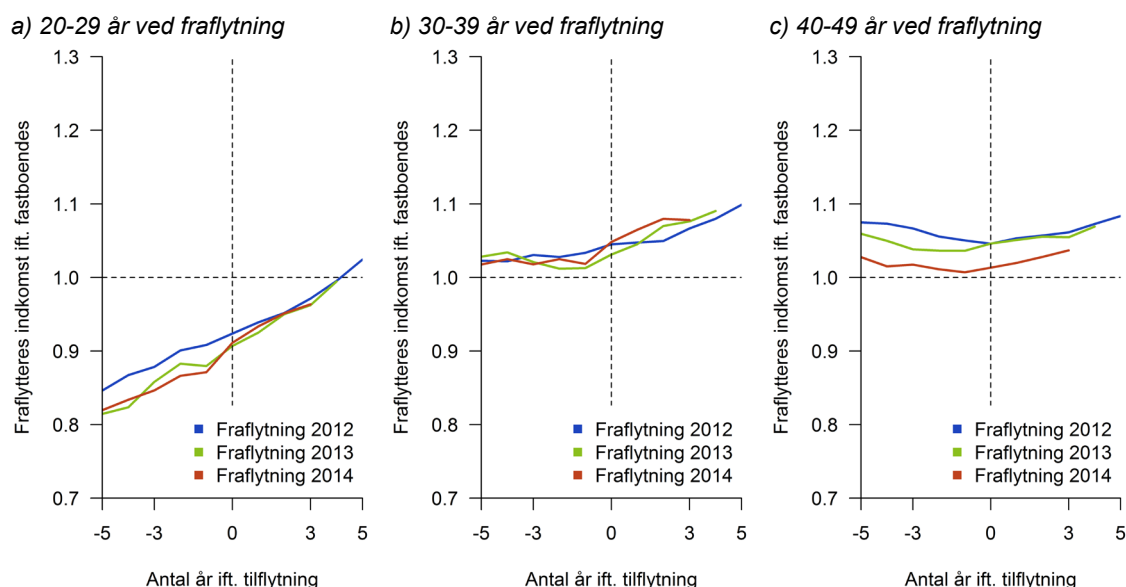
Her betragtes udviklingen i indkomst efter fraflytning fra det varigt udsatte almene boligområde. I Figur 3.16 betragtes personer, som fraflytter et varigt udsat alment boligområde i perioden 2012-2014. Der ses på den gennemsnitlige indkomstudvikling for disse personer i årene efter fraflytning, hvilket sættes i forhold til indkomstudviklingen for de personer, som i perioden forbliver boende i området.

Året for fraflytning er i figuren markeret med den lodrette stiplede linje. For alle tre aldersgrupper ses en tendens til, at indkomstudviklingen i den femårige periode frem mod fraflytningstidspunktet er tilnærmelsesvis ens for de personer, som fraflytter, og for de

personer, der vælger at blive boende i området. For alle tre aldersgrupper gælder desuden, at der ses en tendens til en mere positiv indkomstudvikling i årene efter fraflytning for de personer, der fraflytter det varigt udsatte almene boligområde, end for de personer, der bliver boende i området. Tendensen er dog mest tydelig for fraflyttere i 20'erne og 30'erne.

Dette kan indikere, at fraflytterne i gennemsnit er en mere ressourcestærk gruppe end de personer, som bliver boende i det varigt udsatte almene boligområde, selvom personer i samme alder havde samme gennemsnitlige indkomst på flyttetidspunktet.

Figur 3.16. Indkomstudvikling i årene efter fraflytning fra et varigt udsat alment boligområde.



Anm.: Figuren viser bruttoårsindkomsten for tilflyttere til varigt udsat almene boligområder sat i forhold til indkomsten for fastboende i samme områder. År nul er året, hvor personen tilflytter boligen, år 1 er efterfølgende år, år 2 er to år efter tilflytning osv.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

3.5 Flyttemønstre og segregeringsgrad

Tilflytterne til varigt udsatte almene boligområder flytter per definition til et område med høj grad af socioøkonomisk segregering. Inddeles befolkningen i 20 lige store grupper efter det Socioøkonomiske Nabolag-Indeks, vil beboerne i de varigt udsatte almene boligområder per konstruktion være blandt den gruppe med mest socioøkonomisk udsathed (se konstruktion af Socioøkonomiske Nabolag-Indeks ud fra CNN-algoritmen i afsnit 4).

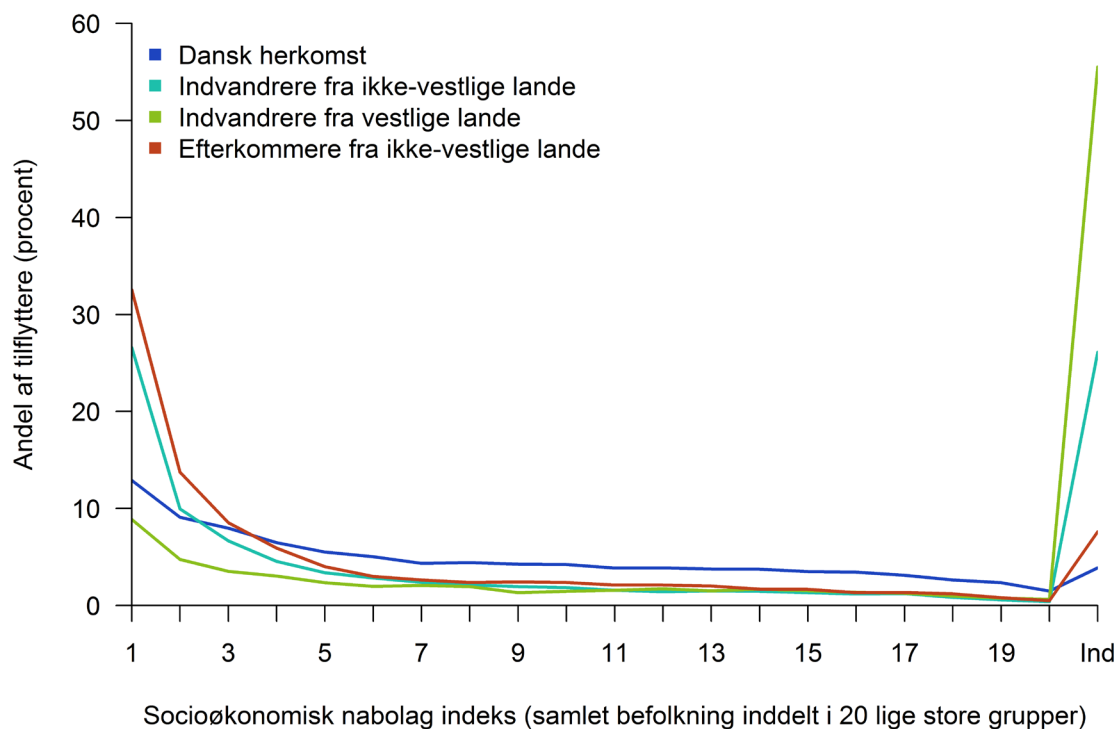
For de beboere, som flytter til et varigt udsat alment område, betragtes graden af socioøkonomisk udsathed i det nabolag, vedkommende flyttede fra. Som mål for fraflytningsområdets udsathed anvendes det Socioøkonomiske Nabolag-Indeks, hvor befolkningen inddeles i 20 lige store grupper. Den første (sidste) gruppe indeholder 5 procent af befolkningen bosat i de mest (mindst) udsatte nabolag. For personer indvandret til Danmark i tilflytningsåret er det ikke muligt at opgøre det Socioøkonomiske Nabolag-Indeks ved fraflytningsnabolaget. Denne gruppe opgøres derfor særskilt.

Tilflyttere til varigt udsatte almene områder fraflytter generelt områder, der er blandt de 25 procent mest udsatte nabolag (indeks 1-5 i figuren nedenfor). Dette gælder 45 procent af alle tilflyttere, jf. Figur 3.17, hvoraf 19 procentpoint flytter fra nabolag, som er udsat i samme grad som tilflytningsnabolaget. Omkring 37 procent flytter fra nabolag, som ikke er blandt de 20

procent mest udsatte (indeks 6-20 i figuren). De resterende 17 procent af tilflytterne indvandrer fra udlandet (markeret yderst til højre i figuren).

At relativt dårligt stillede husstande flytter fra nabolag, som ikke er udsatte, til udsatte nabolag, kan være medvirkende til at øge graden af segregering i de varigt udsatte almene områder. Som det fremgår af Figur 3.18 nedenfor, sker der dog en modsatrettet fraflytning fra udsatte områder, som delvis udligner dette.

Figur 3.17. Tilflyttere til varigt udsatte almene områder efter socioøkonomisk nabolag indeks i fraflytningsområdet, gennemsnit 2012-2016.



Anm.: Socioøkonomiske Nabolag-Indeks inddeles i 20 lige store grupper, som hver indeholder fem procent af befolkningen primo året. Herudover kommer tilflyttere, som er indvandret direkte fra udlandet (indtegnet i figuren som gruppen yderst til højre).

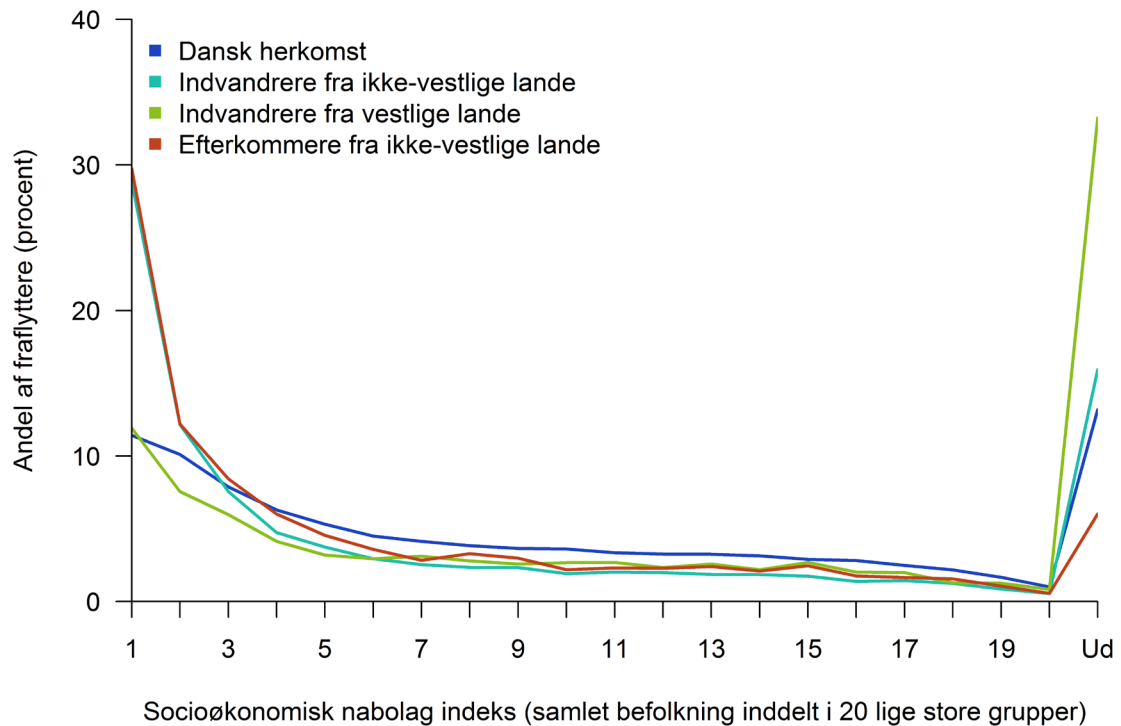
Efterkommere fra vestlige lande udgør en relativt lille andel af det samlede antal tilflyttere, hvorfor denne herkomstgruppe ikke indgår i figuren. Figuren indeholder tilflyttere, som er mindst 18 år gamle.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Personer, som flytter fra varigt udsatte almene områder, fraflytter boliger, som er beliggende i områder, der er blandt de fem procent mest socioøkonomisk udsatte i landet. I Figur 3.18 betragtes graden af socioøkonomisk udsathed i det område, beboere fra varigt udsatte almene områder flytter til. I figuren er det Socioøkonomiske Nabolag-Indeks inddelt i 20 lige store grupper (ventiler), hvor den første (sidste) gruppe indeholder de 5 procent af befolkningen, som bor i de mest (mindst) udsatte områder.

Knap hver femte (18,5 procent), som fraflytter et varigt udsat alment område, flytter til en bolig beliggende i et nabolag med samme grad af socioøkonomisk udsathed, jf. Figur 3.18 (indeks 1 i figuren). Blandt personer af ikke-vestlig herkomst er andelen dog betydeligt større. Knap halvdelen (47 procent) af fraflytterne flytter til nabolag, der er blandt de 25 procent mest udsatte (indeks 1-5 i figuren). Omkring 15 procent fraflytter som følge af udvandring eller dødsfald, andelen er særligt høj for indvandrere fra vestlige lande.

Figur 3.18. Fraflyttere fra varigt udsatte almene områder efter socioøkonomisk nabolag indeks i tilflytningsområdet, gennemsnit 2012-2016.



Anm.: Socioøkonomisk nabolag indeks indeles i 20 lige store grupper, som hver indeholder fem procent af befolkningen ultimo året. Herudover kommer fraflyttere, som er udvandret eller døde (indtegnet i figuren som gruppen yderst til højre).

Efterkommere fra vestlige lande udgør en relativt lille andel af det samlede antal fraflyttere, hvorfor denne herkomstgruppe ikke indgår i figuren. Figuren indeholder fraflyttere, som er mindst 18 år gamle.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

4. Individuelle indkomstforløb

At et individ oplever en ændring i sit Socioøkonomiske Nabolag-Indeks (SNI) over tid er per definition ikke alene funderet i personens egne socioøkonomiske karakteristika, men er tæt knyttet til karakteristika for de nærmeste naboer. Hvis man ønsker at se på, hvorledes et individ klarer sig socioøkonomisk over tid, er SNI altså ikke et ideelt mål at anvende. Gennem anvendelse af forskellige statistiske tilgange søger dette afsnit at udrede, hvorvidt der er et signifikant forhold mellem personkarakteristika og diversiteten i befolkningens socioøkonomiske formåen gennem en udvalgt periode af et livsforløb. Specielt ønskes det at udrede om der er en signifikant relation mellem bopæls erfaring fra et udsat område og forskelle i socioøkonomisk status. Eftersom personlig indkomst spiller en væsentlig rolle i afgrænsningen af SNI, jf. Figur 6.3, er det derfor her valgt at udtrykke udviklingen i socioøkonomisk status ved udviklingen i indkomst.

Udviklingen i indkomst over livet udviser betydelig variation på tværs af individer og disses karakteristika. For den analyserede målgruppe peger resultaterne i retning af, at køn, oprindelse, familietype, uddannelsesniveau, arbejdsmarkedstilknytning, indkomsthistorik og boligtype kan signifikant relateres til indbyrdes forskelle i indkomstforløb over livet. Omvendt kan der ikke påvises et signifikant forhold mellem nuværende eller tidligere bosættelse i et udsat område og indbyrdes forskelle i indkomstforløb. Det vises også, at det ved en algoritmisk tilgang er muligt at klassificere indkomstforløbene i fire grupper, der meningsfuldt kan identificeres som omfattende forløb kendetegnet ved ophold i de henholdsvis laveste, næstlaveste, næsthøjeste og højeste indkomstdeciler. Førnævnte baggrundskarakteristika påvises også at have signifikant betydning for dette gruppetilhørsforhold, mens det ikke entydigt kan påvises, at en person klassificeres i én gruppe frem for en anden alene, fordi vedkommende har erfaring med bosættelse i et udsat område. Konklusionen beror på, hvorvidt samvariationen mellem område- og arbejdsmarkedserfaring indregnes. Der er endvidere tale om en gennemsnitsbetragtning, hvorfor man godt kan opnå andre resultater, hvis områdeeffekten analyseres for undergrupper af befolkningen eventuelt afgrænset ved forskelle i oprindelse, uddannelse eller arbejdsmarkedstilknytning. Eksistensen af sådanne betingede effekter vedrørende forholdet mellem områdeerfaring og tilhørsforholdet til indkomstgrupperne afsøges ikke i dette afsnit.

Det er centralt at bemærke, at resultatet af analyserne udelukkende udtaler sig om, hvorvidt der er en signifikant sammenhæng mellem baggrundskarakteristika og indkomstforløb og ikke om kausaliteten heraf. Dette afsnit har dermed ikke til formål at udlede kausal inferens, men i de tilfælde, hvor der er oplagte endogenitetsproblemer i form af simultaneitet, vil analyserne søges udført på en sådan måde, at denne type af bias mindskes. I praksis er det dog, dette til trods, ikke muligt at opretholde en antagelse om eksogenitet og resultaterne skal derfor kun benyttes med dette væsentlige forbehold¹³. Spørgsmålet om kausalitet adresseres i afsnit 5.

Indledningsvist præsenteres i afsnit 4.1 udviklingen i målgruppens indkomststruktur fordelt på udvalgte personkarakteristika, hvilket giver en indikation på hvilke forhold, der giver anledning til diversitet i indkomstfordelingen. Der ses i denne forbindelse bl.a. på, hvorledes indkomststrukturen varierer på tværs af køn, alder, oprindelse, uddannelsesniveau, arbejdsmarkedstilknytning, område- og boligkarakteristika. Der er primært tale om et deskriptivt afsnit, som det ikke er en forudsætning at læse for at få et udbytte af de efterfølgende afsnit.

¹³ Forholdet mellem de uafhængige variable og den afhængige variabel omtales i flæng som *sammenhæng*, *forhold*, *relation* eller *forklaring* uden at dette må tillægges en kausal betydning.

Efterfølgende gennemføres i afsnit 4.2 en algoritmisk gruppering af statistisk ensartede indkomstforløb og der redegøres for, hvorledes individernes baggrundskaraktistika varierer på tværs af disse grupper både ud fra en deskriptiv og en statistisk betragtning. Således vil det blive klarlagt hvilke personkaraktistika, der har signifikant betydning for sandsynligheden for at tilhøre én gruppe frem for en anden.

Klassifikationsanalysen har til formål at skabe et overblik over typiske indkomstforløb og udrede, hvorvidt udvalgte baggrundskaraktistika signifikant kan forklare tilhørsforholdet til grupperne. Tilgangen bortser dog fra diversiteten i indkomstforløb inden for hver gruppe og suppleres derfor med en analyse, der imødeser dette. Ved anvendelse af en slags *variansanalyse* redegøres i afsnit 4.3 for, hvorvidt udvalgte baggrundskaraktistika signifikant kan knyttes til forskelle mellem indkomstforløb uafhængigt af tilhørsforholdet til klassifikationsgrupperne. Resultatet klarlægger således i modsætning til ovenstående om de enkelte karaktistika kan relateres til indbyrdes forskelle mellem samtlige livsforløb og ikke blot om de har betydning for en overordnet kategorisering af indkomstforløb.

Analyserne fokuserer på individer, der er 30 år i begyndelsen af 1986 og som uden afbrydelser opholder sig i landet frem til begyndelsen af år 2017, hvor de pågældende er 61 år. Nedenfor betegnes denne fokusgruppe som målgruppen, der omfatter i alt 47.479 personer. I praksis vil der dog, af årsager, som senere vil blive klarlagt, blive fokuseret på udviklingen i indkomststrukturen i aldersgrupperne 35-61 år, dvs. i perioden 1991-2017.

Valget af perioden og kravet om opholdstid er truffet af forsimplende hensyn, idet analysetilgangen ikke er begrænset til, at opholdstidskravet skal være opfyldt for samtlige forløb. Således er det også muligt, at man betragter personer, der både har opholdt sig inden og uden for landets grænser i perioden og dermed føjer opholdstid til de karaktistika, der kan forklare forskelle mellem klassifikationsgrupperne og de indbyrdes forløb. Endvidere indbyder tilgangen til, at man simultant betragter udviklingen i indkomstforløb for forskellige generationer, men i de samme aldersgrupper, hvilket muliggør identifikation af en generationseffekt. Sidstnævnte er afsøgt gennem en analyse af indkomstforløbene i aldersklasserne 35-50 år for generation 1956-1967 og det er her fundet, at der *ikke* kan påvises en signifikant sammenhæng mellem fødselsår og indbyrdes forskelle i indkomstforløb, heller ikke ved kontrol for øvrige karaktistika. På den baggrund er det derfor valgt alene at fokusere på en enkel generation og dermed muliggøre, at den pågældende generation kan følges gennem en længere periode.

På ethvert tidspunkt i løbet af den betragtede periode karakteriseres målgruppen ved en række beskrivende karaktistika vedrørende person- og familiespecifikke forhold samt karaktistika for bopælsområdet og boligen¹⁴. En persons bopælsområde er på ethvert tidspunkt karakteriseret som værende udsat eller ikke-udsat iht. definitionen introduceret i de foregående afsnit. Såfremt der for en given person observeres et skift i bopælsområdets status (udsat/ikke-udsat) over tid, sondres der her ikke mellem, hvorvidt dette skyldes et skift i områdets karaktistika eller om det skyldes, at personen er flyttet fysisk til et andet område. En person, der har boet de første 15 år i et udsat område og efterfølgende er flyttet til et ikke-udsat område, vil fra analysens synspunkt have de samme områdekarakteristika, som en person, der har boet det samme sted i hele sit liv, men hvor områdets karaktistika efter 15 år har ændret sig fra udsat til ikke-udsat grundet ændring i beboersammensætningen eller i disses karaktistika.

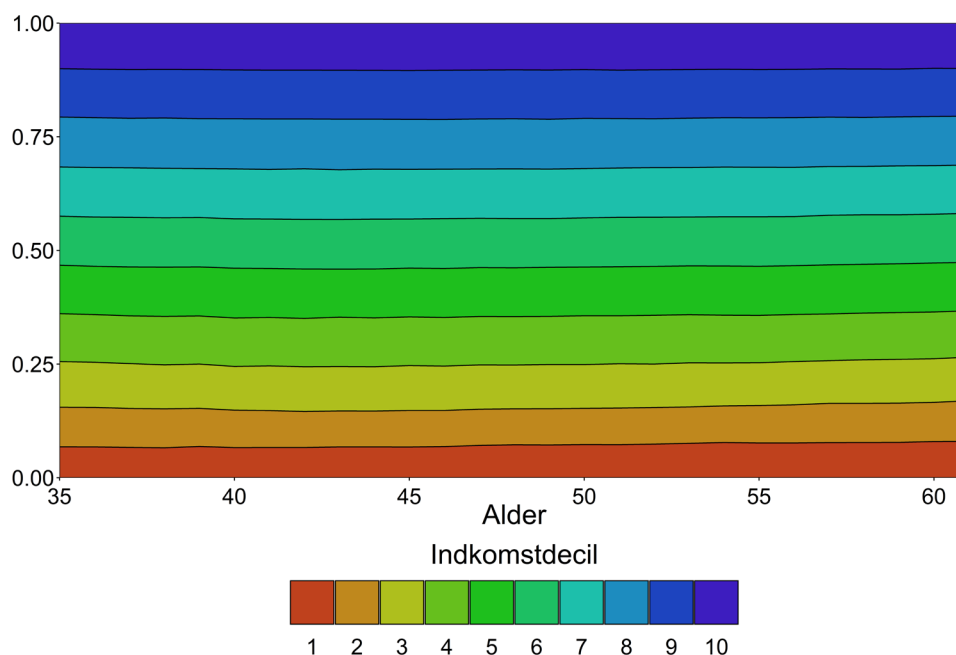
¹⁴ Der inddrages udelukkende personer i analysen, hvorom der udover konstant ophold i landet også kræves komplette informationer for alle beskrivende karaktistika gennem hele perioden. Personer med ukendt boligtype er således udeladt.

4.1 Indkomstforløb og baggrundskarakteristika

Som optakt til de statistiske analyser betragtes i dette afsnit forholdet mellem indkomstfordelinger og udvalgte baggrundskarakteristika. Dette kan give en indikation på, hvorvidt indkomstforløbene varierer på tværs af bestemte karakteristika, der i så fald kan spille en væsentlig rolle i at forklare, hvorfor en person klassificeres i én bestemt gruppe relativt til en anden, eller som kan forklare diversiteten i indkomstforløbene uafhængigt af gruppetilhørsforholdet. Det er dog centralt at bemærke, at indkomststrukturens variation nedenfor illustreres for et karakteristikum ad gangen. Findes der for et givet karakteristikum en forskel i indkomststrukturen, kan man altså ikke slutte, at dette entydigt stammer fra variation i det pågældende karakteristikum. Der kan meget vel være tale om en korrelation med andre forklarende variable, der blot medieres gennem det analyserede karakteristikum.

Målgruppens indkomstforløb i perioden 1991-2017 er karakteriseret ved individernes placering i aldersbetingede indkomstdeciler. Indkomstdecilerne er baseret på hele befolkningens indkomstfordeling i de enkelte aldersgrupper¹⁵. Grundet opholdstidskravet udgør målgruppen ikke en fuldkommen repræsentativ stikprøve af befolkningen, og målgruppens overordnede fordeling på indkomstdeciler vil derfor ikke være konstant over tid eller ligeligt fordelt på deciler, jf. Figur 4.1. At man således udelukker personer med midlertidigt ophold fra analysen betyder, at der er tendens til en relativt større repræsentation af personer i de fem øverste indkomstdeciler.

Figur 4.1. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

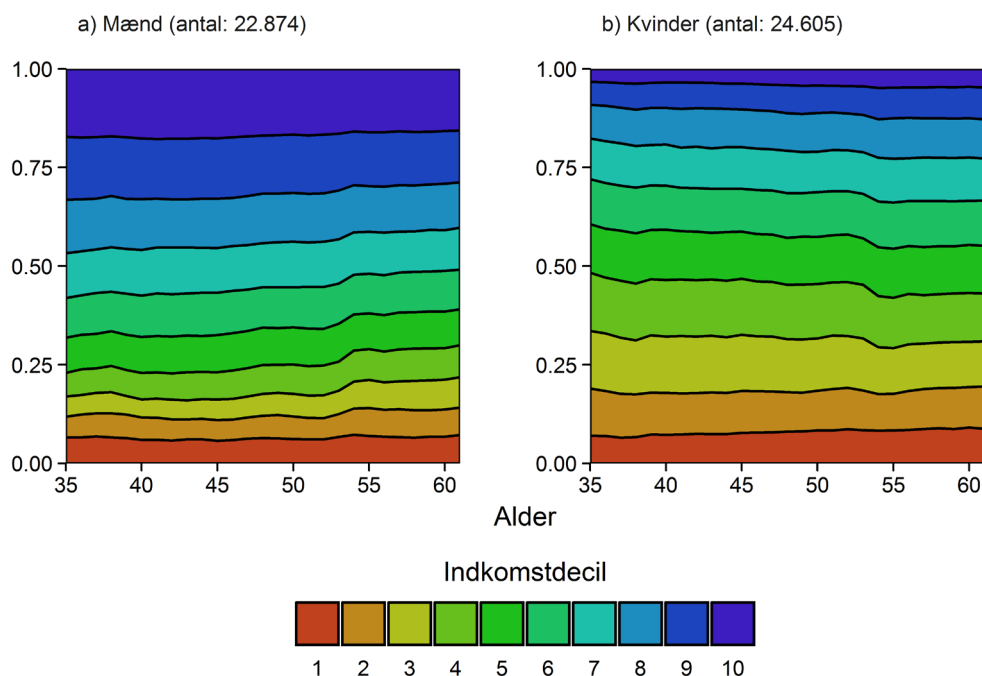
At der ikke er komplet repræsentation understreger, at det er vigtigt at fortolke analysens resultater med afsæt i kriterierne for valg af målgruppe. Der benyttes aldersbetingede indkomstdeciler for at eliminere tendensen til, at placering i et højt indkomstdecil alene kan

¹⁵ Indkomst er udtrykt ved den samlede personindkomst ekskl. beregnet lejeværdi af egen bolig og før fradrag af renteudgifter. Personindkomsten består således af erhvervsindkomst (herunder løn og nettooverskud af egen virksomhed), offentlige overførsler, udbetalte private pensioner, renteindtægter og realiserede tab/gevinster på værdipapirer samt restindkomst.

være udtryk for en alderseffekt. Eliminering af denne aldersbias vil dog primært have betydning for grupperingen af livsforløb i en situation, hvor indkomstforholdene for forskellige generationer følges gennem den samme tidsperiode. Eftersom der her udelukkende betragtes indkomstforløb for én generation er betydningen altså mindre.

Målgruppen omfatter en lidt større andel af kvinder (51,8 procent) end mænd (48,2 procent). Omkring 47 procent af mændene befinder sig som 35-årige i de tre øverste indkomstdeciler. For kvindernes vedkommende er det blot omkring 18 procent, jf. Figur 4.2. Over tid forbedres kvindernes position i indkomstfordelingen, hvilket delvist medvirker til en fortrængning af mændene i de højeste indkomstdeciler. Dette kan med stor sandsynlighed tilskrives en interaktion med andre personkarakteristika, der har ændret sig gennem den betragtede periode. På baggrund af Figur 4.2 skønnes køn umiddelbart at have betydning for indkomststrukturen og dermed forklare en del af forskellen i de individuelle indkomstforløb.

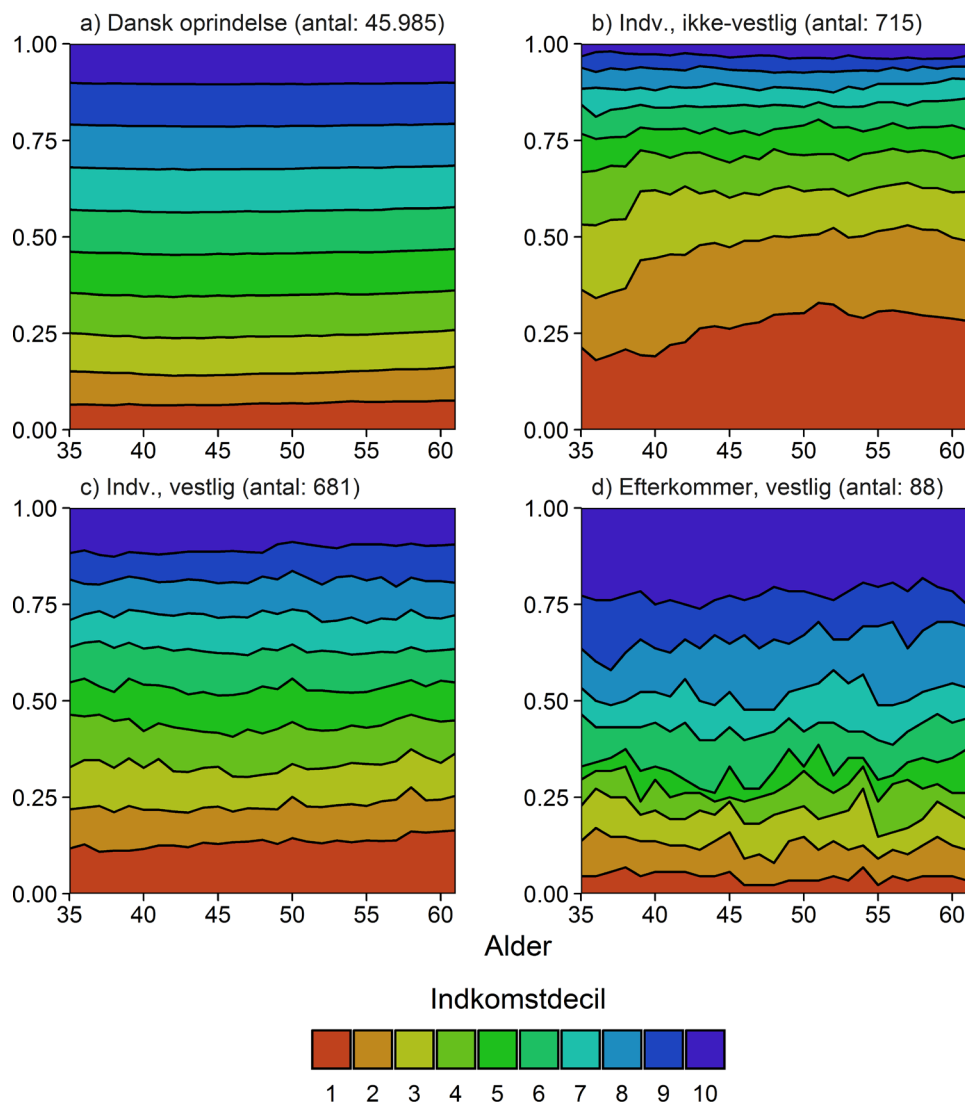
Figur 4.2. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986 fordelt på køn.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Antallet af personer, der ikke har dansk oprindelse, omfatter færre end 1.500 personer. Dette giver for de pågældende grupper anledning til en del fluktuation i indkomststrukturen over tid. Overordnet minder vestlige indvandreres udvikling i indkomstsammensætningen om den, der er gældende for personer af dansk oprindelse, jf. Figur 4.3. Ikke-vestlige indvandrere er kendetegnet ved i højere grad at optræde i de nederste indkomstdeciler, mens efterkommere fra vestlige lande har en relativt større tilbøjelighed til at optræde i de øverste deciler end personer af dansk oprindelse. Det skal dog bemærkes, at målgruppen omfatter færre end 100 vestlige efterkommere, hvorfor det kan være en udfordring at drage robust statistisk inferens omkring, hvorvidt indkomststrukturen i denne oprindelsesgruppe afviger signifikant fra strukturen i de øvrige befolkningsgrupper. Umiddelbart synes oprindelse dog at kunne forklare en del af variationen i indkomstforløb, men idet der specielt for ikke-vestlige indvandreres vedkommende er tale om en ændring i strukturen over tid, må det formodes, at effekten stammer fra korrelation med et andet personkarakteristikum.

Figur 4.3. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986 fordelt på oprindelse.



Anm.: Der er blot 10 ikke-vestlige efterkommere i målgruppen, som er udeladt ovenfor.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

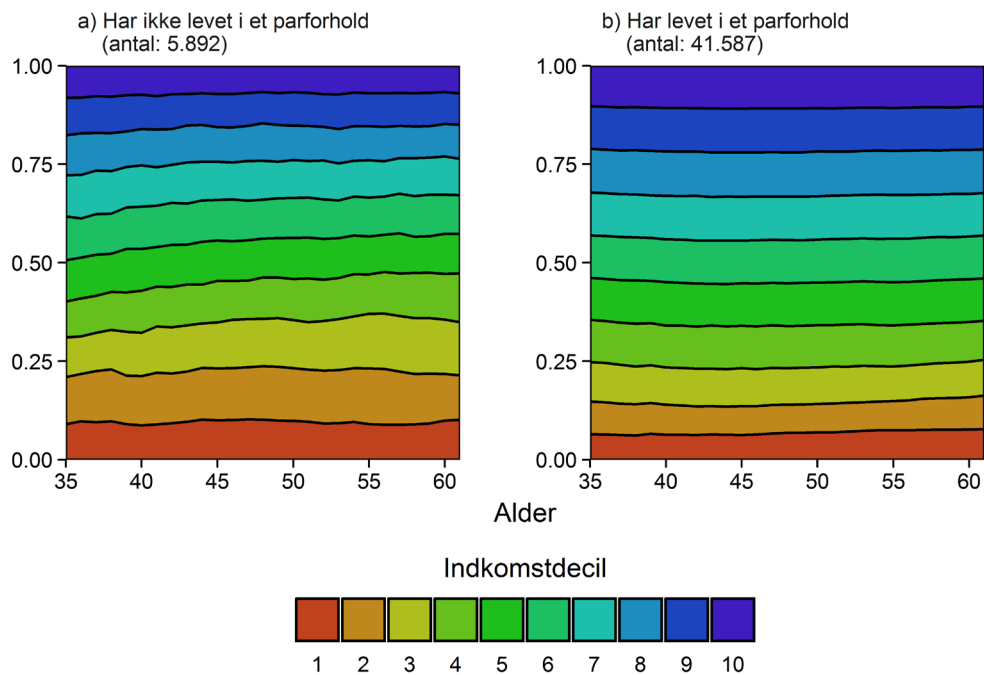
I modsætning til køn og oprindelse er målgruppens familietypestatus en dynamisk størrelse, der kan variere gennem såvel analyseperioden som perioden forud herfor. Således vil indkomststrukturen her blot blive sammenlignet på tværs af to grupper, der er kendetegnet ved henholdsvis ingen og mindst ét års parforholdserfaring i alderen 30-34 år. Det viser sig, at der er en betydelig sammenhæng mellem parforholdserfaring i alderen 30-34 år og 35-61 år. Således vil 98 procent af de personer, der har parforholdserfaring fra den unge aldersgruppe også have det i den ældre gruppe. Omvendt vil 93 procent, der lever i par i alderen 35-61 år også have gjort det i perioden forud herfor. Andelen bliver mindre set fra den anden periodes synspunkt, da antallet med parerfaring her er øget relativt til den første periode. Det er tænkeligt, at der kan være en gensidig påvirkning mellem indkomst og tilbøjeligheden til at leve i par. Graden af kontinuitet i parforholdsstatus mellem aldersgrupperne retfærdiggør, at man kan lave en vurdering af forholdet mellem familietype og indkomst baseret på parforholdserfaring i alderen 30-34 år, der per definition er upåvirket af indkomstudviklingen i alderen 35-61 år. Således kan simultanitetsbias i de senere multivariate analyser undgås.

Som nævnt indledningsvist leder dette dog ikke til kausal inferens pga. den fortsatte tilstedeværelse af andre typer af bias.

Med afsæt i ovenstående er indkomststrukturen nedenfor nuanceret på tværs af parforholdserfaring i alderen 30-34 år. Betragtes variationen i indkomststrukturen på tværs af familietypeerfaring, kan det konstateres, at der er stor forskel på det antal, der i mindst et år i løbet af perioden har indgået i et parforhold og det antal, der konsekvent har levet som enlige. Eftersom hovedparten af befolkningen på et tidspunkt i løbet af perioden har indgået i en parrelation, vil indkomststrukturen for de pågældende naturligt minde om strukturen for målgruppen som helhed, jf. Figur 4.4.

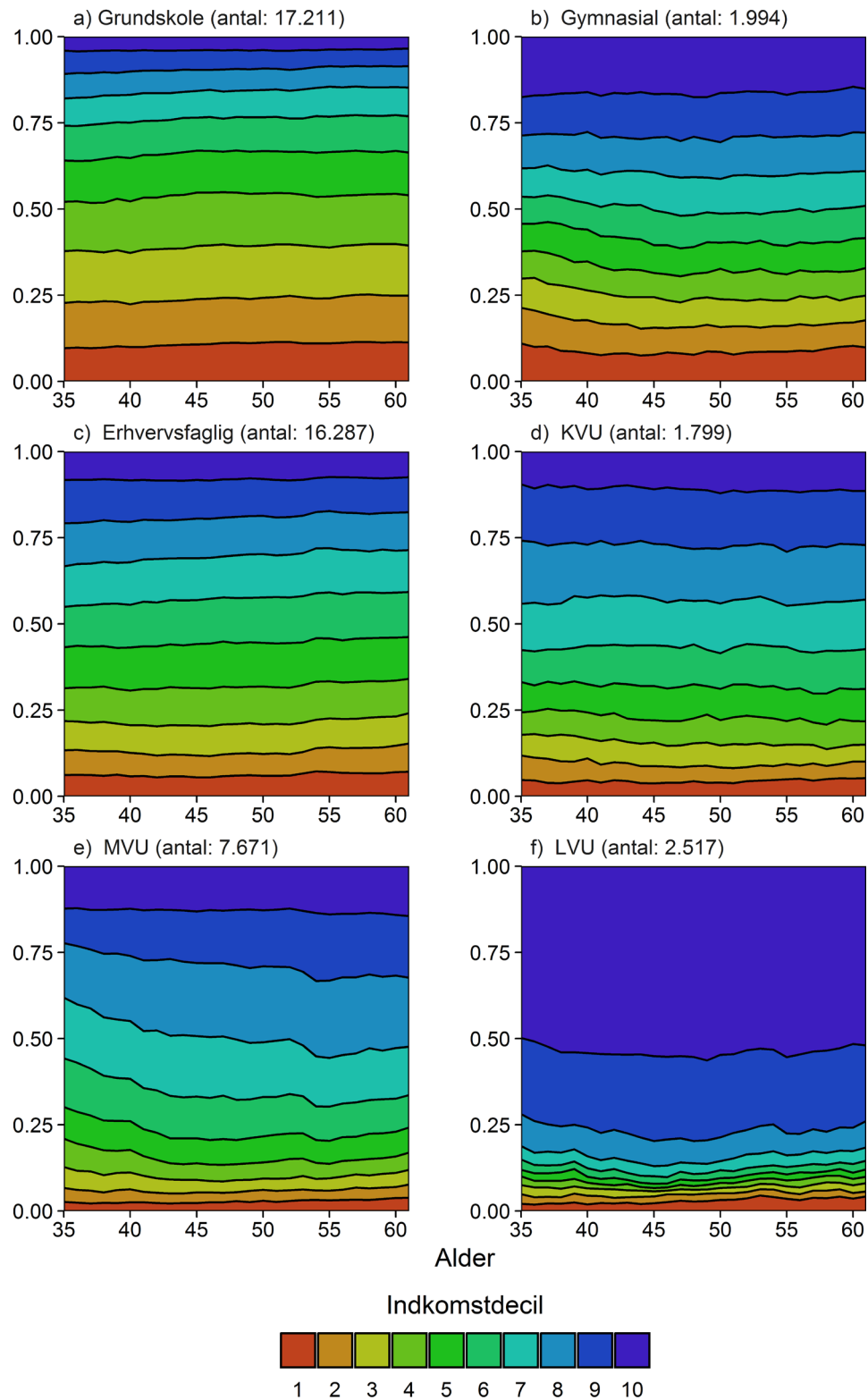
Indkomststrukturen for enlige er præget af en relativt større tilbøjelighed til at opholde sig i de nederste indkomstdeciler. Således vil omkring 57 procent på sigt være tilknyttet de fem nederste indkomstdeciler, mens det blot er tilfældet for 46 procent af gruppen med parforholdserfaring. På denne baggrund kan parforholdserfaring i evt. samvariation med andre karakteristika formodes at indvirke på grupperingen af indkomstforløb og forklare en del af den indbyrdes diversitet.

Figur 4.4. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986 fordelt på parforholdserfaring i alderen 30-34 år.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 4.5. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986 fordelt på uddannelsesniveaue som 35-årig.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

At målgruppen er afgrænset til at omfatte personer, der i starten af deres indkomstforløb er 35 år, udspringer af, at befolkningen typisk har afsluttet deres uddannelsesforløb på dette tidspunkt og dermed har en lav tendens til at skifte niveau for højst fuldførte uddannelse i løbet af resten af deres livsforløb. Således kan i nogen udstrækning bortses fra, at det for visse personer er skift i uddannelsesniveaue, der afstedkommer skift i individets placering i indkomstfordelingen. Indkomststrukturen i perioden kan således ideelt blot beskrives ved uddannelsesniveaue på et vilkårligt tidspunkt. Ovenfor er indkomstforløbene opdelt efter uddannelsesniveaue ved 35-årsalderen og billedet er nogenlunde det samme, hvis der i stedet fokuseres på uddannelsesniveaue ved et højere alderstrin¹⁶. Med en vis rimelighed kan man derfor opfatte uddannelsesniveaue som en statistisk variabel. Selv om det således ikke har den store betydning hvorvidt det uddannelsesniveaue, der anvendes i de senere estimationer, er opgjort ved 35-årsalderen eller i slutningen af perioden, er det valgt at benytte niveauet ved 35-årsalderen for at mindske graden af simultanitetsbias¹⁷. Det er således også dette uddannelsesniveaue, der benyttes til nedenstående kategorisering af indkomstforløb. En gensidig relation mellem indkomstforskelle og uddannelsesniveaue vurderes i øvrigt at være af mindre betydning.

Ikke overraskende er der overordnet tendens til, at andelen af befolkningen i de øverste indkomstdeciler er stigende i uddannelsesniveaue og der er derfor belæg for at forvente, at uddannelsesniveaue spiller en rolle i kategoriseringen af indkomstforløb og yderligere må formodes i nogen udstrækning at kunne redegøre for indbyrdes forskelle i indkomstforløb over den betragtede periode.

Målgruppens arbejdsmarkedstilknytning kan potentielt ændre sig gennem analyseperioden og de fem år umiddelbart forud for denne. Der eksisterer oplagt en gensidig påvirkning mellem arbejdsmarkedstilknytning og udviklingen i indkomstforløb. Med henblik på at kunne kontrollere for simultanitetsbias i de senere analyser, vil det derfor være nyttigt, hvis der også her kunne påvises en vis grad af persistens mellem arbejdsmarkedstilknytning i alderen 30-34 år og i alderen 35-61 år. Tilknytning til arbejdsmarkedet defineres her som mindst ét års kategorisering som enten beskæftiget uden nærmere angivelse, beskæftiget studerende eller ledig¹⁸.

Der viser sig at være en betydelig sammenhæng mellem erhvervstilknytning i alderen 30-34 år og i resten af den betragtede periode. Således vil 99 procent af de personer, der er tilknyttet arbejdsmarkedet i den yngre aldersgruppe også have et eller flere års tilknytning til arbejdsmarkedet i alderen 35-61 år. Af de, der ikke har mindst ét års arbejdsmarkedstilknytning i alderen 30-34 år, vil 41 procent heller ikke have det efterfølgende.

¹⁶ For at sikre en større stabilitet i uddannelsesniveaue over tid kunne alternativt betragtes en ældre generation i 1986, eksempelvis 40-årige. Dette ville dog have den ulempe, at man specielt for personer uden en lang videregående uddannelse ville miste vigtig information om indkomstforholdene tidligere i arbejdslivet, der potentielt kan være afgørende for den videre indkomstkarrere. Endvidere må anvendelsen af en ældre generation som målgruppe formodes at mindske andelen af indvandrere og efterkommere yderligere og dermed umuliggøre, at oprindelsesstatus kan identificeres som et karakteristika med betydning for den indbyrdes forskel i indkomstforløb.

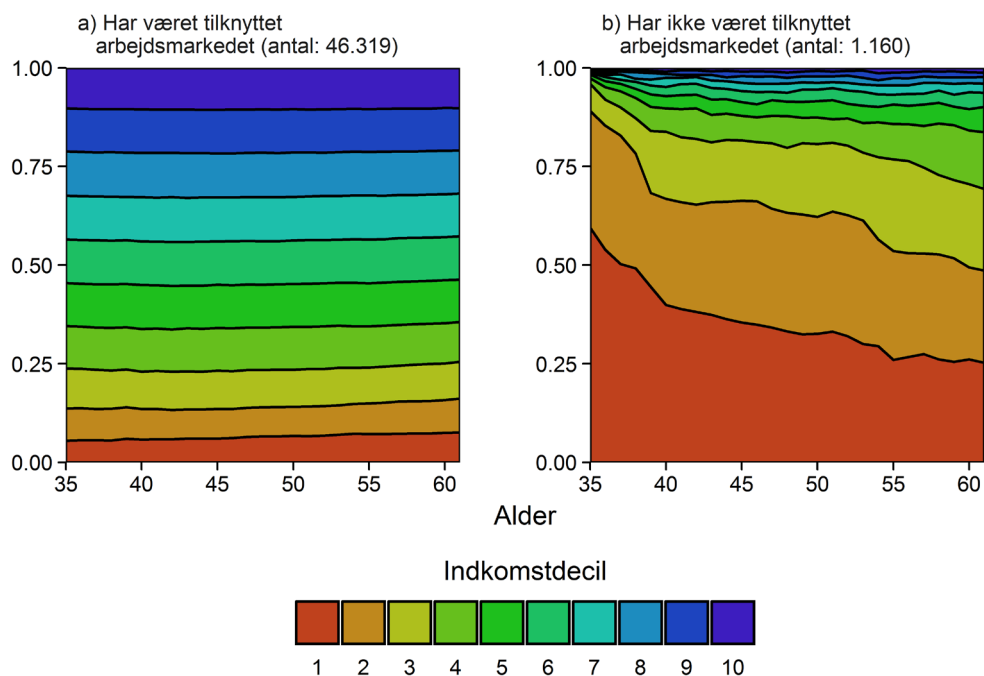
¹⁷ Man kunne af hensyn til simultanitetsproblemet have valgt uddannelsesniveaue som 30-årig, men idet der sker en del forskydninger i målgruppens niveau, er niveauet for 35-årige valgt som en bedre repræsentation for det endeligt opnåede uddannelsesniveaue, selv om dette betyder et års overlap med den betragtede indkomstperiode.

¹⁸ Overordnet er tilordningen af befolkningen på arbejdsmarkedsstatus fastlagt på baggrund af Registerbaseret ArbejdsstyrkeStatistik (RAS). Der er bortset fra antallet af arbejdstimer i perioden, hvorfor nuanceringen udelukkende afspejler forskelle i indkomststrukturen på tværs af deltagelse og ikke fuldtidsdeltagelse. Indkomststrukturen vil dog afspejle kombinationen af deltagelse, timer og produktivitet.

I afsnit 4.2 ses nærmere på, hvorledes arbejdsmarkedstilknytningen i aldersgruppen 35-61 år er sammensat på anciennitet og tilhørsforhold til arbejdsstyrken i alderen 30-34 år.

På baggrund af ovenstående er indkomststrukturen i Figur 4.6 nuanceret på, om man i alderen 30-34 år har været tilknyttet arbejdsmarkedet. Som forventeligt varierer strukturen betydeligt afhængigt af graden af arbejdsmarkedstilknytning. Langt hovedparten af målgruppen har arbejdsmarkedserfaring og indkomststrukturen for denne del gengiver derfor den overordnede tendens fra Figur 4.1. Personer uden erhvervstilknytning er karakteriseret ved stort set entydigt at befinde sig i de tre laveste indkomstdeciler i 35-årsalderen, jf. Figur 4.6b. På sigt vil dette dog kun vedrøre ca. 75 procent, hvor faldet skal forklares af en mindsket andel i det laveste indkomstdecil på bekostning af en stigning i decilerne omfattende personer med de 20-40 procent laveste indkomster. Den ændrede struktur kan meget vel forklares af, at personer uden tilknytning til arbejdsmarkedet i alderen 30-34 år gradvist indtræder på arbejdsmarkedet efterfølgende.

Figur 4.6. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986 fordelt på tilknytning til arbejdsmarkedet i alderen 30-34 år.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Betragtes i stedet indkomststrukturen fordelt på tilknytning til arbejdsmarkedet i alderen 35-61 år (ikke illustreret) fås samme kvalitative struktur som ovenfor. Dog er gruppen helt uden arbejdsmarkedserfaring blevet væsentligt mindre og er karakteriseret ved næsten konsekvent at omfatte personer i de tre nederste indkomstdeciler. I Figur 4.7 ses udviklingen i indkomststrukturen for personer i alderen 35-61 år fordelt på arbejdsmarkedsanciennitet. Der betragtes udelukkende personer med mindst ét års erhvervstilknytning i løbet af den betragtede periode, mens der ikke er betinget på graden af tilknytning til arbejdsmarkedet i den yngre aldersgruppe. Der sondres heller ikke mellem, hvorvidt ancienniteten er optjent gennem en sammenhængende periode eller om forløbet på arbejdsmarkedet har været afbrudt.

Generelt gælder, at personer med færre end både 10 og 20 års tilknytning til arbejdsmarkedet har en udpræget tilknytning til de laveste indkomstdeciler, som enten øges midlertidigt eller permanent gennem perioden, jf. Figur 4.7a og b.

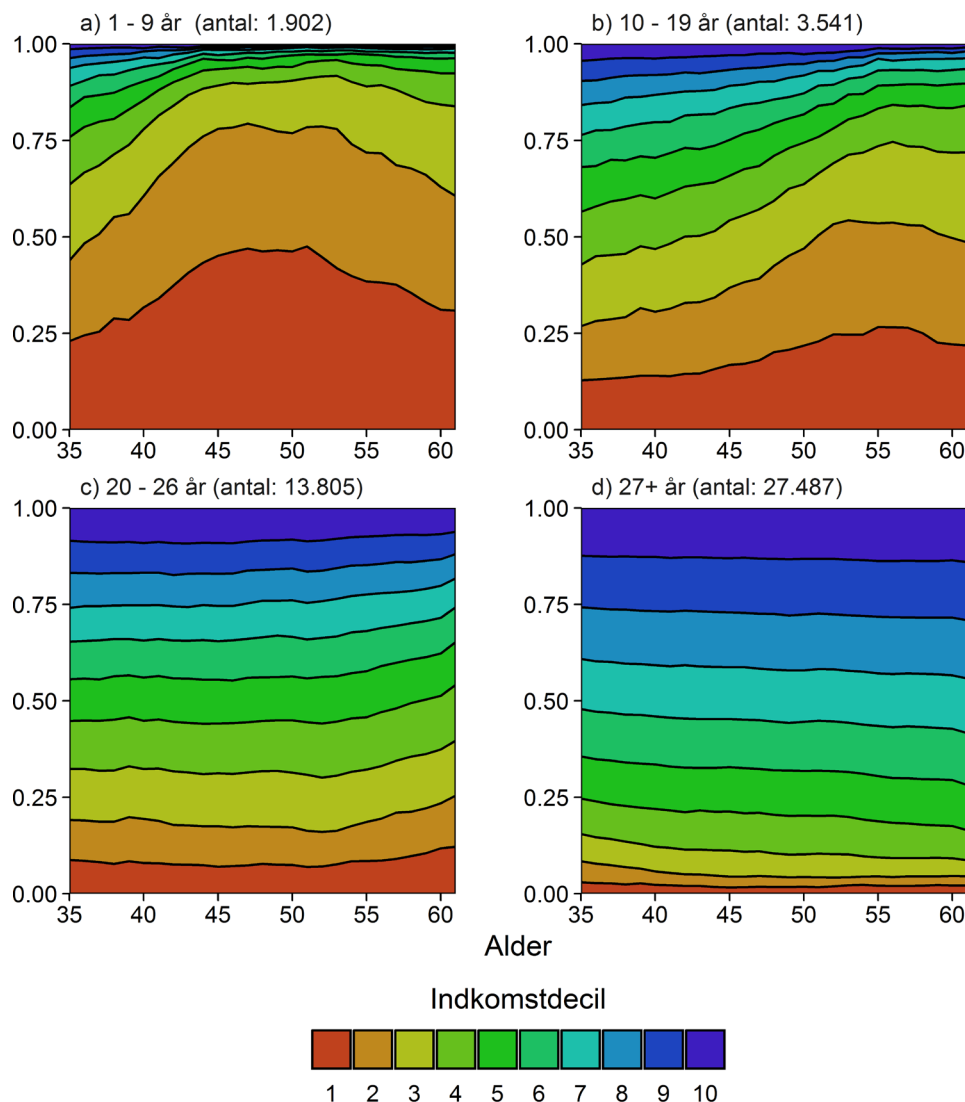
Gruppen med den laveste anciennitet er kendetegnet ved allerede i periodens begyndelse at omfatte næsten 30 procent uden tilknytning til arbejdsmarkedet, hvoraf den største del ud over førtidspension betegnes som øvrige udenfor arbejdsstyrken, der her bl.a. omfatter personer, der modtager kontanthjælp, sygedagpenge eller som ikke er berettiget til ydelser. På sigt vil stort set alle i anciennitetsgruppen modtage førtidspension. 87 procent af gruppen har mindst ét års tilknytning til arbejdsmarkedet i alderen 30-34 år. Gruppen består primært af personer, der på sigt har en grundskoleuddannelse og i mindre grad en erhvervsfaglig uddannelse som højst fuldførte.

Gruppen omfattende personer med 10-19 års anciennitet er initialt kendetegnet ved at omfatte 87 procent med tilknytning til arbejdsmarkedet, hvor en relativt stor del heraf er ledige. På sigt vil næsten 70 procent dog være tilknyttet førtidspension, mens tilbagetrækning i form af efterløn er sjælden, idet arbejdsmarkedsancienniteten i gruppen udfordrer opfyldelse af kravet for deltagelse i efterlønsordningen. Hele 95 procent af gruppen har været tilknyttet arbejdsmarkedet i et eller flere år i alderen 30-34 år. Således synes der at være tale om en gruppe, der initialt har haft et fornuftigt fodfæste på arbejdsmarkedet, men hvor gradvis overgang til enten midlertidigt eller permanent ophold uden for arbejdsmarkedet har givet anledning til en gradvis forøgelse af antallet af personer i de laveste indkomstdeciler. Gruppen har stort set samme uddannelsessammensætning som den del af målgruppen, der har den laveste anciennitet, men med en mere ligelig fordeling mellem grundskole- og erhvervsfagligt uddannede.

Den næststørste anciennitetsgruppe med 20-26 års erfaring på arbejdsmarkedet omfatter næsten 30 procent af målgruppen. Flertallet i gruppen vil være tilknyttet de fem laveste indkomstdeciler, hvilket er en tendens, der øges svagt over tid. Dette afspejler delvist en gradvis tilbagetrækning fra arbejdsmarkedet i form af førtidspension, men vil mod slutningen af perioden primært være udtryk for overgang til efterløn. 98 procent af gruppen vil også have været tilknyttet arbejdsmarkedet i alderen 30-34 år. Næsten 40 procent af gruppen har på sigt opnået en erhvervsfaglig uddannelse, mens cirka 30 procent og 27 procent har opnået hhv. en grundskoleuddannelse eller en videregående uddannelse.

Mere end halvdelen af målgruppen har været tilknyttet arbejdsmarkedet gennem hele den betragtede periode og har en indkomststruktur, der er kendetegnet af, en gennem perioden, stadigt større tilknytning til de øverste indkomstdeciler på bekostning af tilknytningen til de laveste deciler, jf. Figur 4.7d. Dette er på trods af en svag stigning i antallet af ledige over tid. Næsten samtlige personer i gruppen har også erhvervs erfaring fra perioden i alderen 30-34 år. Erhvervsfaglige udgør også her næsten 40 procent, mens andelen med en videregående uddannelse udgør 33 procent og grundskoleuddannede 24 procent af gruppen.

Figur 4.7. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986 fordelt på antal års tilknytning til arbejdsmarkedet i alderen 35-61 år.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Målgruppens boligtype varierer betydeligt gennem livet. Dette afspejler blot, at søgningen mod de enkelte boligtyper er knyttet til forskellige faser i livsforløbet, der bl.a. er afgrænset af tidspunktet for stiftelse af familie, ændring i antallet af hjemmeboende børn, udvikling i indkomstforhold, tilbagetrækning fra arbejdsmarkedet og alderdom. De to sidstnævnte faser er grundet valget af målgruppe ikke aktuelle i det givne tilfælde. At målgruppens boligkarakteristika dermed er en dynamisk størrelse, betyder alt andet lige, at der med stor sandsynlighed vil være forskel på, om indkomststrukturen opgøres fordelt på boligtypen i starten eller slutningen af perioden. Man kan sige, at boligtypen for 35-årige i et vist omfang afspejler målgruppens fremtidige muligheder på boligmarkedet. Har man allerede en ejerbolig på dette tidspunkt, vil man med stor tilbøjelighed også have det fremadrettet, mens skift til ejerbolig er mindre sandsynligt, hvis man i starten af perioden bor i en almen bolig. Boligtypen som 61-årig er omvendt en sammenfatning af de begivenheder, der er indtruffet forud for dette tidspunkt. Eksempelvis kan gælde, at sandsynligheden for at bo i en ejerbolig tæt på tilbagetrækningstidspunktet er større, hvis man har haft en relativt stabil tilknytning til arbejdsmarkedet.

Det viser sig, at indkomststrukturen varierer betydeligt mere både inden for og på tværs af boligtype, hvis denne opgøres som 61-årig end som 35-årig. Dette synes at afspejle, at boligtypen som 61-årig omfatter en større variation i livshistorier end boligtypen som 35-årig. Relativt til ønsket om at undgå simultanitetsbias, er der umiddelbart ikke noget, der peger entydigt i retning af, at opgørelsen af indkomst på boligtype er mest fordelagtig i starten eller slutningen af analyseperioden: boligtypen som 35-årig kan påvirke den fremtidige indkomst ved eksempelvis salg af ejer- eller andelsbolig, men fremtidig indkomst ikke vil påvirke boligsituationen for en 35-årig direkte¹⁹. Indkomsten i perioden 35-61 år, vil dog have indflydelse på, hvilken boligtype man vil bo i mod slutningen af perioden, mens boligtypen som 61-årig ikke direkte kan påvirke indkomsten forud herfor²⁰. For at skabe konsistens med nuanceringen af indkomst på de øvrige karakteristika, er det her valgt at betragte boligtypen som 35-årig.

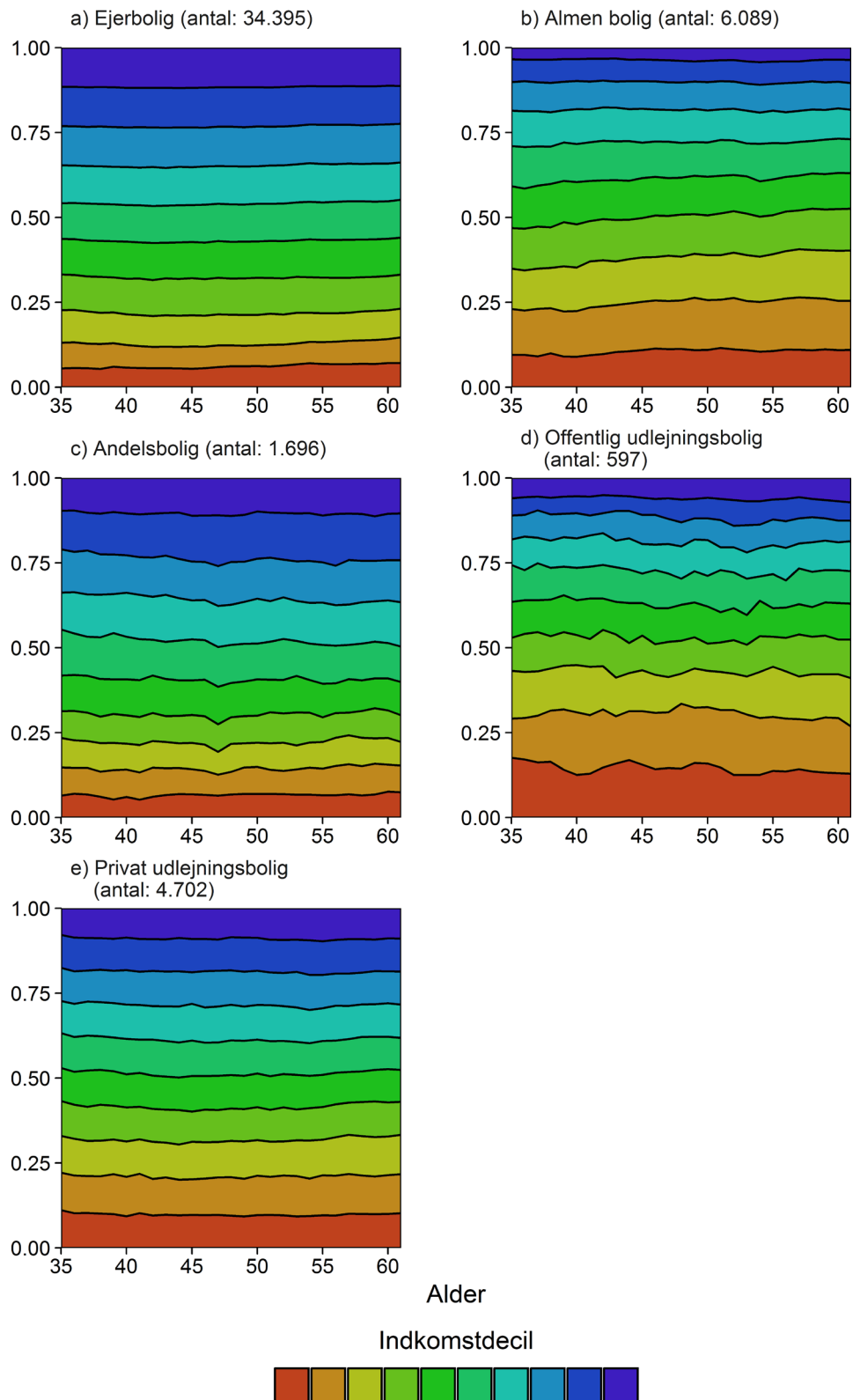
I Figur 4.8 er indkomststrukturen illustreret fordelt på boligtype i starten af perioden. Hvis man betragter gruppernes størrelse, er den væsentligste forskel til en opgørelse baseret på 61-åriges boligtype, at målgruppen som 35-årige har en mindre tendens til at bo i andelsbolig og en større tendens til at bo i en udlejningsbolig. Antallet af personer i en almen bolig eller en ejerbolig er stort set upåvirket af opgørelsestidspunktet. Uafhængigt af opgørelsestidspunktet vil ejerbolig være den alt dominerende boligtype og indkomststrukturen afspejler dermed i vidt omfang målgruppens overordnede struktur.

Indkomststrukturen synes relativt ensartet for den del af målgruppen, der bor i enten ejerbolig, andelsbolig eller privat udlejningsbolig, mens strukturen for almen bolig og offentlig udlejningsbolig domineres af ophold i de laveste indkomstdeciler. Målgruppens boligtype som 35-årig kan på denne baggrund repræsentere en del af variationen i indkomstforløb og være med til at fastlægge, hvorfor nogle forløb i det kommende afsnit klassificeres i én gruppe frem for en anden.

¹⁹ Med direkte menes, at der her er bortset fra den bias, der er knyttet til ikke-observerbare forhold: Hvis en 35-årig eksempelvis har en bestemt forventning til sine fremtidige indkomstmuligheder, kan dette meget vel påvirke valg af boligtype på nuværende tidspunkt.

²⁰ Medmindre den er opnået ved køb en del år forud herfor. Er man eksempelvis bosat i en ejerbolig som 61-årig, er der en udbredt tendens til, at man også var det tidligere eller helt i starten af perioden. Har man oplevet gentagne køb og salg inden for den samme boligtype undervejs, vil boligtypen som 35-årig både have indflydelse på det fremtidige indkomstforløb og på boligtypen som 61-årig. Selv om simultanitetsbias ikke er oplagt, vil der altså her være tale om endogenitet fra en 3. variabel, der både er korreleret ved fremtidig indkomst og boligtypen i slutningen af perioden. En sådan 3. variabel er også enkel at finde i forholdet mellem boligtype som 35-årig og fremtidig indkomst, eksempelvis uddannelse.

Figur 4.8. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986 fordelt på boligtype som 35-årig.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Områdetyper for en persons bopæl, her afgrænset som udsat/ikke-udsat, er som flere af de øvrige karakteristika dynamisk og kan dermed ændre sig over tid. Dette kan enten skyldes, at en person flytter fysisk fra én områdetype til en anden, men kan også skyldes, at området skifter kategorisering pga. en ændring i beboernes karakteristika. Som nævnt tidligere vil der her blot blive set på egenskaberne ved en persons bopælsområde og ikke blive sondret mellem, hvorvidt individuelle skift skal tilskrives ændringer i områdekategoriseringen eller en flytning.

Det kan konstateres, at der også her er en vis sammenhæng mellem erfaring med bosættelse i et udsat område i alderen 30-34 år og bosættelse i et udsat område i resten af det betragtede livsforløb. Således vil 59 procent af de personer, der i den unge aldersgruppe har været bosat et eller flere år i et udsat område også være det senere i livet. Har man ikke bopæls erfaring fra et udsat område i alderen 30-34 år, vil kun omkring 5 procent få det senere i livet. Omvendt vil omkring 40 procent af de personer, der i alderen 35-61 år bor i et udsat område også have gjort det tidligere i livet. Blandt dem, der ikke bor i et udsat område i alderen 35-61 år, er der færre end 3 procent, der har gjort det som 30-34-årige.

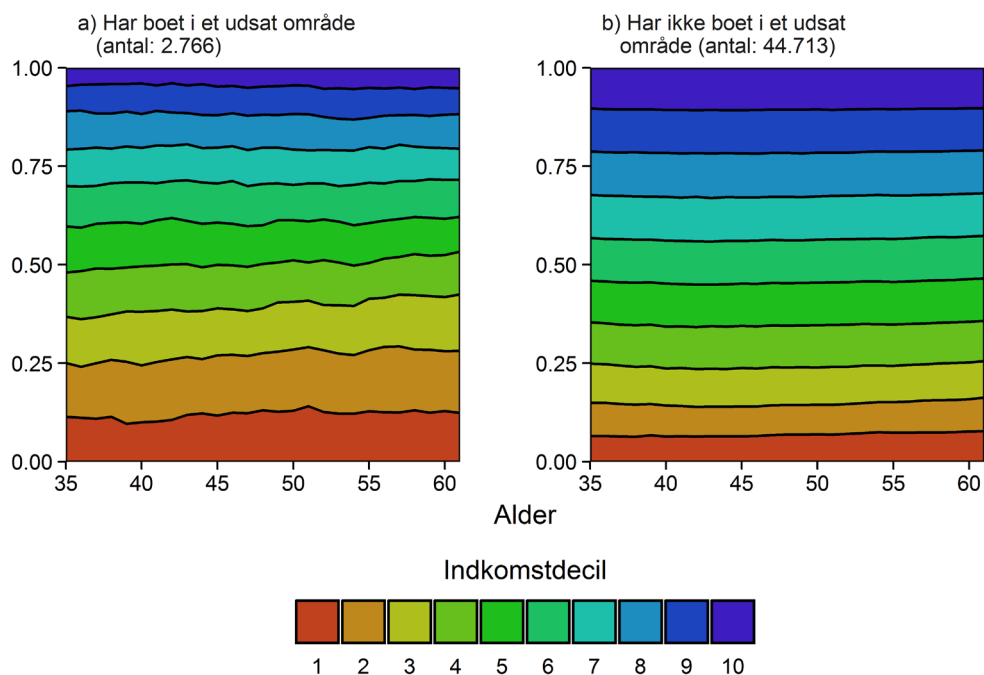
Et af de store spørgsmål, der rejser sig, når forholdet mellem indkomst og bosætning i et udsat område betragtes, er udredning af det kausale forhold. Påvirker områdets karakteristika beboernes indkomstmuligheder eller er indkomststrukturen i et udsat område et resultat af, at områdetypen tiltrækker en bestemt type af beboere? Flere undersøgelser peger på, at bosætningsområdet kan påvirke et individs socioøkonomiske karriere. Dette skal tilskrives, at individuelle beslutninger og adfærdsmønstre tilpasses i forhold til hvorledes andre i nabolaget agerer, jf. Akerlof (1997) og Banerjee (1992). Der er en tendens til, at personer bosat i et udsat område socioøkonomisk vurderet er relativt svagere end andre, jf. Vasiljeva & Mairey (2016). Dette behøver dog ikke nødvendigvis at kunne henføres til påvirkning fra området, men kan potentielt tilskrives, at der er en tendens til, at de svageste vælger at bosætte sig i de udsatte områder og der dermed er tale om selvselektion, der både kan udspringe af præferencer for bestemte naboer og begrænsede muligheder for bosætning i andre områder.

Som tidligere nævnt vil vi i dette afsnit ikke beskæftige os med udredning af det kausale forhold og udgangspunktet vil derfor blot være, at der muligvis eksisterer en gensidig påvirkning og dermed en risiko for simultaneitetsbias. Sammenhængen mellem personens områdestatus tidligt og senere i livsforløbet udnyttes derfor i de senere estimationer for at mindske simultaneitetsproblemerne. I Figur 4.9 er indkomststrukturen illustreret for personer, der har enten ingen eller mindst ét års bopæls erfaring i et udsat område i alderen 30-34 år.

Næsten 6 procent af målgruppen har på et tidspunkt været bosat mindst 1 år i et udsat område, mens de var i alderen 30-34 år, jf. Figur 4.9. I og med, at hovedparten af målgruppen således ikke har boet i et udsat område i den unge aldersgruppe, vil indkomststrukturen for denne gruppe minde om strukturen for målgruppen som helhed. Personer med ophold i et udsat område vil typisk have en større tendens til at opholde sig i de nederste indkomstdeciler.

I gennemsnit har omkring 60 procent af gruppen på et vilkårligt tidspunkt opholdt sig i en af de fem nederste deciler. Dette er kun tilfældet for godt 45 procent af befolkningen uden ophold i et udsat område. Kvalitativt minder forskellen i indkomststruktur om den, der var gældende for parforholdserfaring og det viser sig faktisk også, at personer uden parforholdserfaring i alderen 30-34 år har større tendens til at have boet i et udsat område i dette aldersinterval (10,8 procent) end dem med parforholdserfaring (5,1 procent).

Figur 4.9. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986 fordelt på bopælsområdeerfaring i alderen 30-34 år.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Blandt de personer, der i aldersgruppen 35-61 år er bosat i et udsat område i et eller flere år, er indkomststrukturen i Figur 4.10 fordelt på antal års bosættelse. Der er i opgørelsen ikke betinget på, hvorvidt man i den unge aldersgruppe har været bosat i et udsat område. Desuden afspejler opgørelsen ikke nødvendigvis længden af en sammenhængende bosætningsperiode.

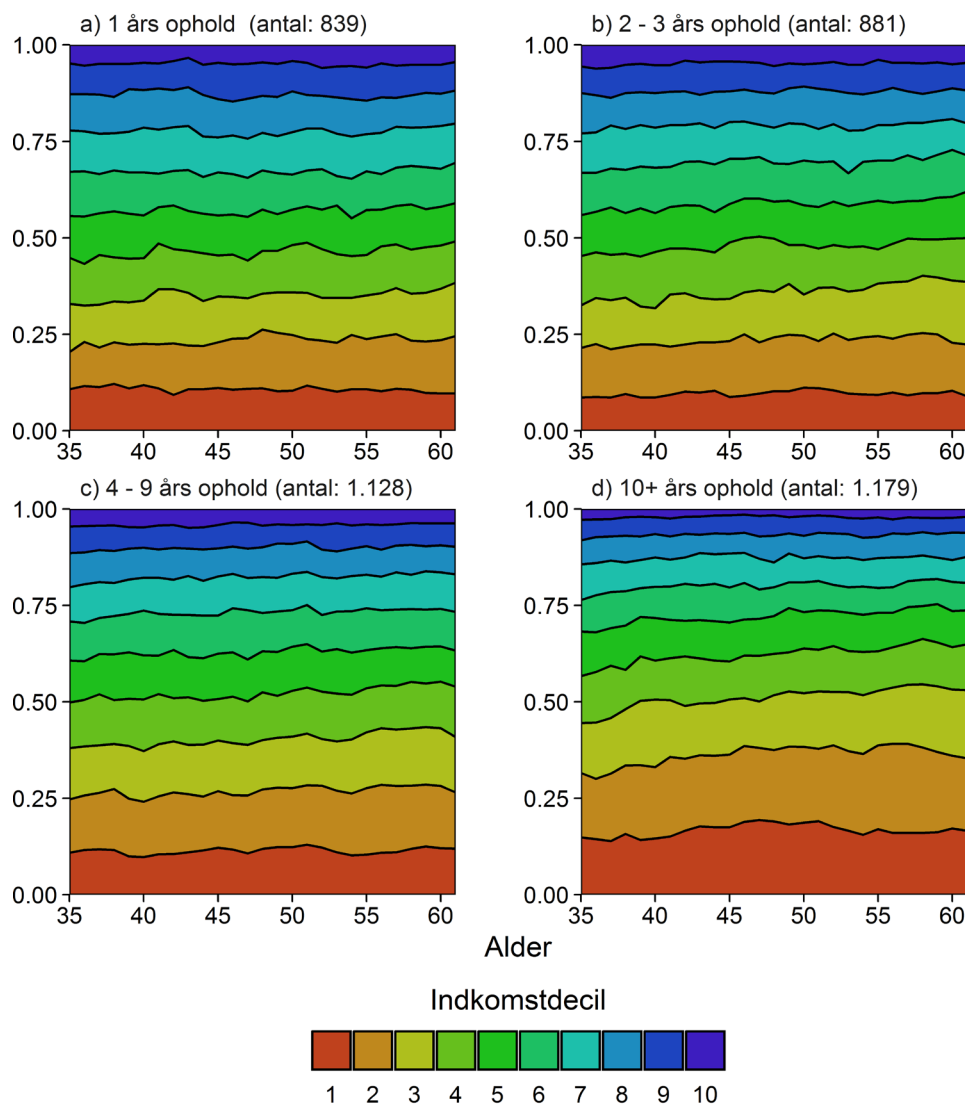
Første gruppe omfatter personer, der har boet i et udsat område i 1 år, 2. gruppe omfatter personer med 2-3 års bosættelse, mens 3. og 4. gruppe omfatter personer med hhv. 4-9 og 10+ års bosættelse. Der synes at være en tendens til, at andelen af målgruppen i de nederste indkomstdeciler er stigende i antallet af år i et udsat område, men tendensen er mest udtalt mellem 2., 3. og 4. gruppe. Det viser sig, at der er en positiv sammenhæng mellem antallet af års bosættelse i alderen 35-61 år og sandsynligheden for også at have boet i et udsat område mindst ét år som 30-34-årig. Således vil ca. 24 procent af de personer, der i alderen 35-61 år har opholdt sig 1 år i et udsat område også have boet i et sådant område på et eller andet tidspunkt i alderen 30-34 år, mens det samme er tilfældet for 34 procent af den del af målgruppen, der har 2-3 års bosættelse i et udsat område i alderen 35-61 år. Har man boet i et udsat område 4-9 år eller 10+ år i alderen 35-61 år, så vil henholdsvis 40 procent og 59 procent også have gjort det mindst ét år som 30-34-årige.

Indkomststrukturen er oplagt forskellig over livet afhængig af om man har erfaring med bosætning i et udsat område i den unge aldersgruppe. Heraf kan man dog ikke konkludere, at det er forskellen i bosætningserfaring, der giver anledning til forskellen i sammensætningen på indkomstdeciler. Jævnfør kausalitetsdebatten er det ikke nødvendigvis områdeegenskaberne i sig selv, der giver anledning til et bestemt indkomstforløb. Der kan være tale om, at personer, der har boet i et udsat område kan have strukturelle karakteristika, der uagtet deres tidligere bosætningsforhold vil gøre, at de adskiller sig socioøkonomisk fra resten af målgruppen. Selv om det i de senere estimationer er muligt at kontrollere for graden af arbejdsmarkedstilknytning og uddannelsesniveaue, vil der fortsat være effekter, der kan

være vanskelige at måle og derfor tage højde for, eksempelvis positive effekter fra bedre muligheder for at udnytte et socialt netværk og negative effekter fra dårlig *ability*²¹.

Når indkomststrukturen i Figur 4.9 og Figur 4.10 varierer på tværs af områdeerfaring og opholdstid, kan dette altså afspejle, at bosætning er korreleret med andre socioøkonomiske forhold. At dette synes at være tilfældet motiveres af, at indkomststrukturen varierer over tid inden for grupperne afgrænset af hhv. bosætningserfaring og opholdstid.

Figur 4.10. Indkomststruktur i perioden 1991-2017 for personer, der var 30 år i 1986 fordelt på antal års bosættelse i et udsat område i alderen 35-61 år.



Anm.: Der ses udelukkende på personer, der på et tidpunkt i løbet af deres liv i perioden 1991-2017 har været bosat i et udsat område. Længden af opholdet er ikke nødvendigvis opgjort over en sammenhængende periode.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

²¹ Dårlig *ability* anvendes her som et bredt begreb, der kan dække over flere forhold såsom manglende kendskab til, hvorledes der kan skaffes hjælp til at få fodfæste på arbejdsmarkedet eller svære betingelser for at indtræde på arbejdsmarkedet pga. social arv, dårlige evner eller fysisk og psykisk sårbarhed.

4.2 Gruppering af individuelle indkomstforløb

Det foregående afsnit så nærmere på udviklingen i indkomststrukturen for målgruppen som helhed i løbet af den periode, hvor gruppen er i alderen 35-61 år. Dette kan imidlertid dække over en betydelig variation i individuelle indkomstforløb og i dette afsnit ønsker vi derfor at afdække i hvilken udstrækning, nogle individer gennemlever livsforløb, der fra et indkomstmæssigt perspektiv minder om hinanden. En eventuel systematik i livsforløb kan afsløres ved anvendelse af sekvensanalyse, jf. eksempelvis Gabandinho et. al (2011a). Sekvensanalyse stammer oprindeligt fra bio-videnskabelige analyser af DNA, men har de seneste 10 år fundet anvendelse i de sociale videnskaber. Analyseformen er således typisk anvendt til belysning af samfundsvidenskabelige forhold gennem et livsforløb såsom udviklingen i uddannelses- og arbejdsmarkedstilknytning, jf. Gabandinho et. al (2011c), udviklingen i familiemønstre, jf. Fulda (2016) eller udviklingen i bosættelsesmønstre gennem livet, jf. Zwierny et al. (2017). Så vidt vides har tilgangen ikke været anvendt på indkomstforløb beskrevet ved skift mellem indkomstdeciler.

Dette afsnit afdækker systematikken i de individuelle indkomstforløb for målgruppen ved uafhængigt af baggrundskarakteristika at klassificere forløbene i grupper, der hver især repræsenterer bestemte typer af forløb. Efterfølgende vil der blive set nærmere på i hvilket omfang udvalgte baggrundskarakteristika signifikant kan redegøre for, at et indkomstforløb tilordnes én gruppe frem for en anden.

Det viser sig, at målgruppens indkomstforløb overordnet kan opdeles i fire forskellige typer af forløb. Den første gruppe er karakteriseret ved at omfatte personer, der typisk opholder sig i de lavere indkomstdeciler gennem perioden, hvor de pågældende har en relativt høj ledighed, en løs tilknytning til arbejdsmarkedet og en relativt høj tilbøjelighed til gradvist at overgå til førtidspension. Den anden og tredje gruppe omfatter personer, der har indkomstforløb, der typisk er karakteriseret af indkomster lidt over og under medianen. Begge grupper har en stabil tilknytning til arbejdsmarkedet, men specielt den tredje gruppe er typisk tilknyttet arbejdsmarkedet en stor del af den betragtede periode. Forskellen mellem de to grupper skal primært findes i gruppernes uddannelsesstruktur, hvor den tredje gruppe er kendetegnet ved en lavere andel med grundskoleuddannelse og en større andel med en videregående uddannelse end den anden gruppe. Personer kategoriseret i den fjerde gruppe har indkomstforløb præget af ophold i de højeste indkomstdeciler. Herudover svarer tilknytningsgraden til arbejdsmarkedet stort set til, hvad var gældende i den tredje gruppe, men uddannelsesniveaet er væsentligt højere. Lavindkomstgruppen omfatter den største andel af personer, der på et tidspunkt i løbet af perioden er bosat i et udsat område. Den anden og tredje gruppe omfatter de næsthøjeste andele. Der er stort set ingen personer i gruppe 4, der på et tidspunkt i løbet af perioden bor i et udsat område. Blandt personer, der på et tidspunkt bor i et udsat område, er det individerne i gruppe 1, der har de længste opholdstider i området. Kvinder udgør flertallet af individerne i de to laveste indkomstgrupper, mens mænd dominerer de to højeste indkomstgrupper. Ikke-vestlige indvandrere er stærkest repræsenterede i gruppe 1, mens personer af dansk oprindelse er stærkest repræsenterede i grupperne omfattende forløb med middelhøje og høje indkomster.

Grupperingen af indkomstforløb er overordnet baseret på hierarkisk clustering iht. tilgangen beskrevet i Gabandinho et. al (2011a). Relativt til en traditionel clusteranalyse, er det specielle dog her, at de emner, der skal grupperes, tager form af individuelle sammenhængende sekvenser frem for individuelle enkeltstående observationer. Fremgangsmåden tillader dermed, at der inddrages en sekventiel dimension i data, der typisk vil udgøres af alder eller tid. Når man her ønsker at gruppere individer baseret på indkomstforløb frem for et mere sammenfattende mål som eksempelvis samlet livsindkomst,

er det motiveret af, at personer med den samme livsindkomst sagtens kan have opbygget deres indkomst på forskellig vis betinget af udviklingen i deres socioøkonomiske forhold gennem livet. Omvendt kan individer også have ensartede indkomstforløb i en stor del af livet for så at divergere og dermed ikke opnå samme livsindkomst. Når man betragter forskelle i livsforløb frem for forskelle i et statisk eller sammenfattende indkomstbegreb, er det muligt at identificere, hvilke socioøkonomiske forhold, der er knyttet til forskelle i individuelle livsforløb eller typer af individuelle forløb. Endvidere kan man også analysere, hvorledes indkomstforløb i en tidlig fase af livet er knyttet til indkomstforløb i en senere fase af livet og derigennem få afdækket en detaljeret grad af persistens. Her vil der primært blive fokuseret på at afdække, hvorvidt der findes typiske kategorier af indkomstforløb og i hvilket omfang personkarakteristika kan knyttes til en sådan kategorisering og evt. også til individuelle forskelle.

Enhver clustering-tilgang tager afsæt i et afstandsmål og kravet til det her anvendte er specielt idet, det simultant skal kunne håndtere afstande for samtlige tilstande mellem livsforløb. Det her anvendte afstandsmål benævnes *optimal matching (OM) edit distance*, hvor afstanden mellem to forløb er defineret som den samlede omkostning, der er forbundet med at transformere det ene forløb til det andet gennem udskiftning af de enkelte tilstande. Omkostningen for at skifte én tilstand ud med en anden er baseret på den gennemsnitlige sandsynlighed for, at man overgår fra en indkomstdecil til en anden i løbet af livet. I det aktuelle tilfælde er omkostningen baseret på en 10x10-matrix, der angiver tilbøjeligheden for overgang mellem ethvert par af indkomstdeciler. Som anført i appendiks C, så kan afstanden mellem et cluster og et punkt opgøres på forskellige måder. I det aktuelle tilfælde er forbindelsen opgjort ved *average-linkage*²², som også anvendes i CNN-algoritmen.

Antallet af clustre eller grupper er valgt ud fra en række mål for clusterkvalitet²³. Anbefalingerne fra målene er ikke entydige, men indikerer typisk, at 2, 3, 4 eller 9-10 clustre er et hensigtsmæssigt valg. Opdeling på mange grupper, vil typisk lede til større ensartethed inden for de enkelte grupper og dermed alt andet lige større forskel på tværs af grupper. Såfremt der er en signifikant sammenhæng mellem baggrundskarakteristika og indkomststrukturen, kan der være tendens til, at forskellen i indkomstforløb mellem grupper afspejler en tilsvarende forskel i baggrundskarakteristika. I så fald tydeliggøres det, at individer med bestemte karakteristika er mere tilbøjelige til at have et indkomstforløb i én gruppe frem for en anden. Ulempen ved at opdele indkomstforløbene i mange grupper kan dog være, at enkelte grupper bliver så små, at det ikke bliver meningsfuldt at drage inferens. Herudover kan det være hensigtsmæssigt, at opdelingen er i overensstemmelse med en fortolkningsmæssigt meningsfuld karakterisering af grupperingen. Med afsæt i disse overvejelser og test for clusterkvalitet vælges en opdeling af indkomstforløbene på fire grupper.

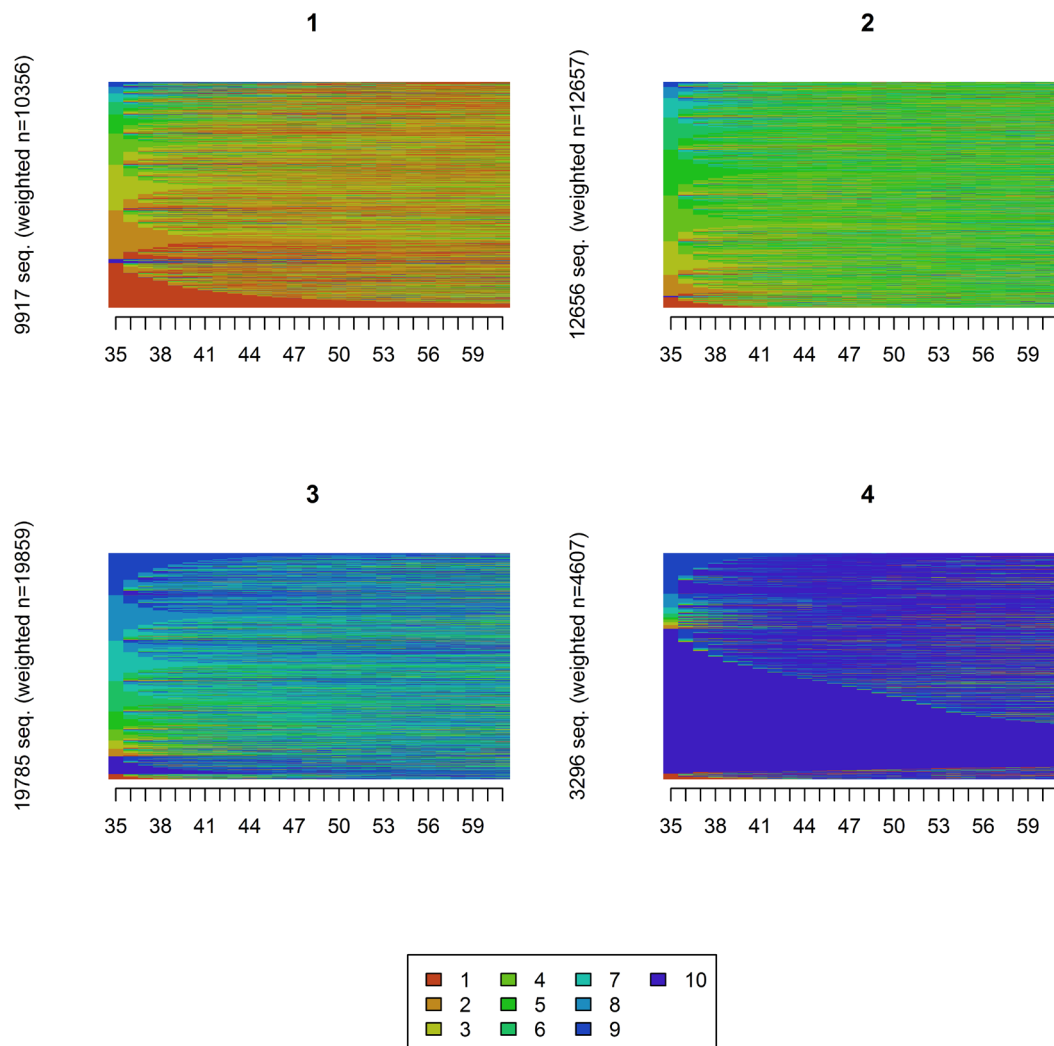
De individuelle indkomstforløb inden for hver gruppe er afbilledet i Figur 4.11. I Figur D.1 og Figur D.2 kan ses, hvorledes de fire grupper fremkommer ved sekventielle binære split af indkomstforløbene for hele målgruppen. Kriterierne anvendt til udvælgelse af målgruppen resulterer i, at analysen kan foretages på 47.479 personer, der repræsenterer 45.654 unikke indkomstforløb. Omkring 1.800 forløb kan altså genfindes en eller flere gange i data. Sandsynligheden for forekomsten af gentagne forløb mindskes betydeligt med antallet af

²² Muligheden for at benytte *Ward-linkage* er også afsøgt. Den resulterende gruppering minder meget om resultatet opnået ved *average-linkage*, hvorfor sidstnævnte er foretrukket i kraft af sin hurtigere afviklingstid.

²³ Jf. <http://mephisto.unige.ch/weightedcluster/>

perioder (her 27) og med antallet af mulige tilstande inden for hver periode (her 10). Antallet af gentagelser er både absolut og forholdsmæssigt størst i gruppe 1 og i gruppe 4, der overordnet repræsenterer individer med tilknytning til hhv. de laveste og de højeste indkomstdeciler. Forløb i gruppe 2 og gruppe 3 er kendetegnet ved, at befolkningen primært er knyttet til de midterste indkomstdeciler, hvor individer i gruppe 2 er positioneret lidt svagere i indkomstfordelingen relativt til personer i gruppe 3.

Figur 4.11. Individuelle decilindkomstforløb i perioden 1991-2017 fordelt på fire grupper.



Anm.: Den laveste indkomstdecil repræsenteres ved "1". I hver af de fire grupper betragtes udviklingen i individuelle livsforløb (2. aksens) over perioden 1991-2017, her svarende til, at 35-årige i 1991 følges frem til de er 61 år i 2017 (1. aksens). Ud for hver gruppe er anført såvel antallet af unikke som vægtede forløb i gruppen. Forskellen repræsenterer antallet af gentagelser i gruppen.

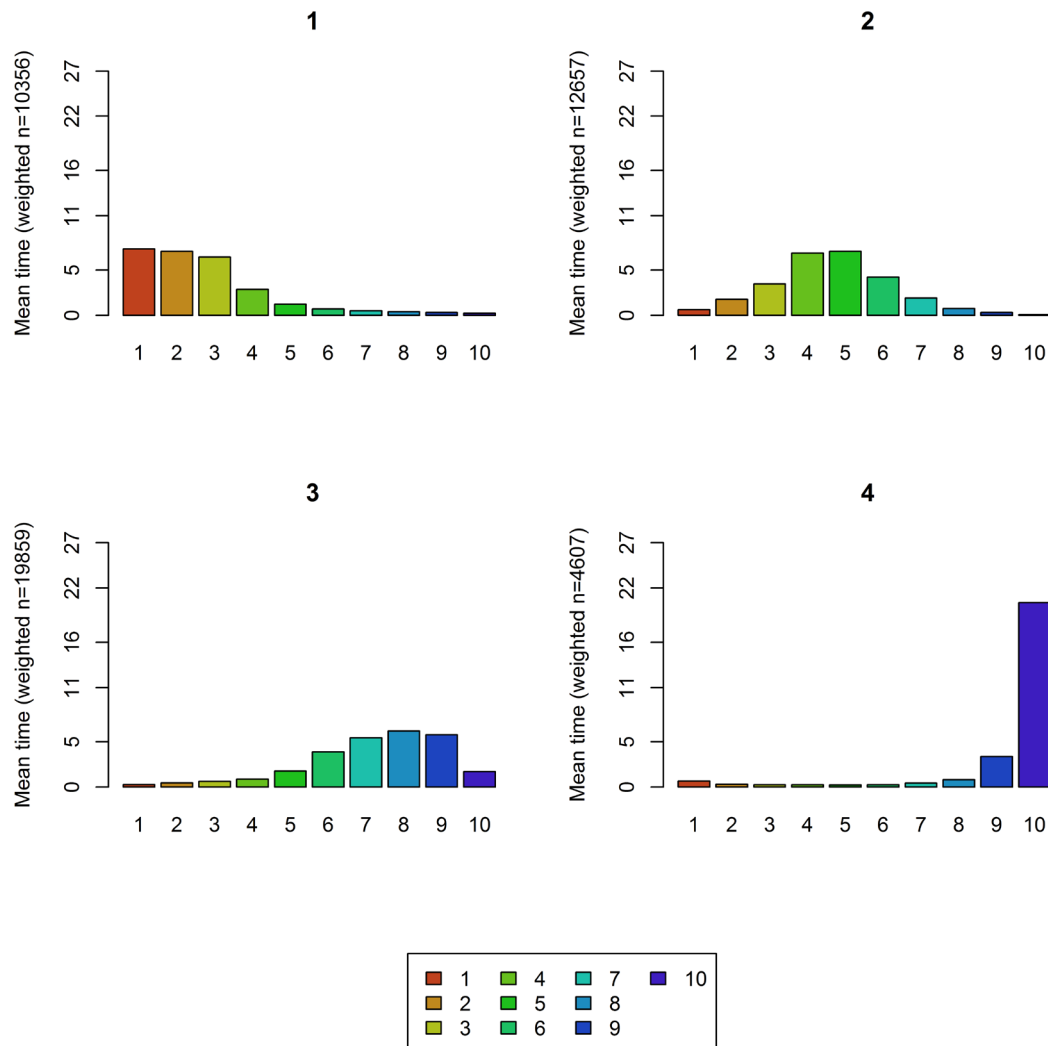
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Gruppe 1 og 2 er nogenlunde ensartede i størrelse og omfatter mellem 10.300 og 12.700 forløb, mens antallet af forløb i gruppe 3 er omkring dobbelt så stort som antallet af forløb i gruppe 1. Gruppen kendetegnet ved ophold i de øverste indkomstdeciler er omkring halvt så stor som gruppe 1.

Forskellen mellem de fire grupper kan alternativt repræsenteres ved en opgørelse af det antal år, som individerne inden for en gruppe i gennemsnit har tilbragt i hver decil, jf. Figur 4.12. Individerne i gruppe 1 har gennemsnitligt tilbragt mere end 2/3 af perioden i første, anden og

tredje indkomstdecil, heraf 7,3 år i det første, 7,1 år i den anden og 6,4 år i den tredje. Individene i den 4. gruppe har gennemsnitligt tilbragt 23,8 år i det 9. og 10. indkomstdecil, heraf 20,4 år i den højeste decil. Gruppe 2 og 3 er kendetegnet ved en større variation i deciltilknytning. Opgørelsen indikerer, hvor personerne i hver af de fire grupper typisk opholder sig, men afslører ikke i hvilken udstrækning tilknytningen til de pågældende deciler er stabil over livet for det enkelte individ.

Figur 4.12. Gennemsnitligt antal år i hver decil i perioden 1991-2017 fordelt på fire grupper.

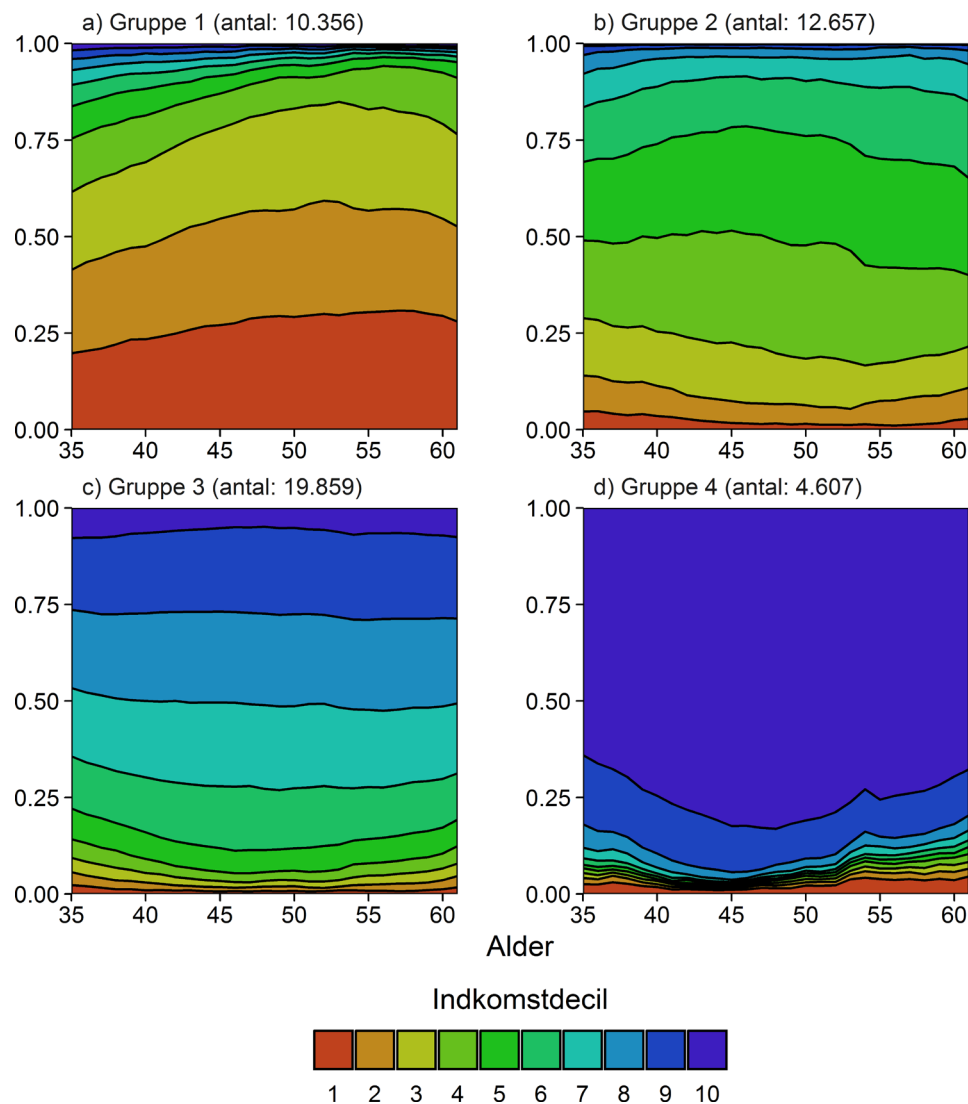


Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Gruppekarakteristika

Med afsæt i opdelingen af indkomstforløb ovenfor ses her nærmere på hvilke baggrundkarakteristika, der kendetegner individerne i hver af de fire grupper. Betydelige forskelle i karakteristika mellem grupperne kan indikere, at de pågældende karakteristika kan være med til at forklare, hvorfor et individ tilordnes én gruppe frem for en anden.

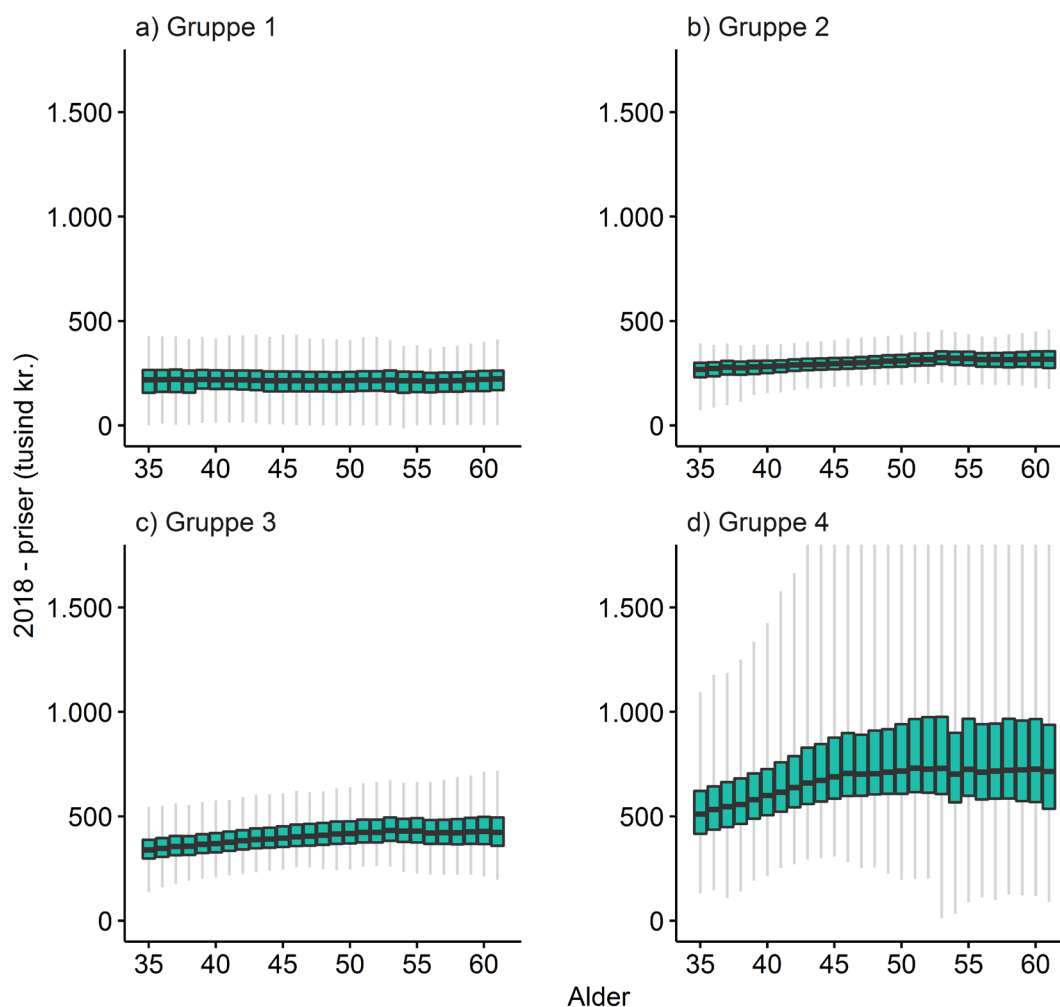
Figur 4.13 sammenholder indledningsvist udviklingen i målgruppens sammensætning på indkomstdeciler over tid i de fire grupper. Figuren er en alternativ fremstilling af indholdet i Figur 4.11, idet fordelingen på indkomstdeciler på tværs af individer er beregnet på hvert tidspunkt i livsforløbet omfattende aldersgrupperne 35-61 år.

Figur 4.13. Gruppebetinget udvikling i målgruppens fordeling på indkomstdeciler.

Kilde: *Egne beregninger på baggrund af registerdata fra Danmarks Statistik.*

Ud over at bekræfte de foreløbige konklusioner om den kvalitative forskel i indkomstforløb mellem de fire grupper, tydeliggør Figur 4.13 hvor stor en andel af befolkningen i de enkelte grupper, der befinder sig i de respektive deciler. Mindst 75 procent af befolkningen i gruppe 1 befinder sig til hver en tid i de fire laveste indkomstdeciler. I gruppe 4 vil mindst 80 procent af befolkningen og typisk flere befinde sig i de fire øverste indkomstdeciler. Den første og den fjerde gruppe er dermed karakteriseret af nogenlunde den samme indkomststruktur, men i hver sin ende af indkomstfordelingen. Noget tilsvarende kan sluttes om den indbyrdes forskel mellem personerne i gruppe 2 og 3, der primært omfatter forløb, der ligger hhv. under og over medianindkomsten. Gruppe 4 udmærker sig ved, at befolkningen er relativt koncentreret i de to øverste indkomstdeciler, hvilket selvfølgelig er i overensstemmelse med den relativt høje gennemsnitlige opholdstid i netop disse deciler.

Figur 4.14. Gruppebetinget udvikling i målgruppens medianindkomst fra 35-61 år, 2018-kr.



Anm.: Den personlige indkomst er omregnet til 2018-kr. ved anvendelse af forbrugerprisindekset, tabel PRIS112 i Statistikbanken.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

En alternativ karakterisering af gruppernes indkomststruktur kan opnås ved at betragte udviklingen i medianindkomsten over tid inden for hver gruppe i Figur 4.14. Medianindkomsten i gruppe 1 er relativt stabil over tid, mens der for de resterende grupper indtræder en gradvis stigning over alder, en tendens, der er mest udpræget for gruppen karakteriseret af ophold i de øverste indkomstdeciler. For gruppe 2, 3 og 4 indtræder en stabilisering af medianindkomsten først i 50'erne. Samme kvalitative billede kan ses, hvis man alene betragter udviklingen i medianen for erhvervsindkomsten. Det kan dermed ikke umiddelbart afklares om tendensen stammer fra en aldersbetinget ændring i produktiviteten eller deltagelsestilbøjeligheden. Som forventet øges det absolutte niveau for medianindkomsten fra gruppe 1 til 4, mens der for gruppe 2 til 4 ses en markant forøgelse af variation i indkomstniveauet. Variationen i gruppe 4 må formodes delvist at skulle tilskrives, at antallet af personer i denne gruppe er lavere end i de øvrige.

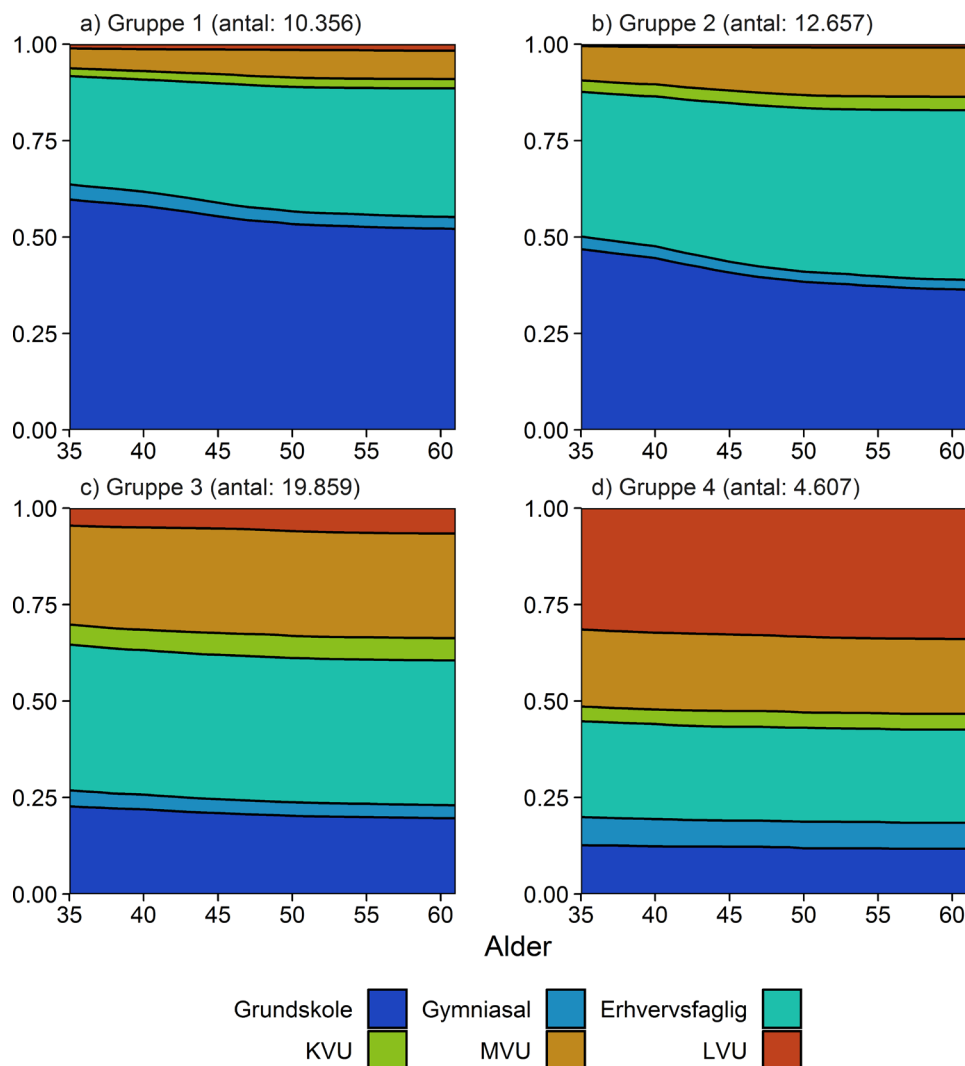
Forskellen i indkomststrukturen mellem grupperne gengives i nogen grad, når forskellen i gruppernes uddannelsessammensætning betragtes, jf. Figur 4.15. Befolkningen i gruppe 4 har i højere udstrækning end befolkningen i de øvrige grupper opnået en videregående uddannelse og har i mindre grad end de øvrige en grundskoleuddannelse som højst fuldførte

uddannelse. Mellem 50 og 60 procent af befolkningen i gruppe 1 har på et givet tidspunkt en grundskoleuddannelse som højst fuldførte, hvorefter en erhvervsfaglig uddannelse er den næsthyppest forekommende. Uddannelsesstrukturen i gruppe 1 og gruppe 4 afspejler derfor i vidt omfang indkomststrukturen i grupperne. Uddannelsessammensætningen i gruppe 2, der ovenfor karakteriseredes som en mellemindkomstgruppe under medianen, har en uddannelsesstruktur, der minder særdeles meget om strukturen i gruppe 1. Den primære forskel synes at være, at andelen af erhvervsfagligt uddannede og personer med en mellemlang videregående uddannelse i gruppe 2 er lidt større end i gruppe 1. Omvendt er der en større andel af befolkningen i gruppe 1, der har en grundskoleuddannelse. At uddannelsesstrukturen i gruppe 1 og 2 er relativt ensartet, mens indkomststrukturen synes noget mere forskellig kan indikere, at uddannelsesniveaueet ikke er det karakteristiskum, der forklarer, hvorfor et individ tilordnes gruppe 1 frem for gruppe 2. Denne hypotese betragtes senere fra et statistisk perspektiv.

Bortset fra en lavere andel med en grundskoleuddannelse og en højere andel med en videregående uddannelse minder uddannelsesstrukturen i gruppe 3 om strukturen i gruppe 2. Dette er i overensstemmelse med, at indkomststrukturen i begge tilfælde er præget af deciler umiddelbart over og under medianen. Således vil det ikke være overraskende, hvis uddannelse spiller en rolle i at forklare forskelle i gruppetilhørsforholdet mellem gruppe 2 og 3.

Som tidligere nævnt er valget af målgruppe delvist funderet i, at den pågældende generation fra forløbets start har en alder, hvor den enkeltes uddannelsesforløb typisk er afsluttet. At dette ikke helt er tilfældet afsløres i Figur 4.15 ved, at andelen med en grundskoleuddannelse i specielt gruppe 1, 2 og 3 er aftagende gennem livet på bekostning af en større andel med en videregående uddannelse, hvilket kan indikere, at man i de pågældende grupper afslutter sit uddannelsesforløb senere eller er tilbøjelig til at føje en overbygning til en tidligere opnået uddannelse.

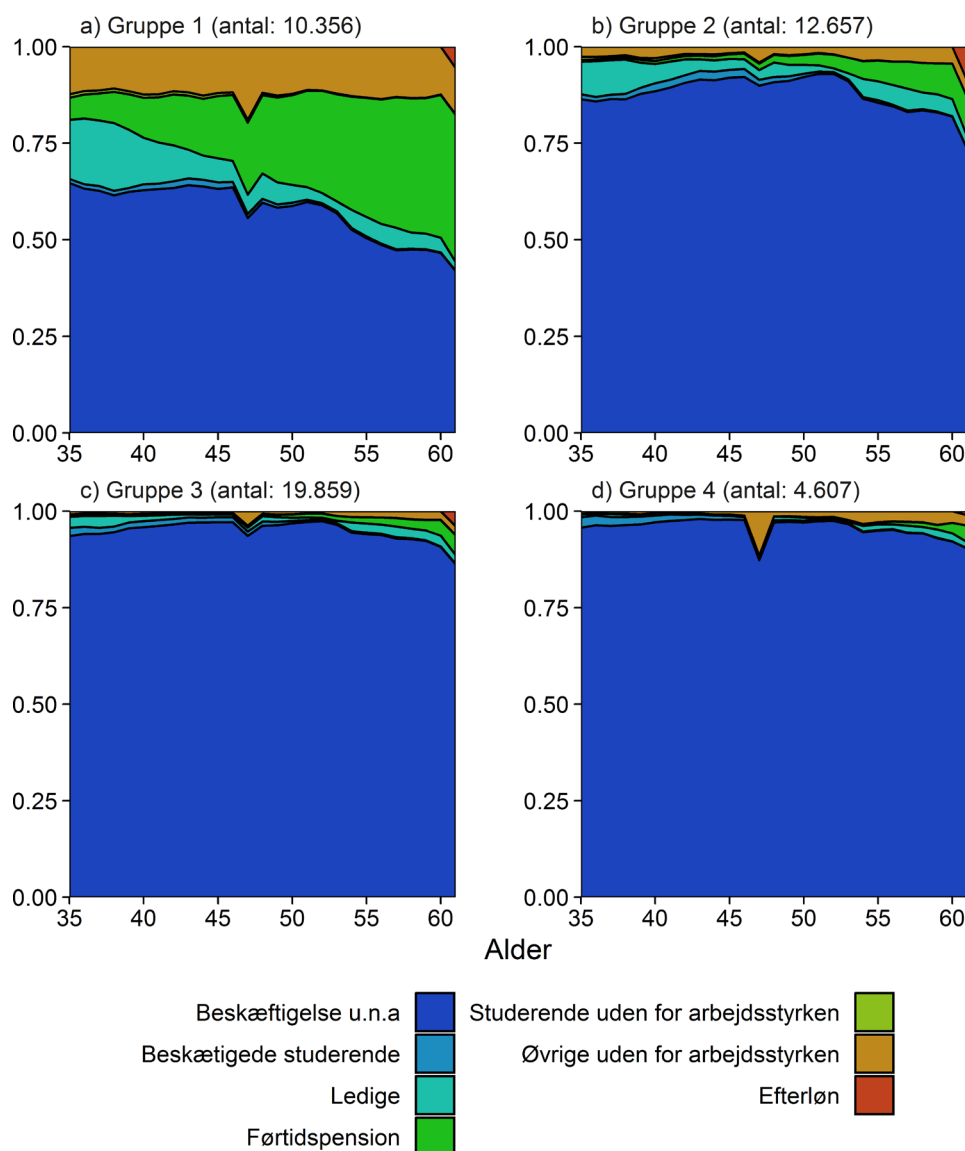
Figur 4.15. Gruppebetinget udvikling i målgruppens fordeling på højest fuldførte uddannelse.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Betragter man i stedet gruppernes sammensætning på arbejdsmarkedsstatus, afsløres en langt større grad af forskellighed mellem gruppe 1 og gruppe 2 end uddannelsesstrukturen giver indtryk af. Endvidere synes der at være en langt større grad af ensartethed i arbejdsmarkedsstrukturen mellem gruppe 2, 3 og 4 end uddannelsessammensætningen umiddelbart tilsiger. Specielt er der en stor grad af overensstemmelse i den socioøkonomiske struktur mellem gruppe 3 og 4. Beskæftigelsesgraden er dog stigende fra gruppe 2 til gruppe 4 og udgør dermed en del af forklaringen på den indbyrdes forskel i indkomststrukturen mellem de tre grupper. Udover en forskel i beskæftigelsesomfanget må forskellen i gruppernes uddannelsessammensætning forventes at afspejle sig i deres produktivitet. Eftersom produktiviteten alt andet lige er stigende i uddannelsesniveaue, understøtter uddannelsesstrukturen dermed sammen med beskæftigelsesomfanget, at personerne i grupperne præget af længerevarende uddannelser typisk befinder sig i de øverste indkomstdeciler.

Figur 4.16. Gruppebetinget udvikling i målgruppens fordeling på socioøkonomisk status.



Anm.: Der skal bortses fra den midlertidige stigning i "Øvrige uden for arbejdsstyrken" omkring 47-årsalderen, der skal tilskrives en uregelmæssighed i data.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

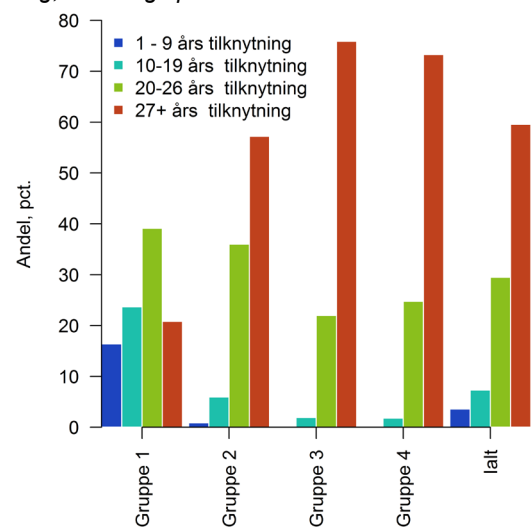
Den socioøkonomiske struktur i gruppe 1 er kendetegnet ved en relativ lav grad af tilknytning til arbejdsmarkedet sammenholdt med de øvrige grupper. Herudover er andelen knyttet til arbejdsmarkedet faldende over tid på bekostning af en stort set lineær stigning i tilknytningen til førtidspension. Personer knyttet til denne ordning udgør således mod periodens slutning 38 procent af befolkningen i gruppen. I periodens begyndelse er andelen af ledige i gruppen aftagende og en del heraf må formodes at overgå til førtidspension. Gruppe 1 er desuden kendetegnet ved at omfatte en relativt stor og stabil andel af øvrige uden for arbejdsstyrken. Datagrundlaget tillader desværre ikke en yderligere nuancering af denne kategori, men den omfatter bl.a. kontanthjælpsmodtagere uden dagpengeret og sygedagpengemodtagere. Overgang til førtidspension vil også hyppigt ske fra kontanthjælp og dette kan være i overensstemmelse med den relative stabile udvikling i andelen af øvrige uden for arbejdsstyrken, såfremt der er en samtidig tilgang til "Øvrige uden for arbejdsstyrken", der

udligner faldet i antallet af kontanthjælpsmodtagere. En sådan tilgang kan evt. udgøres af personer, der overgår fra ledighed til kontanthjælp.

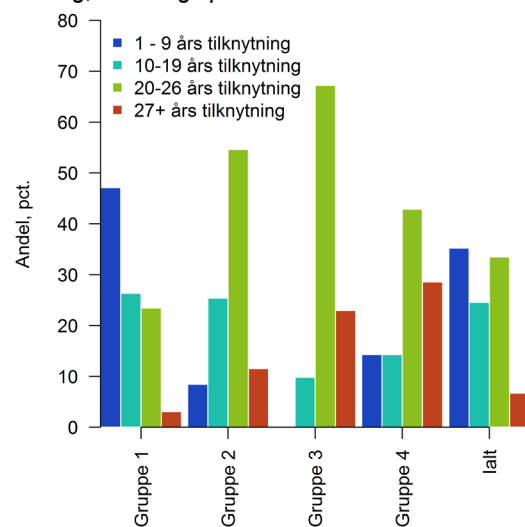
Tendensen til overgang til førtidspension ved stigende alder ses også i gruppe 2, 3 og 4, men i langt mindre udstrækning end i gruppe 1. Desuden starter overgangen til førtidspension først omkring 45-50-årsalderen i de tre grupper. Tilbøjeligheden er mest udbredt i gruppe 2, der også har en større andel af ledige forud for efterlønsalderen end gruppe 3 og 4.

Figur 4.17. Arbejdsmarkedsanciennitet for 35-61-årige fordelt på indkomstgrupper og graden af arbejdsmarkedstilknytning i alderen 30-34 år, fordeling og antal.

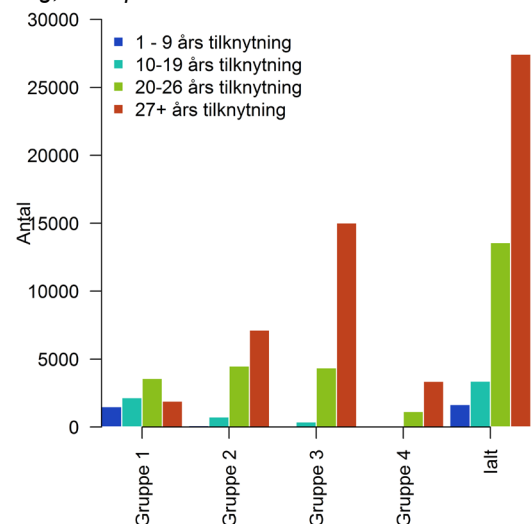
a) Mindst ét år på arbejdsmarkedet som 30-34-årig, fordeling i procent



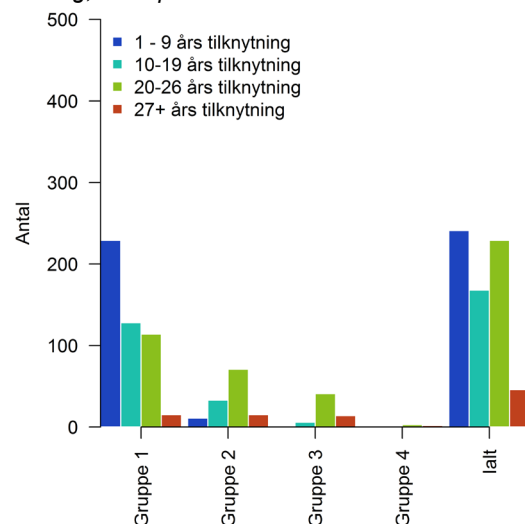
b) Ingen tilknytning til arbejdsmarkedet som 30-34-årig, fordeling i procent



c) Mindst ét år på arbejdsmarkedet som 30-34-årig, antal personer



d) Ingen tilknytning til arbejdsmarkedet som 30-34-årig, antal personer



Anm.: Der ses udelukkende på ancienniteten for personer, der har haft mindst et år på arbejdsmarkedet i alderen 35-61 år. Anciennitet er ikke nødvendigvis udtryk for længden af et sammenhængende forløb på arbejdsmarkedet.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

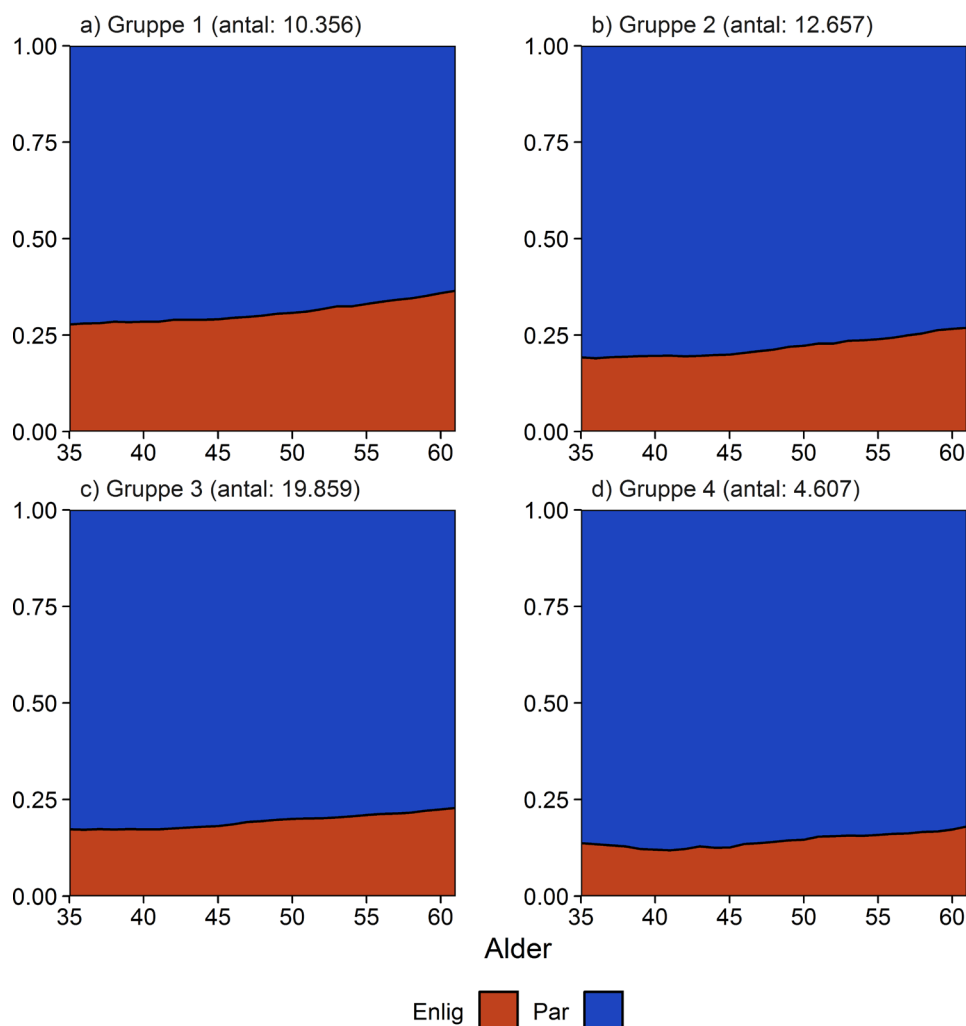
Figur 4.17 illustrerer arbejdsmarkedsancienniteten i hver af de fire grupper for personer, der i alderen 35-61 år har haft mindst et år på arbejdsmarkedet. Opgørelsen nuanceres på hvorvidt individerne har været tilknyttet arbejdsmarkedet i alderen 30-34 år og omfatter i alt 46.735

personer svarende til antallet i Figur 4.7. Samlet set har 46.051 personer været tilknyttet arbejdsmarkedet i alderen 30-34 år, hvorfor blot 684 personer ikke har været det, jf. Figur 4.17c og Figur 4.17d. For personer med arbejdsmarkedstilknytning i alderen 30-34 år adskiller anciennitetsstrukturen i de fire grupper sig ved, at andelen med konstant arbejdsmarkedstilknytning dominerer i gruppe 2, 3 og 4, mens personerne i gruppe 1 typisk har mellem 10 og 26 års tilknytning, jf. Figur 4.17b. Den relativt lavere anciennitet i gruppe 1 er i overensstemmelse med den socioøkonomiske struktur i denne gruppe sammenholdt med de andre grupper. Andelen med konstant arbejdsmarkedstilknytning er størst i den gruppe, der er kendetegnet af indkomster over medianen. At andelen med den højeste anciennitet ikke er højest i gruppe 4, skal givetvist tilskrives, at denne gruppe omfatter den største andel med en længerevarende uddannelse, som dermed er længst tid om at optjene en høj anciennitet på trods af indregningen af anciennitet optjent som beskæftigede studerende.

Har man ikke været tilknyttet arbejdsmarkedet i alderen 30-34 år, er det kun et fåtal, der i den efterfølgende periode har optjent fuld anciennitet. Andelen med en erhvervs erfaring på 20-26 år er her højest for gruppe 3 (når der bortses fra gruppe 4, hvor der er meget få personer uden tidligere tilknytning). I gruppe 1 er den typiske anciennitet på 1-9 år, mens den i de øvrige grupper er 20-26 år.

Den relativt dårligere anciennitet for personer, der ikke var tilknyttet arbejdsmarkedet forud for den betragtede periode, illustrerer betydningen af et tidligt fodfæste på arbejdsmarkedet for den resterende del af livsforløbet. Overordnet giver forskellen i anciennitetsfordelingen på tværs af grupper anledning til at tro, at varigheden af arbejdsmarkedstilknytningen spiller en rolle for klassificeringen af indkomstforløbene og formentlig også for den indbyrdes forskel mellem forløb.

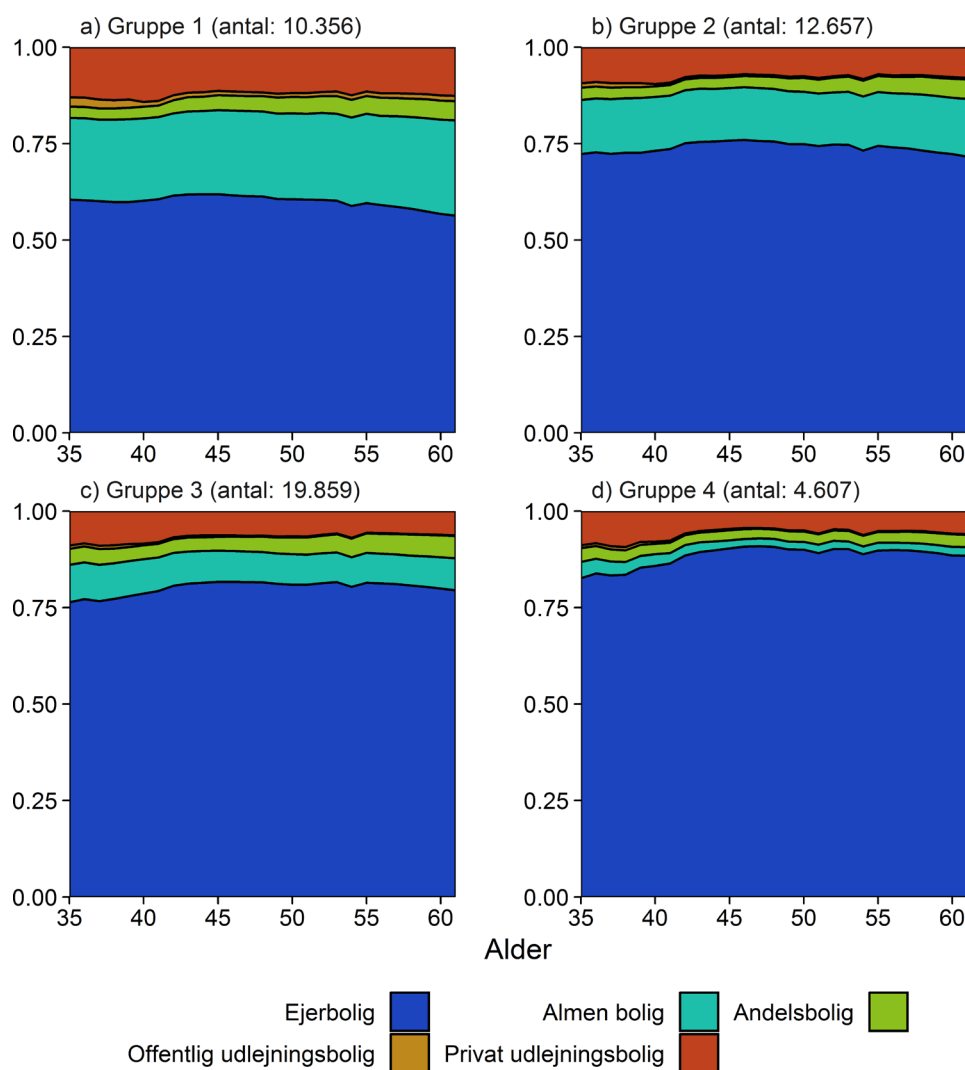
Udviklingen i familiestrukturen inden for grupperne peger i retning af, at andelen af enlige er lidt højere i gruppe 1 og 2 i forhold til de to øvrige grupper, jf. Figur 4.18. At andelen af enlige dominerer i grupperne karakteriseret ved ophold i de lavere indkomstdeciler er i overensstemmelse med, at indkomststrukturen for personer, der ikke har levet i et parforhold i alderen 30-34 år, og dermed typisk heller ikke i alderen 35-61 år, i højere grad er kendetegnet ved ophold i de nederste indkomstdeciler, jf. Figur 4.4. Overordnet synes andelen af enlige gradvist at aftage med øget grad af tilknytning til de øverste indkomstdeciler og der er derfor muligvis anledning til at tro, at forskellen mellem grupperne og indkomstforløbene er knyttet, om end ikke direkte, så indirekte til familietypen.

Figur 4.18. Gruppebetinget udvikling i målgruppens fordeling på familietype.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Betragtes udviklingen i gruppernes fordeling på boligtype kan det ses, at andelen med en ejerbolig er højere i gruppe 3 og 4 end i gruppe 1 og 2, jf. Figur 4.19. Herudover synes fordelingen på boligtype marginalt mere stabil i grupperne karakteriseret ved ophold i de lave indkomstdeciler og decilerne umiddelbart over og under medianen. Dette kan eksempelvis være udtryk for, at der i de pågældende grupper i gennemsnit er en stabil til- og fraflytning inden for den samme boligtype eller, at der for de individuelle forløb er en høj grad af stabilitet i boligtypetilhørsforholdet over livet. Bosætning i almene boliger er relativt mere udpræget i lavindkomstgrupperne, mens andelen af personer i en almen bolig udgør en meget lille andel af gruppe 3. Figur 4.19 synes at motivere, at boligtype kan forklare tilhørsforholdet til højindkomstgrupperne relativt til gruppe 1, mens det fra en visuel betragtning er svært at slutte om strukturen i gruppe 2 er signifikant anderledes end i gruppe 1.

Figur 4.19. Gruppebetinget udvikling i målgruppens fordeling på boligtype.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

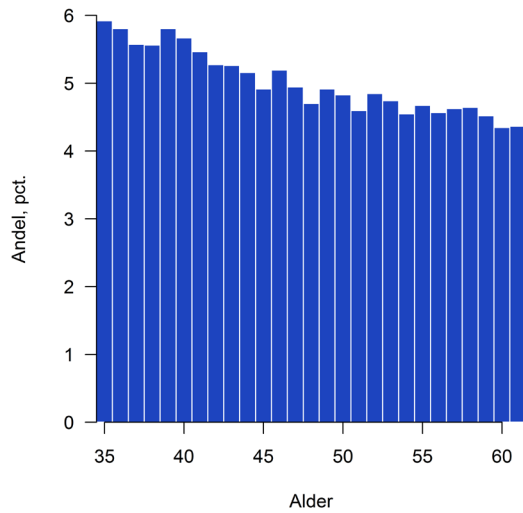
Andelen af målgruppen med erfaring fra bosættelse i et udsat område, vil være højest i gruppe 1 og derefter aftagende både med alder og med stigende gruppenummer, jf. Figur 4.20. Det er således i gruppen kendetegnet af de laveste indkomster, at man på ethvert tidspunkt i løbet af livsforløbet finder den største andel, der er bosat i et udsat område. At andelen, der bor i et udsat område er faldende over tid, kan både tilskrives, at der er en negativ nettotilflytning til de udsatte områder for den betragtede generation og at områderne skifter karakter gennem perioden. For varigt udsatte boligområder indikerede tværnsnitsdata en svag negativ nettotilflytning for personer ældre end 35 år, jf. Figur 3.5, hvorfor dette kunne give anledning til at tro, at en tilsvarende tendens også er gældende for den betragtede målgruppe.

Eftersom langt den overvejende del af personerne bosat i et udsat område også bor i en almen bolig, jf. Tabel 2.1, er det i forlængelse af ovenstående i overensstemmelse hermed, at andelen med bopæls erfaring i et udsat område og andelen med erfaring fra en almen bolig, er rangeret på samme vis mellem grupperne. At der er forskel på områdeerfaringen mellem grupperne, er i overensstemmelse med, at indkomststrukturen over livet varierer alt efter, om man har været bosat i et udsat område i den unge aldersgruppe, jf. Figur 4.9, uden at dette

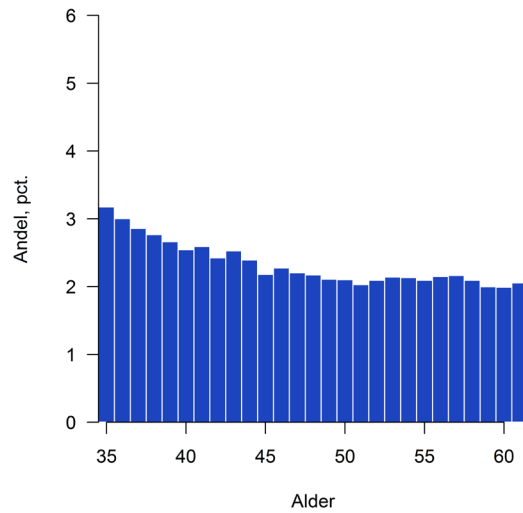
dog nødvendigvis skal tilskrives en effekt fra området i sig selv. Om der er tale om et signifikant forhold mellem indkomstforløbene og områdestatus betragtes senere.

Figur 4.20. Gruppebetinget udvikling i målgruppens andel med bopæl i et udsat område.

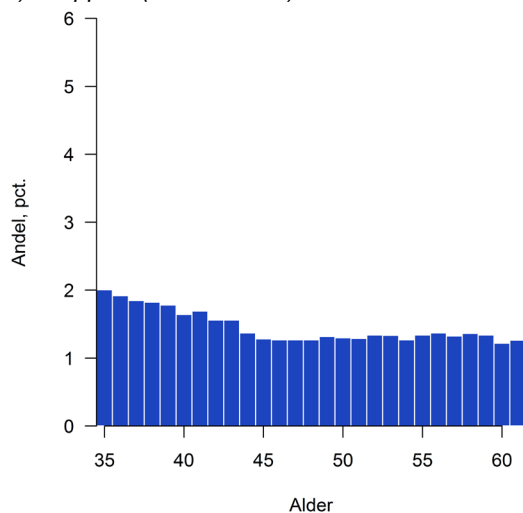
a) Gruppe 1 (antal: 10.356)



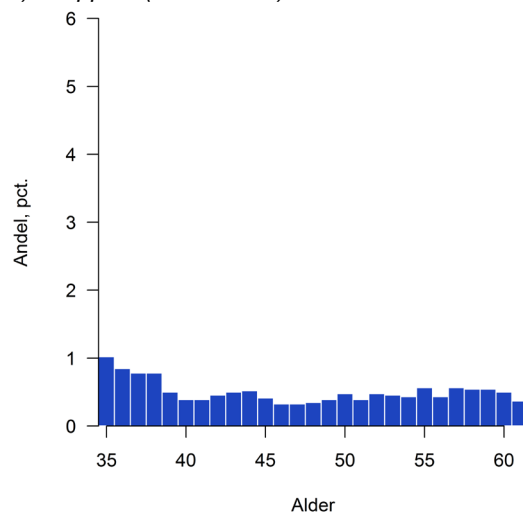
b) Gruppe 2 (antal: 12.657)



c) Gruppe 3 (antal: 19.859)



d) Gruppe 4 (antal: 4.607)

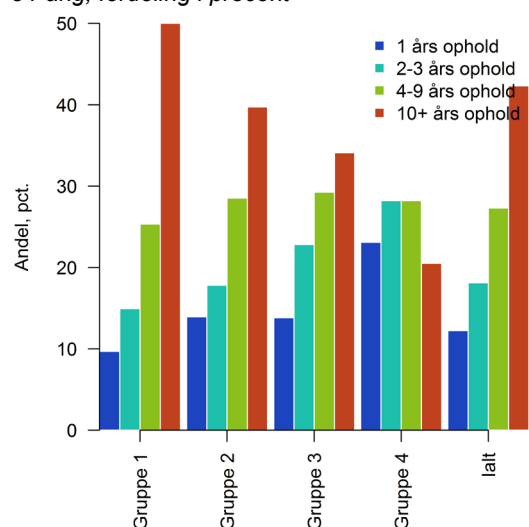


Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

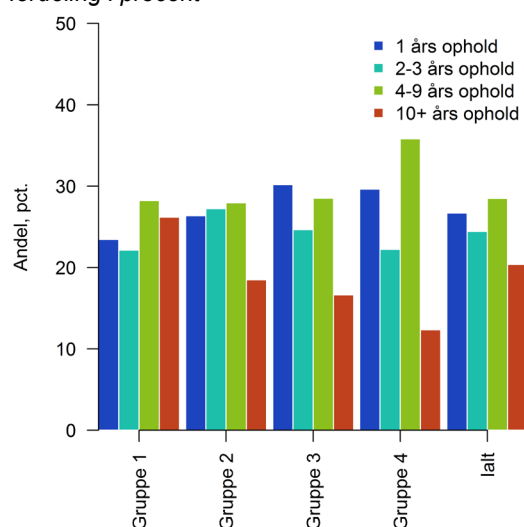
Tidligere betragtedes indkomststrukturen på tværs af varigheden af opholdet i et udsat område for den del af målgruppen, der har boet i et udsat område i alderen 35-61 år, jf. Figur 4.10. I Figur 4.21 nedenfor betragtes for de samme 4.027 individer variationen i opholdslængden inden for hver gruppe og for grupperne som helhed. Endvidere er opgørelsen nuanceret på, hvorvidt man i alderen 30-34 år også har været bosat i et udsat område.

Figur 4.21. Opholdstid i et udsat område for 35-61-årige fordelt på indkomstgrupper og bopæls erfaring i alderen 30-34 år, fordeling og antal.

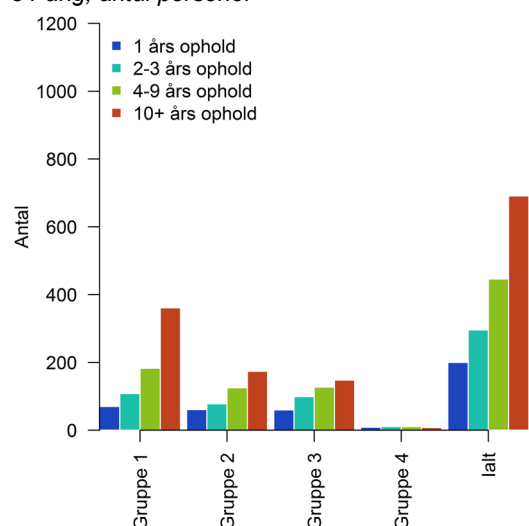
a) Mindst ét års ophold i et udsat område som 30-34-årig, fordeling i procent



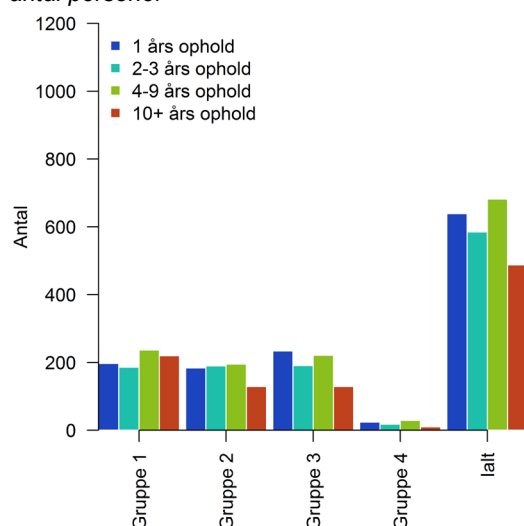
b) Intet ophold i et udsat område som 30-34-årig, fordeling i procent



c) Mindst ét års ophold i et udsat område som 30-34-årig, antal personer



d) Intet ophold i et udsat område som 30-34-årig, antal personer



Anm.: Der ses udelukkende på opholdstiden for personer, der har boet mindst ét år i et udsat område i alderen 35-61 år. Opholdstiden er ikke nødvendigvis udtryk for længden af et uafbrudt bopælsforløb.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Overordnet har 1.633 personer også tidligere været bosat i et udsat område, mens 2.394 personer ikke har opholdt sig i et udsat område i alderen 30-34 år. Blandt de personer, der tidligere har været bosat i et udsat område, vil godt 40 procent have en opholdstid på 10 år eller flere i et udsat område i løbet af den periode, hvor de er 35-61 år gamle, jf. Figur 4.21a. 37 procent vil have boet i et udsat område i 4-9 år, 18 procent i 2-3 år, mens 12 procent blot vil have haft bopæl i et udsat område et enkelt år. At de mellemlange og lange opholdstider dominerer fordelingen er et mønster, der også gengives inden for de enkelte grupper, dog med undtagelse af højindkomstgruppen, hvor fordelingen også må anses for usikkert bestemt, idet gruppen med erfaring fra udsatte områder i begge perioder omfatter færre end

40 personer, jf. Figur 4.21c. Gruppen karakteriseret af ophold i de laveste indkomstdeciler omfatter den største andel med lang bosættelse, gruppe 2 den næsthøjeste andel, mens andelen er mindst i gruppe 3 (når der bortses fra gruppe 4). Faldet i andelen med den længste bosættelsesperiode mellem gruppe 2 og 3 er primært på bekostning af en stigning i andelen med 2-3 års ophold. Generelt er antallet med erfaring fra et udsat område størst i gruppe 1 og altså markant mindst i gruppe 4. Opholdstidsstrukturen synes forskellig i de fire grupper. I afsnit 4.3 vil det i en univariat analyse blive belyst, hvorvidt varigheden af opholdet i et udsat område i alderen 35-61 år kan knyttes signifikant til indbyrdes forskelle i livsforløb. Grundet risikoen for bl.a. simultanitetsbias, vil vi dog ikke statistisk søge at belyse, hvorvidt varigheden af opholdet i et udsat område i denne aldersklasse har signifikant betydning for gruppetilhørsforholdet.

Såfremt man i stedet betragter opholdstidsstrukturen for personer uden forudgående erfaring med bosættelse i et udsat område, synes den indbyrdes forskel mellem grupperne mindre systematisk. Endvidere er der også kvantitativt en større grad af ensartethed mellem fordelingerne i de fire grupper, hvor det også her skal understreges, at gruppe 4 kun omfatter ca. 80 personer, jf. Figur 4.21b. Andelen med lang opholdstid er dog højere i gruppe 1 og aftager herefter med stigende gruppenummer. Antallet af personer med erfaring fra bosættelse i et udsat område er nogenlunde det samme i de tre første grupper, jf. Figur 4.21d. For den del af målgruppen, der ikke har boet i et udsat område i alderen 30-34 år, er der på baggrund af ovenstående ikke umiddelbart anledning til at tro, at varigheden af opholdet i alderen 35-61 år spiller en afgørende rolle for forskellen mellem indkomstforløbene. Dette spørgsmål søges ligeledes besvaret i et senere afsnit gennem en univariat analyse.

I Tabel 4.1 er sammenfattet statistiske karakteristika for de fire grupper og for befolkningen som helhed. Andelen af mænd ses at være højere i gruppe 3 og 4 relativt til målgruppen som helhed, der består af 48 procent mænd. Personer med dansk oprindelse repræsenterer også en marginal større andel af grupperne med forløb i de højeste indkomstdeciler end målgruppens overordnede oprindelsessammensætning tilsiger.

Kvinder dominerer gruppe 1 og gruppe 2, der er kendetegnede ved at omfatte omkring 65 procent kvinder. Vestlige indvandrere og efterkommere er generelt relativt svagt repræsenteret i målgruppen, mens ikke-vestlige indvandrere og efterkommere er stærkest repræsenteret i gruppe 1, hvor andelen er højere end for målgruppen som helhed.

Tabel 4.1. Grubebetinget fordeling af målgruppen på køn og oprindelse.

	Antal	Køn		Oprindelse		
		Mænd	Kvinder	Dansk	Ikke-vestlig	Vestlig
Gruppe 1	5.169	0,34	0,66	0,94	0,04	0,02
Gruppe 2	13.591	0,36	0,64	0,97	0,01	0,02
Gruppe 3	14.044	0,56	0,44	0,98	0,01	0,01
Gruppe 4	13.747	0,82	0,18	0,98	0,00	0,02
I alt	46.551	0,48	0,52	0,97	0,02	0,02

Anm.: Indvandrere og efterkommere er sammenlagt og kategoriseret som enten vestlige eller ikke-vestlige.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Gruppernes karakteristika med hensyn til familietype- og arbejdsmarkedserfaring er sammenfattet i Tabel 4.2. Selv om tabellen ikke på individniveau sammenkobler erfaring fra hhv. den unge aldersgruppe og den analyserede aldersgruppe, så bekræfter tabellen, at der på gruppeniveau er en betydelig sammenhæng. Som tidligere nævnt er andelen af

målgruppen, der i mindst ét år har indgået i et parforhold, stigende i indkomst. Dette gælder for begge aldersgrupper. Således udgør andelen med parforholdserfaring i gruppe 1 81 pct. i den unge aldersgruppe og 87 pct. i den ældre aldersgruppe, men de tilsvarende andele i gruppe 4 er henholdsvis 92 pct. og 97 pct. Blandt de personer, der har indgået i et parforhold i mindst ét år, ses varigheden at udgøre en betydelig del af perioden. Den gennemsnitlige parforholdserfaring for personer i den unge aldersgruppe er 4,5 år, mens den gennemsnitlige erfaring i aldersgruppen 35-61 år er 22,7 år.

Andelen af målgruppen, der har mindst ét års tilknytning til arbejdsmarkedet er inden for hver gruppe og som helhed nogenlunde den samme i de to aldersgrupper. Dette understøtter delvist den persistens i arbejdsmarkedstilknytningen, der tidligere er påvist. Blandt personer med arbejdsmarkedserfaring, er den gennemsnitlige varighed i specielt gruppe 2-4, meget tæt på at repræsentere hele længden af perioden, dvs. 27 år, hvorfor det kan sluttes, at de personer, der har en tilknytning, har en stabil en af slagsen.

Tabel 4.2. Grubebetinget fordeling af målgruppen på parforholds- og arbejdsmarkedserfaring.

Alder	Antal	Parforholdserfaring				Arbejdsmarkedstilknytning			
		Andel, 1+ år		Gns. varighed, år		Andel, 1+ år		Gns. varighed, år	
		30-34	35-61	30-34	35-61	30-34	35-61	30-34	35-61
Gruppe 1	10.356	0,81	0,87	4,4	21,5	0,91	0,93	4,6	19,1
Gruppe 2	12.657	0,88	0,93	4,5	22,6	0,99	1,00	4,9	25,3
Gruppe 3	19.859	0,89	0,94	4,5	23,0	1,00	1,00	4,9	26,3
Gruppe 4	4.607	0,92	0,97	4,5	23,9	1,00	1,00	5,0	26,3
I alt	47.479	0,88	0,93	4,5	22,7	0,98	0,98	4,8	24,5

Anm.: Den gennemsnitlige varighed er ikke nødvendigvis udtryk for længden af et uafbrudt forløb. Varigheden beregnes udelukkende for individer med erfaring.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

I Tabel 4.3 sammenfattes centrale forhold vedrørende målgruppens bopæls erfaring fra et udsat område og beboelse i en almen bolig. Overordnet har 6 procent af målgruppen været bosat mindst et år i et udsat område i alderen 30-34 år, mens det samme gælder for 8 procent af de 35-61-årige. Personerne kategoriseret i gruppe 1 udmærker sig ved at omfatte den største andel med bopæls erfaring fra et udsat område. Blandt de personer, der har været bosat i et udsat område, er den gennemsnitlige varighed af opholdet højest i gruppe 1 og derefter faldende i takt med, at indkomststrukturen er præget af ophold i højere deciler. Den gennemsnitlige opholdstid i gruppe 1 i aldersgruppen 35-61 år er således 9 år mod 5,3 år i gruppe 4. Det er udelukkende i gruppe 1, at den gennemsnitlige varighed af et ophold i et udsat område overstiger længden af det typiske ophold for målgruppen som helhed.

Som tidligere beskrevet er andelen af almene boliger størst i gruppe 1, der er domineret af ophold i de laveste indkomstdeciler. Således vil 27 procent af målgruppen have boet et eller flere år i en almen bolig i alderen 30-34 år og 37 procent vil have gjort det i alderen 35-61 år. Denne tendens er en del højere end tilbøjeligheden for målgruppen som helhed. Således vil 18 procent af målgruppen have været bosat mindst ét år i en almen bolig i den unge aldersgruppe, mens det samme er tilfældet for 24 procent af den ældste aldersgruppe. Har man boet i en almen bolig i alderen 30-34 år er der 0,8 års forskel på den gennemsnitlige længde af opholdet mellem gruppe 1 og gruppe 4. Forskellen er 8,3 år, når den gennemsnitlige varighed for 35-61-årige opgøres. Varigheden af det gennemsnitlige ophold er dermed væsentligt lavere for den del af målgruppen, hvis indkomstforløb er kategoriseret i gruppen kendetegnet af ophold i de øverste indkomstdeciler.

Tabel 4.3. Grubebetinget fordeling af målgruppen på bopæls erfaring fra et udsat område og almen bolig.

Alder	Antal	Udsat områdeerfaring				Almen bolig erfaring			
		Andel, 1+ år		Gns. varighed, år		Andel, 1+ år		Gns. varighed, år	
		30-34	35-61	30-34	35-61	30-34	35-61	30-34	35-61
Gruppe 1	10.356	0,10	0,15	3,0	9,0	0,27	0,37	3,8	16,2
Gruppe 2	12.657	0,06	0,09	2,7	7,0	0,19	0,26	3,7	14,3
Gruppe 3	19.859	0,05	0,06	2,5	6,5	0,15	0,19	3,4	12,0
Gruppe 4	4.607	0,02	0,03	2,3	5,3	0,09	0,08	3,0	7,9
I alt	47.479	0,06	0,08	2,7	7,6	0,18	0,24	3,6	14,0

Anm.: Den gennemsnitlige varighed er ikke nødvendigvis udtryk for længden af et uafbrudt forløb. Varigheden beregnes udelukkende for individer med erfaring.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Overordnet motiverer ovenstående, at en lang række af personkarakteristika har betydning for, hvorledes indkomstforløbene klassificeres. Selv om gruppernes sammensætning på karakteristika i visse tilfælde varierer ganske betydeligt, er det centralt at hæfte sig ved, at samtlige dimensioner af et karakteristikum typisk vil være repræsenteret i større eller mindre grad inden for alle grupper. Eksempelvis kategoriseres personer med bosættelseserfaring i et udsat område ikke entydigt i én gruppe, hvorfor man kan slutte, at de ikke alene i deres egenskab af at have boet i et udsat område, vil have indkomstforløb, der afviger signifikant fra andres. Dette understreger blot, at der er variation i indkomststrukturen inden for de enkelte karakteristika, noget der også blev tydeliggjort ved, at uddannelsessammensætningen i gruppe 1 og 2 var relativt ensartet, mens den socioøkonomiske struktur var ganske forskellig. Dette understreger vigtigheden af ikke blot at se på sammenhængen mellem indkomstklassificeringen og et enkelt karakteristikum, men derimod kontrollere simultant for andre personspecifikke forhold. Dette forfølges nærmere nedenfor.

Statistisk kvalificering af forskelle i gruppetilhørsforhold

Ovenstående afsnit har givet en indikation på hvilke baggrundskarakteristika, der kan tænkes at have sammenhæng med målgruppens tilhørsforhold til hver af de fire grupper. Ved anvendelse af en *multinomial logistisk regression* redegøres nedenfor for i hvilket omfang udvalgte baggrundskarakteristika – herunder bopæls erfaring fra et udsat område - signifikant kan relateres til målgruppens tilbøjelighed til at tilhøre de fire grupper. Konkret ses der på, hvorvidt der er en signifikant sammenhæng mellem køn, oprindelse, uddannelsesniveau som 35-årig, erhvervstilknytning, lavindkomsterfaring, tilknytning til almen bolig, parforholdserfaring samt erfaring med bosætning i et udsat område og tilbøjeligheden til, at individerne i målgruppen tilknyttes gruppe 2, 3 eller 4 relativt til gruppe 1. Referencegruppen for estimationen er således gruppe 1, der som tidligere nævnt er kendetegnet ved, at ophold i de nederste indkomstdeciler er udpræget. Erhvervstilknytning, tilknytning til almen bolig, parforholdserfaring, lavindkomsterfaring og bosættelse i et udsat område er her definerede som mindst ét års erfaring/tilknytning i alderen 30-34 år og er boolske variable. Som tidligere nævnt er der for en del af disse variable en betydelig korrelation med den tilsvarende opgørelse i aldersgruppen 35-61 år, hvilket sammen med ambitionen om at mindske simultanitetsbias har ført til valget af opgørelsen af karakteristika i målgruppens unge aldre, der ligger tidsmæssigt forud for de indkomstforløb, der er grupperede. Bias foranlediget af udeladte variable og ikke målbare forhold spiller dog fortsat en rolle, hvorfor resultaterne skal fortolkes med forsigtighed og væsentligst uden vurdering af kausalitet. Gennemgangen af resultaterne vil fokusere på omfanget af signifikans og vil ellers være rettet mod kvalitative konklusioner.

Ved at inddrage flere forskellige typer af karakteristika i estimationen søges bias foranlediget af udeladte variable mindsket. Dette er dog på bekostning af en øget risiko for multikollinearitet. Analysen omfatter med en enkel og væsentlig undtagelse ikke krydseffekter mellem de forklarende variable. Der kan eksempelvis ikke drages inferens omkring hvorvidt kvinder, der har været bosat i et udsat område har en større tilbøjelighed til at tilhøre gruppe 2 fremfor gruppe 1 end kvinder, der ikke har været bosat i et sådant område. De forudgående afsnit har motiveret, at bopæls erfaring fra et udsat område ikke nødvendigvis i sig selv giver anledning til, at et indkomstforløb klassificeres i en gruppe frem for en anden. For at belyse en potentiel samvariation nærmere, er bopæls erfaring i et udsat område nedenfor inddraget i kombination med arbejdsmarkedstilknytning, hvilket dermed udgør modellens eneste interaktionsled. Således er der ikke set på, hvorvidt bopæls erfaring og erhvervsdeltagelse samvarierer med eksempelvis oprindelse eller uddannelsesniveau. Resultaterne af estimationen, der præsenteres i Tabel 4.4 skal derfor fortolkes med tanke på dette²⁴.

Estimationen viser, at ikke-vestlige indvandrere relativt til personer af dansk oprindelse har en signifikant lavere tilbøjelighed til at optræde i gruppe 2, 3 eller 4 relativt til gruppe 1. Sagt på en anden måde er det for ikke-vestlige indvandrere mere sandsynligt, at deres indkomstforløb er klassificeret i gruppen kendetegnet ved de laveste indkomstdeciler frem for i de øvrige grupper end tilfældet er for personer med dansk oprindelse. Vestlige indvandrere vil have en større tilbøjelighed til at optræde i gruppe 4 frem for gruppe 1 end personer af dansk oprindelse. Vestlige indvandrere vil dermed i større udstrækning end personer af dansk oprindelse have et indkomstforløb blandt de højeste end blandt de laveste indkomstdeciler. Konklusionerne omkring klassifikationen af indvandreres indkomstforløb relativt til forløbene knyttet til personer af dansk oprindelse er i overensstemmelse med gruppernes oprindelsesstruktur, jf. Tabel 4.1 og den indbyrdes forskel mellem oprindelsesgruppernes indkomststruktur, jf. Figur 4.3. Som tidligere nævnt omfatter målgruppen relativt få efterkommere og det er således også blot for ikke-vestlige efterkommere, at der kan drages signifikant inferens. Ikke-vestlige efterkommere har en signifikant lavere tendens til at optræde i gruppe 4 relativt til gruppe 1 end tilfældet er for personer med dansk oprindelse. Der er ingen signifikant forskel mellem ikke-vestlige efterkommere og personer af dansk oprindelse hvad angår placering i medianindkomstgrupperne relativt til gruppe 1.

Kvinder har en lavere tilbøjelighed end mænd til at have et indkomstforløb, der er kategoriseret i grupperne med de højeste eller middelhøje deciler fremfor i gruppe 1. Denne slutning er også i overensstemmelse med gruppernes kønsstruktur, jf. Tabel 4.1 og forholdet mellem mænd og kvinders indkomststruktur i Figur 4.2. Der kontrolleres både for erhvervsdeltagelse i de unge aldre og uddannelsesniveau i estimationen. Ud over at være knyttet til en deltagelseseffekt i den ældre aldersgruppe, kan den større relative forskel mellem tilordning til højindkomstgrupperne fremfor lavindkomstgruppen for kvinders vedkommende afspejle en produktivitsforskelle, som der ikke er kontrolleret for.

²⁴ Bemærk, at inddragelse af interaktionen mellem arbejdsmarkeds- og områdeerfaring ikke svarer til, at effekten af at have boet i et udsat område er estimeret for personer hhv. inden og uden for arbejdsstyrken. Analysen i Tabel 4.4 udtaler sig alene om, hvorvidt effekten af arbejdsmarkedserfaring er betinget af områdeerfaring under den centrale forudsætning, at alle andre forhold er konstante og dermed ikke tillades at variere på tværs af hverken arbejdsmarkeds- eller områdeerfaring. En vurdering af effekten af områdeerfaring for personer hhv. med og uden arbejdsmarkedserfaring indebærer, at effekten af de øvrige forklarende variable også tillades at variere på tværs af arbejdsmarkedserfaring.

Tabel 4.4. Multinomial logistisk regression af individkarakteristika på gruppetilhørsforhold. Log odds. Model med interaktion.

	Gruppe 2 vs. gruppe 1	Gruppe 3 vs. gruppe 1	Gruppe 4 vs. gruppe 1
Indvandrere, ikke-vestlig (ref, dansk)	-0,846*** (-0,104)	-1,009*** (-0,122)	-0,767*** (-0,286)
Indvandrere, vestlig	-0,078 (-0,11)	0,122 (-0,115)	0,609*** (-0,186)
Efterkommer, ikke-vestlig	-0,084 (-1,432)	1,291 (-1,234)	-6,561*** (-0,002)
Efterkommer, vestlig	0,076 (-0,396)	0,499 (-0,374)	0,674 (-0,502)
Kvinde	0,041 (-0,031)	-0,964*** (-0,03)	-2,384*** (-0,056)
Gymnasial (ref, grundskole)	0,256*** (-0,07)	1,466*** (-0,066)	2,445*** (-0,095)
Erhvervsfaglig	0,270*** (-0,032)	0,803*** (-0,033)	0,793*** (-0,063)
KVU	0,398*** (-0,093)	1,640*** (-0,087)	1,905*** (-0,128)
MVU	0,553*** (-0,062)	2,530*** (-0,057)	3,113*** (-0,083)
LVU	-1,308*** (-0,213)	1,927*** (-0,125)	4,437*** (-0,135)
Mindst 1 år i de to laveste indkomstdeciler, 30-34 år	-1,060*** (-0,029)	-1,795*** (-0,032)	-1,674*** (-0,061)
Mindst 1 år i almen bolig, 30-34 år	-0,230*** (-0,037)	-0,362*** (-0,038)	-0,894*** (-0,076)
Mindst 1 år i et parforhold, 30-34 år	0,427*** (-0,041)	0,702*** (-0,041)	1,174*** (-0,077)
(a) Mindst 1 år i arbejdsstyrken, 30-34 år	1,058*** (-0,099)	1,545*** (-0,154)	1,612*** (-0,427)
(b) Mindst 1 år i et udsat område, 30-34 år	-0,261 (-0,271)	0,583 (-0,369)	0,861 (-1,124)
(c) Mindst 1 år i arbejdsstyrken x mindst 1 år i et udsat område, 30-34 år	0,14 (-0,275)	-0,760** (-0,372)	-1,475 (-1,13)
Konstant	-0,767*** (-0,107)	-0,986*** (-0,159)	-3,259*** (-0,432)

Anm.: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standardafvigelse er angivet i parentes. Forløb med ukendt boligtype (48 stk.) er frasorteret i estimationen.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Har man en gymnasial uddannelse frem for en grundskoleuddannelse er det signifikant mere sandsynligt at have et indkomstforløb, der er kategoriseret i gruppe 2 til 4 fremfor gruppe 1. Som tidligere nævnt er andelen med en grundskoleuddannelse som højst fuldførte lavere i gruppe 2 til 4 relativt til gruppe 1, mens andelen af gymnasialt uddannede er højere i såvel

gruppe 3 som 4 relativt til gruppe 1, men graden af signifikans ville vanskeligt kunne vurderes ud fra Figur 4.15. En erhvervsfaglig uddannelse fremfor en grundskoleuddannelse har også en signifikant betydning for, hvorvidt indkomstforløbet er kategoriseret i gruppe 2 til 4 fremfor gruppe 1. Dette er i overensstemmelse med, at grupperne med såvel median som de højeste indkomstforløb er kendetegnet ved, at andelen med en erhvervsfaglig uddannelse her dominerer andelen med en grundskoleuddannelse, jf. Figur 4.15. Endvidere er indkomststrukturen for erhvervsfaglige præget af en større grad af ophold i de højere indkomstdeciler end tilfældet er for grundskoleuddannede, jf. Figur 4.5.

Med en enkelt undtagelse, vil personer med en videregående uddannelse have en større tilbøjelighed end grundskoleuddannede til at optræde i gruppe 2 til 4 fremfor gruppe 1. Hvor andelen med en grundskoleuddannelse er betydeligt lavere i gruppe 3 og 4 sammenholdt med gruppe 1, mens andelen med en videregående uddannelse er højere, ville det være en udfordring at drage signifikante konklusioner på baggrund af en visuel sammenligning af uddannelsessammensætningen i gruppe 1 og gruppe 2. Har man en lang videregående uddannelse, vil man have en mindre tendens til at være i gruppe 2 end i gruppe 1, relativt til en person med en grundskoleuddannelse. Som tilfældet er det i gruppe 2, udgør andelen med en LVU blot en beskedent andel af befolkningen i gruppe 1.

Boligtype er repræsenteret i estimationen gennem erfaring fra en almen bolig i alderen 30-34 år. Har man boet i en almen bolig i mindst ét år i denne periode, har man en relativt lavere tilbøjelighed til at optræde i gruppe 2 til 4 fremfor lavindkomstgruppen relativt til en person, der ikke har boet i en almen bolig. Andelen, der har boet i en almen bolig i alderen 35-61 år er højest i de indkomstgrupper, hvor andelen også var højest i den unge aldersgruppe, jf. Tabel 4.3. Eftersom der således kan være en tendens til, at personer, der har boet i almen bolig i de unge aldre også vil gøre det senere og da almene boliger er mest udbredte i gruppe 1, jf. Figur 4.19, er dette i overensstemmelse med, at personer med almen erfaring fra 30-34-årsalderen hyppigere tilordnes lavindkomstgruppen relativt en person uden almen erfaring.

Har man på et tidspunkt i løbet af alderen 30-34 år indgået i et parforhold, har man en signifikant større tilbøjelighed til at være tilordnet gruppe 2, 3 eller gruppe 4 relativt til gruppe 1. Udviklingen i familiestrukturen over tid i hver af de fire grupper i Figur 4.18 er i overensstemmelse med den analytiske konklusion.

Har man opholdt sig i et af de to laveste indkomstdeciler i løbet af perioden, hvor man var 30-34 år, er man relativt mere disponeret for, at indkomstforløbet i den efterfølgende periode tilordnes lavindkomstgruppen frem for gruppe 2 til 4. Dette kunne indikere en vis persistens i indkomsten og dermed også arbejdsmarkedsdeltagelsen over tid.

Omfanget af arbejdsmarkedsdeltagelse i den unge aldersgruppe indgår dels i estimationen gennem en angivelse af, hvorvidt man i den pågældende aldersklasse har været tilknyttet arbejdsmarkedet i mindst ét år og dels i kombination med variabelen indikerende ophold i et udsat område for at belyse eventuelle samspilseffekter mellem de to forhold. Koefficienten knyttet til den selvstændige variabel for arbejdsmarkedsdeltagelse er signifikant positiv for gruppe 2 til 4. Dette betyder, at tilknytning til arbejdsmarkedet blandt personer, der *ikke* har boet i et udsat område i alderen 30-34 år, giver anledning til en klassificering i gruppe 2, 3 eller 4 fremfor gruppe 1, jf. (a) i Tabel 4.4. Den selvstændige variabel angivende ophold i et udsat område er omvendt ikke signifikant, hvilket betyder, at bosætning i et udsat område ikke har betydning for gruppetilhørsforholdet, såfremt man *ikke* har været tilknyttet arbejdsstyrken i alderen 30-34 år, jf. (b) i Tabel 4.4. At interaktionsleddet mellem de to forhold heller ikke er signifikant (på 1 pct. niveau) betyder, at effekten af erhvervs erfaring på

gruppetilhørsforholdet ikke er betinget af områdeerfaring ligesom betydningen af områdeerfaring ikke er betinget af tidligere af erhvervserfaring, jf. (c) i Tabel 4.4²⁵.

Opsummerende indikerer estimationen, at der er en signifikant sammenhæng mellem arbejdsmarkedstilknytning og gruppetilhørsforhold uafhængig af områdeerfaring. Uanset om man i alderen 30-34 år har været på arbejdsmarkedet eller ikke, er der ingen sammenhæng mellem at have boet i et udsat område og tilhørsforhold til de fire indkomstgrupper. Der er tale om en marginalbetragtning betinget af, at de øvrige karakteristika er uændrede.

For at belyse betydningen af interaktionsleddet for konklusionerne er der også estimeret en model uden interaktion. Denne model indikerer, at der som ovenfor er en signifikant positiv effekt af arbejdsmarkedstilknytning, der motiverer tilordning i gruppe 2 til 4 fremfor gruppe 1. Herudover vil personer med tidligere bosætning i et udsat område dog have en signifikant lavere tilbøjelighed for at optræde i gruppe 3 og 4 fremfor gruppe 1, jf. (a) og (b) i Tabel E.1 i appendiks E²⁶. Relativt til resultaterne i Tabel 4.4 kan det ikke udelukkes, at den signifikante effekt fra beboelse i et udsat område stammer fra korrelation med andre bagvedliggende forhold end arbejdsmarkedserfaring. Søgningen efter eventuelle heterogene eller betingede effekter forfølges dog ikke nærmere i dette afsnit.

Det skal understreges, at konklusionerne ovenfor alene er relateret til den betragtede målgruppe og vil derfor være betinget af de antagelser, der er lagt til grund for udvælgelsen af denne. Det skal endvidere fremhæves, at effekten af bopæls- og erhvervserfaring på tilhørsforholdet til indkomstgrupperne er betinget af, at øvrige forklarende variable er uændrede. Der er dermed tale om en gennemsnitsbetragtning på tværs af øvrige karakteristika. Der tages således ikke højde for samvariation mellem de øvrige karakteristika og hhv. område- og erhvervserfaring. Vurderingen af forholdet mellem områdeerfaring og gruppetilhørsforhold kan derfor sagtens vise sig signifikant for visse delpopulationer, mens det ikke er tilfældet for andre²⁷.

4.3 Identifikation af signifikante relationer mellem baggrundskarakteristika og indkomstforløb

Som nævnt indledningsvist ignorerer tilgangen ovenfor den variation i indkomstforløb, der fortsat kan eksistere inden for de enkelte indkomstgrupper. Nedenfor vil der derfor blive set nærmere på, hvilke karakteristika, der signifikant kan relateres til forskelle i indkomstforløb uafhængigt af klassifikationsgrupper. Som ovenfor peger resultaterne i retning af, at ophold i et udsat område ud fra en gennemsnitlig betragtning ikke kan knyttes signifikant til forskelle i individuelle forløb. Forskellen synes i højere grad knyttet til uddannelsesniveaue og

²⁵ Koefficienten til interaktionsleddet udtrykker forskellen mellem log odds, der sammenligner personer med og uden erhvervserfaring blandt ikke-udsatte og log-odds, der sammenligner personer med og uden erhvervserfaring blandt udsatte. Alternativt udtrykker interaktionen forskellen mellem log odds, der sammenligner personer med og uden erfaring fra et udsat område blandt ikke-arbejdsmarkedserfarne og log-odds, der sammenligner personer med og uden erfaring fra et udsat område blandt personer med arbejdsmarkedserfaring.

²⁶ Formålet med den logistiske regression i dette afsnit er primært at illustrere en tilgang til statistisk analyse af tilhørsforholdet til de fire indkomstgrupper frem for optimering af variabelvalget i modellen. Det skal dog nævnes, at forskellen mellem AIC og McFaddens pseudo R^2 mellem modellen i Tabel 4.4 og modellen i Tabel E.1 er hhv. mindre end 1 og mindre end 0,0001, hvorfor andre forhold end disse størrelser bør udgøre grundlaget for et eventuelt modelvalg. Sigtes mod optimering af modelvalget bør der generelt testes for flere typer af interaktioner, end det er tilfældet i nærværende eksempel.

²⁷ Eksempelvis kan man opnå konklusioner, der afviger fra den gennemsnitlige ved at inddrage interaktionen mellem områdestatus og uddannelsesniveaue eller ved at analysere områdeeffekten for hver enkel uddannelsesgruppe.

arbejdsmarkedstilknytning, men resultaterne udelukker ikke, at konklusionen kan variere for undergrupper af befolkningen.

Grupperingen af livsforløb i afsnit 4.2 er som nævnt baseret på en afstandsmatrix, der vurderer den indbyrdes afstand mellem samtlige unikke forløb og som i det givne tilfælde har dimensionerne 45.654 x 45.654. Regresseres baggrundskaraktistika på dette afstandsmål, kan resultatet belyse, hvorvidt der er en signifikant sammenhæng mellem kovariaterne og forskellene mellem indkomstforløbene. Herunder er det muligt at identificere, hvilke kovariater, der har den stærkeste relation til forskellen i indkomstforløb. Der foretages indledningsvist en *univariat analyse*, hvor udvalgte personkaraktistika analyseres enkeltvist. Dernæst udarbejdes for en stikprøve af målgruppen en *multivariat analyse* til belysning af betydningen af samtidig indregning af flere forskellige karaktistika. Igen er det vigtigt at påpege, at analysen ikke giver anledning til at drage konklusioner om det kausale forhold mellem karaktistika og forskelle i indkomstforløb. Ligeledes skal resultaterne af specielt den univariate analyse fortolkes med det forhold, at en signifikant relation blot kan afspejle, at det pågældende karaktistikum er korreleret med en anden baggrundsvariabel, herunder specielt et ikke målbart forhold.

Tilgangen minder om en variansanalyse (ANOVA), men adskiller sig herfra ved, at det her ikke kan forudsættes, at forløbene er normalfordelte. Signifikansen af de enkelte baggrundskaraktistikas bidrag til at forklare den samlede variation, kan dermed ikke vurderes ud fra en *F*-fordeling. I stedet estimeres *p*-værdierne på baggrund af simulationer, jf. Gabandinho (2011c).

Ønsker man eksempelvis at se på forholdet mellem køn og forskelle i indkomstforløb beregnes først en *F*-værdi på baggrund af den observerede relation mellem køn og indkomstforskelle for målgruppen. *F*-værdien sammenholder variabiliteten i indkomstforskelle mellem køn med variabiliteten i forskellen inden for hvert køn. Nulhypotesen er, at der ikke er forskel mellem kønnene, så en stor observeret *F*-værdi er forenelig med, at hypotesen forkastes og køn dermed giver anledning til en signifikant forskel i indkomst. Da man typisk ikke kan vurdere signifikansen ud fra en *Fisher*-fordeling, udarbejdes ved permutationstest en fordeling baseret på antagelserne i nulhypotesen, jf. Anderson (2001). Nulhypotesens antagelse om, at der ikke er forskel mellem kønnene, er ækvivalent med, at kønskaraktistika kan fordeles tilfældigt blandt målgruppens observationer. Man kan med andre ord ombytte indkomstforløbene tilfældigt uafhængigt af køn. Denne proces gennemføres i praksis et stort antal gange og ved hver iteration beregnes en *F*-værdi. Man har da konstrueret en pseudo *F*-fordeling, som man kan sammenholde den observerede *F*-værdi med. Jo flere af de simulerede *F*-værdier, der overstiger den observerede *F*-værdi, jo mere sandsynligt er det, at nulhypotesen om, at køn er uden betydning godkendes.

Den univariate analyse nedenfor er her baseret på 1.000 simulerede *F*-værdier, hvilket er i overensstemmelse med vurdering af signifikans på et 5 procent niveau, jf. Manly (2007). Af Tabel 4.5 fremgår det, at der for målgruppen ikke er en signifikant sammenhæng mellem alle variable og den indbyrdes forskel i indkomstforløbene. Eftersom den afhængige variabel her udgøres af forskelle i indkomstforløb frem for niveauer, er der også betragtet relationen til variable knyttet til aldersgruppen 35-61 år, der ellers tidligere har været udeladt med henblik på at reducere simultanitetsbias. Denne type bias skønnes at spille en mindre rolle ved analyse af forskelle frem for niveauer.

Ikke overraskende er det varigheden af ophold i de to laveste indkomstdeciler i løbet af den analyserede periode, der har den stærkeste association til forskellen i indkomstforløb. Således kan antallet af år blandt befolkningen med de 20 procent laveste indkomster forklare

13 procent af den samlede variation i indkomstforløb, jf. Tabel 4.5. Herudover er der et signifikant forhold mellem forskelle i indkomstforløb og såvel antal års tilknytning til arbejdsmarkedet som uddannelsesniveaue. Køn, oprindelse og såvel indkomst som tilknytning til arbejdsmarkedet i alderen 30-34 år spiller også en signifikant rolle. At man på et tidspunkt har boet i en almen bolig kan også signifikant relateres til indkomstforskelle, men længden af den periode, man har boet i en almen bolig har ingen betydning. At man har boet i et udsat område i alderen 30-34 år og længden af opholdet i samme aldersgruppe ligger på grænsen til at have en betydning for forskellen i indkomstforløb i alderen 35-61 år, mens dette ikke er tilfældet, når tilhørsforholdet til et udsat område betragtes for aldersgruppen 35-61 år.

Det er således i overensstemmelse med de foregående resultater, at det i høj grad er socioøkonomiske og statistiske demografiske karakteristika, der kan relateres til indbyrdes forskelle i indkomstforløb.

Tabel 4.5. Test af univariate relationer mellem baggrundskarakteristika og indkomstforløb.

	<i>F</i>	<i>R</i> ²	<i>p</i>
Antal år i blandt de 20 procent laveste indkomster, 35-61 år	257,18	0,13	0,00**
Mindst 1 år i blandt de 20 procent laveste indkomster, 35-61 år	3360,28	0,07	0,00**
Antal år tilknyttet arbejdsmarkedet, 35-61 år	114,54	0,06	0,00**
Uddannelsesniveau, 35 år	588,30	0,06	0,00**
Antal år i blandt de 20 procent laveste indkomster, 30-34 år	361,00	0,04	0,00**
Mindst 1 år i blandt de 20 procent laveste indkomster, 30-34 år	1519,99	0,03	0,00**
Antal år i parforhold, 35-61 år	38,99	0,02	0,00**
Oprindelse	970,33	0,02	0,00**
Antal år tilknyttet arbejdsmarkedet, 30-34 år	179,76	0,02	0,00**
Boligtype, 61 år	136,33	0,01	0,01**
Mindst 1 år tilknyttet arbejdsmarkedet, 30-34 år	492,52	0,01	0,01**
Mindst 1 år tilknyttet arbejdsmarkedet, 35-61 år	450,53	0,01	0,01**
Mindst 1 år i almen bolig, 35-61 år	347,04	0,01	0,00**
Boligtype, 35 år	79,21	0,01	0,05**
Mindst 1 år i almen bolig, 30-34 år	181,71	0,00	0,00**
Køn	30,10	0,00	0,03**
Antal år i almen bolig, 35-61 år	17,13	0,01	0,22
Antal år i almen bolig, 30-34 år	41,60	0,00	0,12
Antal år i udsat område, 35-61 år	7,50	0,00	0,10
Mindst 1 år i udsat område, 35-61 år	153,17	0,00	0,10
Mindst 1 år i parforhold, 35-61 år	118,84	0,00	0,08
Antal år i parforhold, 30-34 år	23,38	0,00	0,32
Antal år i udsat område, 30-34 år	17,66	0,00	0,07
Mindst 1 år i parforhold, 30-34 år	81,89	0,00	0,12
Mindst 1 år i udsat område, 30-34 år	77,23	0,00	0,06

*Anm.: Baseret på 1.000 permutationer. Signifikansen af de enkelte variable er angivet med ** ($p < 0.05$).*

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Associationstestet i Tabel 4.5 udtaler sig alene om den isolerede betydning af de enkelte variable for diversiteten i indkomstforløb. Kovariaternes samlede forklaringssevne kan klarlægges gennem en slags multifaktoranalyse, der på samme vis som i Tabel 4.5 etablerer, hvorvidt de enkelte baggrundsvariables forklaringssevne er signifikant, men under forudsætning af, at flere variable analyseres på en gang. Signifikansen af et karakteristikum

er dermed betinget af tilstedeværelsen af eventuelt samkorrelerede variable. Som i de øvrige tilfælde, kan en signifikant effekt dog fortsat være foranlediget af korrelation med ikke indregnede karakteristika ligesom den gennemsnitlige konklusion kan afvige fra den, som en analyse af undergrupper afgrænset ved eksempelvis køn, oprindelse, arbejdsmarkedstilknytning og uddannelse giver anledning til.

Som tilfældet var ovenfor er det også nødvendigt at vurdere signifikansen af variablene i en multifaktormodel med afsæt i en simuleret fordeling. Her kræves et simulationstest for den samlede model og for hver enkel baggrundsvariabel, jf. Gabandinho, A. et al. (2011c). Processen viser sig beregningstung for det multivariate tilfælde og etablering af nulhypotesens sandsynlighedsfordelinger er derfor kun baseret på 300 permutationer og en 10 procent stikprøve af målgruppen omfattende ca. 4.600 personer, hvorfor nedenstående skal fortolkes med disse forbehold.

I Tabel 4.6 ses, at modellen som helhed omfatter signifikant information omkring forskelligheden i forløbene og den samlede forklaringsevne er omkring 24 procent. Erfaring med parforhold og bosætning i et udsat område i alderen 30-34 år er fortsat ikke signifikant knyttet til den indbyrdes forskel i indkomstforløb. Ophold i almen bolig har dog i modsætning til det univariate tilfælde ikke længere betydning. Uddannelsesniveau, tidligere arbejdsmarkedstilknytning og indkomstforhold udviser fortsat et signifikant forhold til forskelle i indkomstforløb selv om de nu optræder samtidigt i modellen og er oplagt samvarierende. Uddannelse repræsenterer den variabel, der har den største forklaringsevne i modellen. Med tanke på det bredere konfidensinterval som færre permutationer afstedkommer, er oprindelse nu på grænsen til ikke at have signifikant betydning for indkomstforskelle, mens køn fortsat er signifikant. Dette indikerer, at oprindelseeffekten nu kan være delvist beskrevet gennem korrelation med de øvrige forklarende variable.

Tabel 4.6. Test af multivariat relation mellem baggrundskarakteristika og indkomstforløb.

	<i>F</i>	<i>R</i> ²	<i>p</i>
Uddannelsesniveau, 35 år	199,07	0,14	0,00*
Mindst et år i blandt de 20 procent laveste indkomster, 30-34 år	99,86	0,01	0,00*
Køn	97,55	0,01	0,00*
Mindst et år tilknyttet arbejdsmarkedet, 30-34 år	23,86	0,00	0,01*
Oprindelse	3,21	0,00	0,05*
Mindst 1 år i almen bolig, 30-34 år	7,39	0,00	0,19
Mindst 1 år i parforhold, 30-34 år	4,37	0,00	0,12
Mindst ét år i udsat område, 30-34 år	1,01	0,00	0,95
Total	116,22	0,24	0,00*

Anm.: Baseret på 300 permutationer og en 10 procent stikprøve af målgruppen. For en multivariat model med færre karakteristika er ændringen i p-værdien ved forskellige antal simulationer betragtet og der er her fundet, at der opnås en relativ grad af stabilitet efter 300 permutationer. Få permutationer giver dog generelt anledning til, at intervallet inden for hvilken en p-værdi kan vurderes at ligge over eller under 5 procent bliver bredere. I det givne tilfælde kan modellen på et 5 procent signifikansniveau således ikke sondre mellem p-værdier mellem 0,025 og 0,075, hvorfor signifikansen i tabellen blot er markeret med en enkelt () svarende til signifikans på 10 procent niveau.*

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Mens både den univariate og multivariate analyse kan sige noget om, hvilke personspecifikke karakteristika, der har betydning for indbyrdes forskelle i indkomstforløb, så kan man ikke rigtigt uddrage information om, hvordan indkomststrukturen faktisk varierer på tværs af værdierne inden for det enkelte karakteristikum. Dette kunne delvist afdækkes ved at

gennemføre ovenstående analyser på forskellige undergrupper af eksempelvis køn, oprindelse, arbejdsmarkedstilknytning og uddannelsesniveau, men kan også delvist adresseres ved et regressionstræ. I modsætning til multifaktoranalysen klarlægger et regressionstræ således den indbyrdes relation mellem alle baggrundskarakteristika. Manglen på interaktion mellem baggrundskarakteristika er også et oplagt kritikpunkt for den logistiske regression i det foregående afsnit.

Ved anvendelse af et regressionstræ opdeles indkomstforløbene for målgruppen rekursivt gennem binære split betinget af de forklarende karakteristika således, at der opnås størst mulig forskel mellem forløbene på tværs af grupperne og størst mulig ensartethed i forløbene inden for grupperne. Opdelingen baseres på det karakteristikum, der isoleret set repræsenterer den største andel af variationen i indkomstforløbene. Som i de univariate og multivariate analyser ovenfor opgøres associationen mellem variabelen og indkomstforløbene med afsæt i simulerede fordelinger, mens opdelingen inden for det valgte karakteristikum også er fastlagt på baggrund af permutationer, jf. Gabandinho et al. (2011c).

Proceduren er beregningstung og der er derfor blot udarbejdet et illustrativt eksempel baseret på samme stikprøve, som anvendtes til den multivariate analyse i Tabel 4.6. Regressionstræet omfatter de variable, som var signifikante i den multivariate analyse. Stopkriteriet for de binære split er en vurdering af F -værdien på et 5 procent signifikansniveau, en maksimal dybde på fem niveauer og et krav om, at der efter et split skal være mindst 30 observationer i en gruppe²⁸. Regressionstræet kan ses i Figur D.3 i appendiks D.

Uddannelsesniveauet repræsenterer det første split og det fastlægges, at en adskillelse af personer med en lang videregående uddannelse fra resten af befolkningen giver anledning til den største diversitet mellem forløbene. Det bemærkes, at indkomststrukturen i gruppen bestående af personer med en LVU næsten entydigt er karakteriseret ved ophold i den øverste indkomstdecil. For gruppen bestående af personer med en lang videregående uddannelse identificeres køn dernæst som den variabel, der repræsenterer den størst mulige diversitet i forløbene, hvorefter der ikke sker yderligere opdeling af personer med det højeste uddannelsesniveau.

For personer, der ikke har en LVU, er det indkomsterfaringen fra aldersgruppen 30-34 år, der udvælges som det karakteristikum, der giver anledning til størst mulig indbyrdes forskel mellem indkomstforløbene. Det ses, at godt 1.200 personer med tidligere ophold i de to laveste indkomstdeciler udskilles fra de personer, der ikke har lavindkomsterfaring. Indkomststrukturen i alderen 35-61 år er for personer med lavindkomsterfaring kendetegnet ved primært at omfatte indkomstdeciler under medianen, mens strukturen for personer uden lavindkomsterfaring omvendt er præget af en indkomststruktur omkring eller over medianen.

Blandt personer med lavindkomsterfaring sondres dernæst mellem om man har været tilknyttet arbejdsstyrken i alderen 30-34 år eller man ikke har. Fravær af erhvervs erfaring ses udelukkende at omfatte personer i de fire laveste deciler, hvor det primært er kvinder, der udgør den laveste indkomstdecil. De fleste i lavindkomstgruppen har haft erhvervs erfaring, hvorfor indkomststrukturen for denne del ikke vil adskille sig væsentligt fra lavindkomstgruppen som helhed.

Betragter vi nu den del af målgruppen, der ikke tidligere har opholdt sig i de to nederste indkomstdeciler, udvælges køn som det karakteristikum, der giver anledning til den største

²⁸ Inddragelse af erfaring med ophold i et udsat område giver ikke anledning til, at indkomstforløbene opdeles på denne variabel.

diversitet. Indkomststrukturen for mænd ses at være en anelse mere præget af ophold i de øverste indkomstdeciler end tilfældet er for kvinder. For begge køn sker efterfølgende en udskillelse af personer med en grundskole eller erhvervsfaglig uddannelse som højst fuldførte. For begge køn vil indkomststrukturen blandt personer med disse uddannelsesniveauer være præget af ophold i lavere indkomstdeciler end blandt personer med en gymnasial eller videregående uddannelse.

Tilgangen illustrerer dermed, at forklaringsevnen af et baggrundskarakteristikum, vil variere på tværs af de øvrige karakteristika. Således har tidligere erfaring med indkomst i de laveste indkomstdeciler relativt større betydning, hvis man ikke har en lang videregående uddannelse end hvis man har. Eksempelvis er køn også en vigtigere diskriminationsfaktor, hvis man ikke har lavindkomsterfaring end hvis man har. Tilsvarende vil erhvervs erfaring spille en større rolle end køn, hvis man tidligere har opholdt sig i lavindkomstgruppen. At forklaringsevnen af et karakteristikum påviseligt varierer afhængigt af værdien af andre personspecifikke forhold motiverer, at de statistiske analyser i dette afsnit fremadrettet også gennemføres for udvalgte undergrupper afgrænset af eksempelvis uddannelse, arbejdsmarkedstilknudning eller oprindelse.

Hvor grupperingen af indkomstforløb i afsnit 4.2 skete uafhængigt af individernes baggrundskarakteristika, kan regressionstræet alternativt ses som en mere detaljeret klassifikation, der er funderet i personspecifikke forhold. Som i det foregående afsnit gælder også her, at konklusionerne er knyttet specifikt til målgruppen, hvorfor de kan ændres, hvis kriterierne for valg af målgruppe revideres.

5. Effekt af bosætning i udsatte boligområder: Et quasi-randomiseret eksperiment

Hvordan påvirker det et menneske at bo i et udsat boligområde? Der findes efterhånden en betydelig litteratur omhandlende såkaldte nabolageeffekter, jf. Graham (2018) og Manski (1993). En klassisk problemstilling i denne litteratur er selvselektion: Har man en lav indkomst, fordi man bor i et udsat område, eller er det omvendt? I kapitel 4 understregedes det, at vi ikke havde løst selvselektionsproblemet. I dette kapitel udføres en analyse, hvor vi løser problemet ved at kombinere et dansk quasi-randomiseret datasæt med de i kapitel 2 definerede udsatte områder. Det faktum, at vi definerer udsatte områder ud fra objektive socioøkonomiske kriterier og over mange år, giver en unik mulighed for at vurdere effekten af at bo i sådanne udsatte boligområder. Det er muligt at påvise betydelige effekter på arbejdsmarkedstilknytning, mens der ingen effekter findes på kriminalitetsomfanget. Disse effekter måles både på kort og på langt sigt²⁹.

Analysen foretages på et datasæt, der beskriver flygtninge ankommet til Danmark i perioden 1986-1998. Dansk Flygtningehjælp stod i denne periode for at fordele de nyankomne flygtninge ud over Danmark. Hovedformålet med den anvendte fordelingspolitik var at fordele flygtninge ud i amter og kommuner efter indbyggertal. I analysen indgår i alt 70.753 flygtninge, hvoraf 12.678 blev boligplaceret i et udsat boligområde. Det vurderes i Damm (2005), at datasættet kan betragtes som et såkaldt naturligt eksperiment.

Det vises, at bosættelse i et udsat boligområde har en signifikant negativ effekt på flygtnings erhvervsfrekvens. Analysen finder, at erhvervsfrekvensen for flygtninge allokert til et udsat boligområde på både kort og langt sigt er signifikant lavere end erhvervsfrekvensen for flygtninge, der initialt blev placeret i et ikke-udsat boligområde.

På kort og mellemlangt sigt har placeringen i et udsat boligområde for begge køn en negativ effekt på erhvervsfrekvensen. Det er dog kvinders arbejdsmarkedsadfærd, der påvirkes mest negativt. Effekten af initialt at blive placeret i et udsat boligområde findes på langt sigt at være forskellig for mænd og kvinder. På langt sigt finder analysen ingen signifikant forskel i erhvervsfrekvensen for mandlige flygtninge, der er allokert til et udsat boligområde i forhold til mandlige flygtninge, der blev allokert til et ikke-udsat boligområde. Modsat mænd har kvinder på langt sigt en signifikant lavere erhvervsfrekvens som en konsekvens af den initiale allokering til et udsat boligområde.

Effekten på erhvervsfrekvensens størrelse er forskellig afhængigt af, hvilket land man flygter fra. På kort sigt oplever flygtninge fra Balkanhalvøen og Asien en stor negativ erhvervsfrekvenseffekt, mens der på langt sigt ikke findes nogen signifikant effekt. Flygtninge fra Afrika og Mellemøsten oplever både på kort, mellemlangt og langt sigt en stabil negativ effekt på arbejdsmarkedstilknytningen af initialt at blive placeret i et udsat boligområde.

Der findes ingen samlet signifikant effekt på kriminalitet af at bo i et udsat boligområde. Denne manglende effekt findes stort set også uanset køn, flygtnings oprindelsesområde og hvad enten vedkommende var voksen eller barn på indvandringstidspunktet.

I policy-sammenhæng er det nyttigt at vide, hvorvidt resultaterne kan generaliseres til en større befolkningsgruppe. Det afsøges derfor afslutningsvist, hvorvidt den gennemsnitlige indsatseffekt på erhvervsdeltagelsen og kriminalitetsomfanget kan generaliseres til tre

²⁹ Det er valgt at fokusere på arbejdsmarkedstilknytning, som er mindre konjunkturfølsom end for eksempel beskæftigelsesfrekvens og personindkomst.

forskellige målgrupper, hvoraf de to første omfatter flygtninge og familiesammenførte karakteriseret ved forskellige indvandringsperioder, mens den sidste omfatter alle ikke-vestlige indvandrere indvandret i 1986 eller senere. Det viser sig, at de gennemsnitlige indsatseffekter kan generaliseres til de to grupper omfattende flygtninge og familiesammenførte, men ikke til gruppen bestående af ikke-vestlige indvandrere.

5.1 Det naturlige eksperiment: Geografisk fordeling af flygtninge

Det naturlige eksperiment, som danner grundlag for denne undersøgelse, er tidligere anvendt adskillige gange af Damm, jf. Damm (2005, 2009a, 2009b, 2014b), Damm og Rosholm (2010), Damm og Dustman (2014a) samt Dustman, Vasiljeva og Damm (2016). Den efterfølgende beskrivelse af det naturlige eksperiment tager udgangspunkt heri.

I perioden 1986-1998 fungerede en fordelingspolitik for asylansøgere, som fik tildelt opholdstilladelse som flygtninge. Hovedformålet med fordelingspolitikken var at fordele flygtninge forholdsmæssigt ud i amter og kommuner efter indbyggertal. Fordelingspolitikken blev vedtaget som reaktion på det tidligere bosættelsesmønster, hvor flygtningene hovedsageligt bosatte sig i og omkring de større danske byer.

I den betragtede 13-årige periode blev i alt 73.036 flygtninge fordelt. Heraf er der for 70.753 flygtninge informationer om uddannelsesstatus, oprindelsesland og boligoplysninger. Visse flygtninge ankommer efter familiesammenføring med en herboende flygtning, hvilket typisk omfatter ægtefæller og børn. Der kan stilles spørgsmål ved, hvorvidt familiesammenførte boligplaceres tilfældigt, idet disse ofte bosætter sig ved det familiemedlem, som allerede opholder sig i landet. Familiesammenførte til flygtninge ekskluderes af denne grund af analysens datagrundlag, som herefter indeholder 55.314 flygtninge ankommet i perioden 1986-1998.

Fordelingen af flygtninge blev foretaget af Dansk Flygtningehjælp. I ansøgningen om asyl udfyldte asylansøgerne et spørgeskema, hvori der blev spurgt til alder, ægteskabelig status, antal børn, herkomst og sprog. Dette spørgeskema var tilgængeligt for Dansk Flygtningehjælp, og spørgeskemaet var i princippet bestemmende for placering af flygtningene. Der var i placeringsprocessen intet møde mellem Dansk Flygtningehjælp og flygtningene. Flygtningene blev opfordret til at blive i deres tildelte kommune i 18 måneder. Der var ingen restriktioner til hindring for, at de kunne flytte før ophør af denne periode.

I enkelte tilfælde afviste flygtninge at bo i den anviste bolig. Disse få flygtninge blev tildelt en anden lokation. Damm (2013) bemærker, at dette er et mindre problem, da de fleste af flygtningene, der blev genplaceret, først fik tildelt ny bolig efter, at de var flyttet ud til den først anviste bolig.

Damm (2005) undersøger det initiale bosættelsesmønster for flygtninge, der fik asyl i perioden 1986 til 1998. Formålet med undersøgelsen var at vurdere, om fordelingspolitikken kan udgøre grundlaget for et naturligt eksperiment. Undersøgelsen baserede sig både på registerdata og interviews med ansatte i Dansk Flygtningehjælp, der var med til at placere flygtningene. Undersøgelsen konkluderer, at fordelingspolitikken for flygtninge i perioden 1986 til 1998 resulterede i en tilfældig initial geografisk fordeling af flygtninge betinget på seks personspecifikke karakteristika: Sundhed (behov for speciel medicinsk eller psykiatrisk behandling), specielle uddannelsesbehov, afstand til familiemedlemmer, familiestatus (køn, par/enlig, antal børn), nationalitet og indvandringsår.

For at sikre et så randomiseret eksperiment som muligt er det vigtigt at tage højde for disse karakteristika i den efterfølgende analyse. Sundhedsstatus og ønsket afstand til familiemedlemmer er ikke umiddelbart observerbart i data.

Analyserne udført af Damm har haft et andet sigte end i denne rapport. I Damm (2009b) analyseres flygtnings valg af bolig, idet det analyseres, hvor flygtningene flytter hen, når de forlader den første bolig. I Damm (2009a) påvises positive netværkseffekter, idet det vises, at den årlige indkomst er positivt korreleret med tætheden af personer med samme oprindelsesland. I Damm og Rosholm (2010) opstilles og estimeres en søgeteoretisk arbejdsmarkedsmodel, og det vises, at spredningen af de ankomne flygtninge har en negativ effekt på de lokale arbejdsmarkeders funktionsmåde. I Damm (2014b) genbesøges netværksproblematikken, idet det påvises, at arbejdsløse mandlige flygtnings sandsynlighed for at finde beskæftigelse især er bestemt af beskæftigelsesgraden blandt lokale ikke-vestlige indvandrere og mænd med samme nationale baggrund. I Damm (2014a) vises det, at unge flygtnings kriminalitet især er påvirket af antallet af kriminelle i lokalområdet. Endelig vises det i Dustman, Vasiljeva og Damm (2016), at antallet af flygtninge, der anvises til en given kommune, påvirker den lokale befolknings tendens til at stemme til højre for midten.

5.2 Datagrundlag og balancering af data

Til undersøgelsen er det nødvendigt at identificere personer indvandret som flygtninge og familiesammenførte til flygtninge i perioden 1986 til 1998³⁰.

På baggrund af Danmarks Statistiks individdata er det muligt at identificere i alt 73.036 flygtninge og familiesammenførte til flygtninge, som ankom til Danmark i perioden 1986-1998. Nogle af disse opholder sig imidlertid meget kort tid i landet. For andre er det ikke muligt at identificere alle nødvendige baggrundskarakteristika. Som følge heraf ekskluderes 3,1 procent af den samlede flygtningebestand fra undersøgelsen. Datagrundlaget består herefter af i alt 70.753 flygtninge og familiesammenførte til flygtninge ankommet 1986-1998, for hvem alle nødvendige baggrundskarakteristika er tilgængelige.

Tilstrømningen af flygtninge var særligt stor midt i 1990'erne, hvor der i løbet af blot to år ankom næsten 30.000 flygtninge og familiesammenførte til flygtninge. I den øvrige del af perioden svinger antallet af flygtninge og familiesammenførte mellem 3.000-5.000 årligt ankomne. Det fremgår af Figur 5.1a, der viser antallet af flygtninge og familiesammenførte til flygtninge efter indvandringsår. Flygtningene inddeles i fem grupper efter oprindelsesland.

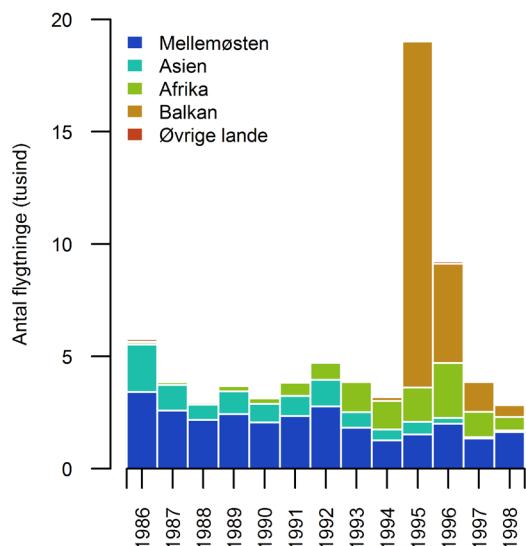
Flygtninge fra Mellemøsten kom hovedsageligt fra Irak, Iran og Libanon. De førstnævnte flygtede fra krigen mellem Iran og Irak og senere fra følgerne af Golfkrigen. Konflikten mellem israelere og palæstinensere øgede sammen med den libanesiske borgerkrig ligeledes flygtningestrømmen. Fra Asien kom der særligt i 1980'erne en større gruppe, bestående af vietnamesiske bådflygtninge og flygtninge fra de Tamilske Tigres oprør i Sri Lanka. Gennem 1990'erne øgedes antallet af flygtninge fra Afrika, hvor særligt borgerkrigen i Somalia resulterede i flygtninge til Danmark. I 1995-1996 skete der en voldsom stigning i antallet af nytilkomne flygtninge fra Balkan, da krigen i Bosnien sendte mange på flugt. Der var primært tale om flygtninge fra Bosnien-Hercegovina og Ex-Jugoslavien. Flygtninge fra øvrige verdensdele udgjorde relativt få personer. Gruppen indeholdt blandt andet statsløse og personer med uoplyst oprindelsesland.

³⁰ Data om opholdsgrundlag findes kun for personer indvandret fra og med 1997. For personer indvandret før 1997 vurderes en persons opholdsgrundlag ud fra historisk viden om indvandringsmønstre kombineret med viden om statsborgerskab ved indvandring og første indvandringsår, jf. Danmarks Statistik (2018).

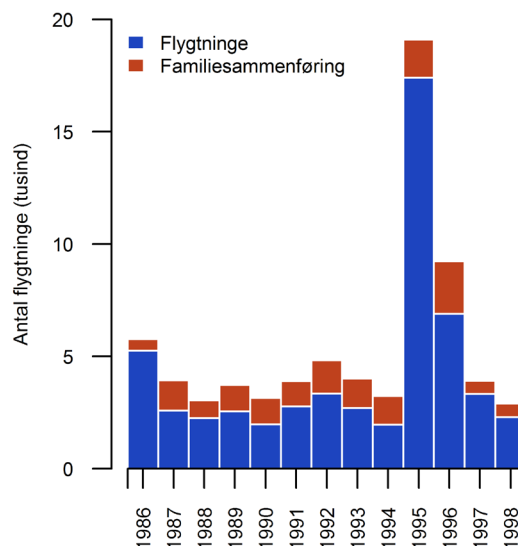
Omkring 22 procent af de betragtede flygtninge vurderes at være familiesammenført til herboende flygtninge³¹. Gennem den betragtede periode var antallet af familiesammenføringer typisk mellem 500-1.500 årligt, jf. Figur 5.1b. Familiesammenførte er typisk kvindelige partnere eller børn. Tabel 5.1 viser de ankomne flygtninge samt familiesammenførte til flygtninge efter familietype ved ankomst.

Figur 5.1. Flygtninge og familiesammenførte ankommet til Danmark, 1986-1998.

a) Antal efter oprindelsesland



b) Antal efter familiesammenføring eller flygtning



Anm.: Figuren viser samtlige 70.753 flygtninge og familiesammenførte ankommet mellem 1986-1998.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Familiesammenførte ekskluderes fra undersøgelsen således, at undersøgelsens datagrundlag omfatter i alt 55.314 flygtninge ankommet i perioden 1986-1998. Dette sker som konsekvens af, at det er tvivlsomt hvorvidt familiesammenførte boligplaceres tilfældigt, hvorfor det kan give anledning til endogenitetsproblemer at inddrage dem i analysen.

En relativt stor andel af flygtningene fra Balkan ankom som hele familier. Dette medfører, at relativt få personer familiesammenføres til disse, og familiesammenførte udgør blot fem procent af det samlede antal personer ankommet fra Balkan. For asiater ansås familiesammenførte at udgøre op mod 40 procent af de ankomne i den betragtede periode. For personer fra Mellemøsten og Afrika vurderes familiesammenførte at udgøre hver fjerde henholdsvis hver tredje ankommet i perioden 1986-1998.

Samlet set placeres hver femte af de ankomne flygtninge eksklusiv familiesammenførte i udsatte boligområder. Gennem den betragtede periode ses ingen systematik i forhold til, hvilke grupper af flygtninge, der blev placeret i udsatte boligområder. For alle grupperinger af oprindelseslande vil andelen af de nytilkomne flygtninge, der placeres i et udsat boligområde, typisk variere mellem 10 og 30 procent.

³¹ I den betragtede periode findes der ikke data for hvilke personer, som er familiesammenført. Opgørelsen baseres derfor på et skøn. En vurderet flygtning (voksen) antages at være familiesammenført, såfremt vedkommende bosætter sig med en partner, som allerede opholdt sig i landet. Dette gælder uanset partnerens status som flygtning/ikke-flygtning. Et barn – der vurderes at være flygtning – antages at være familiesammenført, hvis barnet bosætter sig i en eksisterende familie, hvor mindst en forældre allerede boede i landet, og hvor mindst en af forældrene er flygtning.

Tabel 5.1. Flygtninge efter opholdsgrundlag og familietype ved ankomst.

	1986-88	1989-90	1991-92	1993-94	1995-96	1997-98	Alle år
Antal i alt	12.738	6.888	8.725	7.247	28.336	6.819	70.753
Andel familiesammenført	21 %	34 %	30 %	36 %	14 %	18 %	22 %
<i>Vurderet flygtning</i>							
Enlig mand	4.301	1.457	1.695	1.582	3.619	1.938	14.592
Enlig kvinde	500	399	589	581	1.792	822	4.683
Par uden børn	627	280	517	329	3.178	608	5.539
Enlig mand med børn	68	19	38	13	70	18	226
Enlig kvinde med børn	196	122	183	161	624	189	1.475
Par med børn	1.756	828	1.140	726	7.111	863	12.424
Barn med forældre	2.566	1.382	1.907	1.245	7.843	1.083	16.026
Barn uledsaget	63	43	59	32	58	95	350
<i>Vurderet familiesammenført</i>							
Voksen mand	439	262	403	386	741	395	2.626
Voksen kvinde	899	953	1.273	1.355	1.675	225	6.380
Barn	1.323	1.143	921	837	1.625	583	6.432

Anm.: Persontallene i kategorierne "enlig" og "par" omfatter udelukkende voksne personer. Alle børn er indregnet i kategorierne "barn".

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Med henblik på at optimere flygtningenes initiale geografiske placering korrigeres data for placering i midlertidige boliger, jf. Damm (2009a). Personer, der inden for det første år efter indvandring er flyttet til en bolig i en anden kommune end den oprindelige, antages at være blevet placeret midlertidigt i den første bolig. De pågældende antages derfor i analysen at være blevet initialt placeret i den anden frem for den første bolig. Det kan ikke udelukkes, at personer flytter på eget initiativ inden for det første år, hvorfor der er en risiko for at korrektionen introducerer endogenitet i analysen. Eftersom de personer, der antages placeret i en midlertidig bolig ikke synes systematisk afgrænset af bestemte karakteristika, vurderes endogenitetsproblemet dog at være af begrænset omfang. Man kan endvidere hævde, at det kræver en betydelig indsigt at afsøge bosætningsmulighederne på tværs af kommunegrænser allerede inden for det første år efter indrejse, hvilket styrker formodningen om, at problemet med selvselektion er af relativt beskedent omfang.

Balancering af data

Vi betragter det anvendte datasæt som et naturligt eksperiment, da Dansk Flygtningehjælp havde et relativt overfladisk kendskab til den enkelte flygtning, og for eksempel ikke baserede deres beslutninger på en personlig kontakt. På denne måde får den enkelte flygtning personlige egenskaber ikke nogen indflydelse på geografisk placering, og vi kan tale om et quasi-randomiseret eksperiment. Som nævnt i afsnit 5.1 konkluderede Damm (2005), at den initiale geografiske fordeling af flygtninge var tilfældig betinget på seks personspecifikke karakteristika: Sundhed, specielle uddannelsesbehov, afstand til familiemedlemmer, familiestatus (køn, par/single, antal børn), nationalitet og indvandringsår.

For at kontrollere for de faktorer, vi ved, har haft indflydelse på placeringssted, foretages en såkaldt balancering af data. Først opdeles data i en indsatsgruppe (flygtninge placeret i udsat område) og en kontrolgruppe (resten). Grupperne sammenlignes i Tabel 5.2.

Tabel 5.2. Sammenligning af indsats- og kontrolgruppe, før og efter balancering.

	Matching	Gennemsnit		Ændring		var ratio (Tr/Co)	t-test (p-værdi)
		Indsats	Kontrol	Std. gns.	Gns. (%-point)		
Køn (mænd)	Før		59,2%	-4,30	2,1%	1,01	0,00
	Efter	57,0%	56,7%	0,69	0,3%	1,00	0,57
Alder*	Før		26,29	0,44	0,07	1,10	0,66
	Efter	26,37	26,25	0,67	0,12	1,07	0,57
Familietype (par)	Før		55,1%	7,43	3,7%	0,98	0,00
	Efter	58,8%	58,6%	0,48	0,2%	1,00	0,65
Antal børn*	Før		1,45	10,15	0,18	1,04	0,00
	Efter	1,63	1,65	-0,98	0,02	0,88	0,39
<i>Indvandringsår</i>							
1987	Før		4,8%	-4,11	0,8%	0,84	0,00
	Efter	4,0%	4,2%	-0,80	0,2%	0,96	0,51
1988	Før		4,2%	-3,55	0,7%	0,85	0,00
	Efter	3,6%	3,8%	-1,11	0,2%	0,95	0,37
1989	Før		4,7%	-1,99	0,4%	0,92	0,05
	Efter	4,3%	4,4%	-0,39	0,1%	0,98	0,75
1990	Før		3,6%	-1,68	0,3%	0,92	0,10
	Efter	3,3%	3,2%	0,83	0,1%	1,05	0,50
1991	Før		4,6%	7,15	1,7%	1,35	0,00
	Efter	6,4%	6,0%	1,55	0,4%	1,06	0,17
1992	Før		5,2%	13,34	3,8%	1,67	0,00
	Efter	9,0%	9,5%	-1,79	0,5%	0,95	0,08
1993	Før		4,5%	6,89	1,7%	1,34	0,00
	Efter	6,2%	6,0%	0,79	0,2%	1,03	0,49
1994	Før		3,1%	9,07	2,0%	1,61	0,00
	Efter	5,1%	5,5%	-1,61	0,4%	0,94	0,17
1995	Før		31,5%	-0,30	0,1%	1,00	0,76
	Efter	31,4%	30,7%	1,46	0,7%	1,01	0,19
1996	Før		13,4%	-13,55	4,0%	0,74	0,00
	Efter	9,4%	9,9%	-1,65	0,5%	0,96	0,13
1997	Før		6,1%	-1,11	0,3%	0,96	0,28
	Efter	5,8%	5,8%	0,24	0,1%	1,01	0,85
1998	Før		4,1%	0,36	0,1%	1,02	0,72
	Efter	4,2%	3,9%	1,42	0,3%	1,07	0,25
<i>Uddannelsesniveau</i>							
Grundskole	Før		72,9%	-2,53	1,1%	1,03	0,01
	Efter	71,7%	71,8%	-0,09	0,0%	1,00	0,94
Erhvervsfaglig	Før		11,1%	1,28	0,4%	1,03	0,20
	Efter	11,5%	11,3%	0,67	0,2%	1,02	0,57
Gymnasial	Før		6,4%	1,26	0,3%	1,05	0,21
	Efter	6,8%	6,9%	-0,35	0,1%	0,99	0,78
KVU	Før		2,7%	0,04	0,0%	1,00	0,97
	Efter	2,7%	2,9%	-0,82	0,1%	0,95	0,51
MVU	Før		4,0%	1,59	0,3%	1,08	0,11
	Efter	4,3%	4,3%	0,04	0,0%	1,00	0,97
LVU (inkl. Ph.d.)	Før		2,9%	0,49	0,1%	1,03	0,62
	Efter	3,0%	3,0%	0,23	0,0%	1,01	0,85
<i>Oprindelsesområde</i>							
Afrika	Før		12,3%	5,60	2,0%	1,13	0,00
	Efter	14,3%	14,2%	0,32	0,1%	1,01	0,79
Mellemøsten	Før		35,7%	11,76	5,8%	1,06	0,00
	Efter	41,5%	42,0%	-0,86	0,4%	1,00	0,44
Balkanhalvøen	Før		39,1%	-13,01	6,1%	0,93	0,00
	Efter	33,0%	32,9%	0,15	0,1%	1,00	0,88
Asien	Før		11,5%	-5,69	1,7%	0,87	0,00
	Efter	9,9%	9,6%	0,69	0,2%	1,02	0,55

Anm.: * Alder og antal børn opgøres i absolutte enheder. Gennemsnit er derfor ikke angivet i procent og den gennemsnitlige ændring er ikke i procentpoint.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Indsatsgruppen har i gennemsnit 0,2 børn flere end kontrolgruppen, jf. Tabel 5.2. Desuden består indsatsgruppen før balancering af flere kvinder, i højere grad af par, af personer med lidt højere uddannelsesniveau og af flere flygtninge fra Afrika og Mellemøsten, end det er tilfældet i kontrolgruppen.

For at korrigere for forskellene mellem indsatsgruppen og kontrolgruppen konstrueres en ny balanceret kontrolgruppe. I denne nye kontrolgruppe er fordelingen på køn, alder, ankomstår, uddannelse osv. den samme som i indsatsgruppen. Dette gør det muligt at udføre en direkte sammenligning af de to grupper. Tabel 5.2 sammenligner ligeledes indsatsgruppen med den balancerede kontrolgruppe. Det fremgår for eksempel, at det gennemsnitlige antal børn i begge grupper er lig 1,6.

Metode til balancering

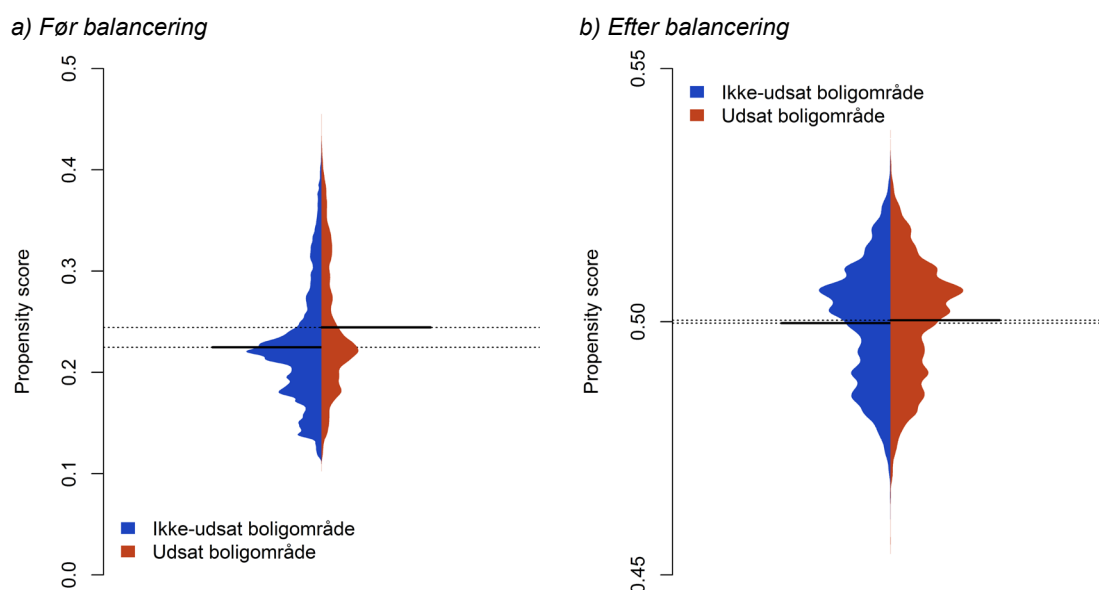
Balancering af kontrolgruppen sker ved metoden *propensity score matching*, jf. Rosenbaum og Rubin (1983). Metoden anvender to skridt: Først tildeles hver flygtning en individuel sandsynlighed for, at vedkommende placeres i et udsat boligområde. På baggrund heraf dannes en gruppe af flygtninge, som blev placeret i et ikke-udsat nabolag, der ud fra baggrundskarakteristika er identisk med indsatsgruppen.

I første skridt estimeres en betinget sandsynlighed (*propensity score*) for at blive placeret i et udsat boligområde givet baggrundskarakteristika. Hertil anvendes en *logit-model*. Den estimerede sandsynlighed for at blive placeret i et udsat boligområde er det mål, den efterfølgende identifikation af en kontrolgruppe magen til indsatsgruppen foretages ud fra. Sandsynligheden kan ligeledes findes med andre metoder som for eksempel *discriminant analysis*, *probit model*, *boosted regression* eller *generiske algoritmer*, jf. McCaffrey m.fl. (2004) og Sekhon (2011). Anvendelse af en logit-model er dog det mest almindelige. Det er ligeledes muligt at foretage balanceringen med baggrund i andre mål end en sandsynlighedsfordeling.

I appendiks F præsenteres estimationsresultaterne fra logit-modellen. Det fremgår, at ved estimation på ubalanceret data er stort set alle kovariaterne signifikant bestemt. Det vil sige, at der før balancering er signifikant forskel på sandsynligheden for at blive placeret i et udsat område for flygtninge initialt placeret i udsatte henholdsvis ikke-udsatte boligområder.

Fordelingen af den estimerede sandsynlighed for at blive placeret i et udsat boligområde er i Figur 5.2a vist for de flygtninge, som rent faktisk blev anvist til et udsat område samt for de flygtninge, der ikke blev (figuren kaldes *common support*). Det fremgår af figuren, at der er et stort overlap mellem sandsynlighedsfordelingerne. Desuden ses enhver flygtning at have en sandsynlighed større end 10 procent for at blive placeret i et udsat boligområde uanset vedkommendes faktiske placering. Herved er der god visuel støtte for at finde et match i kontrolgruppen til alle flygtninge anvist til et udsat boligområde.

Figur 5.2b viser *common support* efter matching. Af figuren ses det, at de to fordelinger efter balancering stort set er identiske.

Figur 5.2. Fordeling af sandsynlighed for placering i udsat boligområde (common support).

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

I andet skridt sammenkædes hver flygtning placeret i en udsat bolig med en person fra kontrolgruppen, som havde tilnærmelsesvis samme betingede sandsynlighed for at blive placeret i et udsat boligområde (*matching*).

Sammenkædningen sker ud fra den logit-model prædikterede sandsynlighed for placering i et udsat boligområde. Sandsynligheden anvendes i en matching-algoritme. Der findes flere forskellige algoritmer til dette formål, for eksempel *Nearest Neighbour Matching*, *Stratification*, *Caliper and Radius Matching* og *Weighting on Propensity Score*. Valget af matching-algoritme har en betydning for, hvor præcist matchet mellem indsatsgruppen og kontrolgruppen er, men det sker på bekostning af præcisionen af estimatet. Det vil sige, at valget af algoritme involverer en afvejning mellem bias og varians på estimatet, jf. Caliendo m.fl. (2005).

I denne analyse vælges *Nearest Neighbour* algoritmen med én-til-én matching uden tilbagelægning. Herved fås et balanceret datasæt med en kontrolgruppe og en indsatsgruppe, der er lige store, og som begge består af unikke personer. I appendiks F argumenteres der for valget af matching-algoritme.

Efter balancering af data er det lykkedes at mindske forskellen mellem kontrolgruppen og indsatsgruppen. Det fremgår af Tabel 5.2, hvoraf det for eksempel ses, at andelen af mænd i kontrolgruppen reduceres med 2,5 procentpoint til 56,7 procent, hvilket modsvarer, at mænd udgør 57,0 procent af indsatsgruppen.

For alle øvrige baggrundskarakteristika ses indsats- og kontrolgruppen ligeledes at være tæt ved identiske. Dette ses for eksempel ved at betragte variansen i indsatsgruppen sat i forhold til variansen i kontrolgruppen (*var ratio*), som skal være lig én, hvis der er perfekt balance. Efter balancering af data ses forholdet klart mindsket og tæt ved én i langt de fleste tilfælde.

Endelig udføres en Welch's t-test, som undersøger, hvorvidt middelværdierne i indsats- og kontrolgruppen kan antages at være identiske. Hvis p-værdien er større end et på forhånd fastsat signifikansniveau (typisk 0,05), accepteres nulhypotesen, og grupperne forudsættes at have samme middelværdi. Af yderste højre kolonne i Tabel 5.2 ses det, at nulhypotesen efter balancering accepteres for alle karakteristika. Den mindste p-værdi efter balancering er 0,08.

Efter matching vurderes data at være balanceret og derfor en god approksimation for et randomiseret eksperiment, jf. Tabel 5.2 og Figur 5.2.

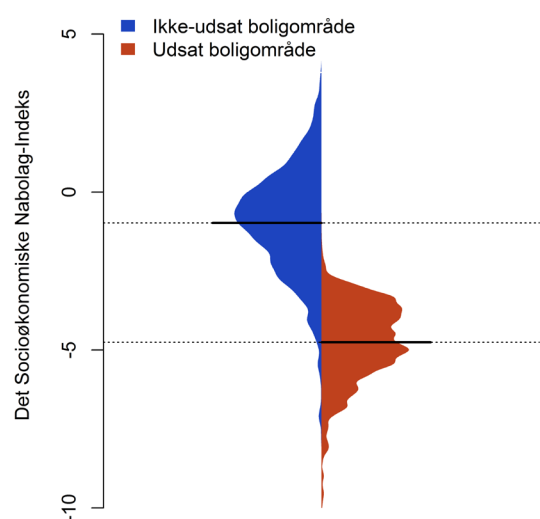
Sammenligning af indsats- og kontrolgruppe efter balancering

Indsatsgruppen er per definition placeret i et udsat boligområde på indvandringstidspunktet. Dette afspejles i fordelingen af det Socioøkonomiske Nabolag-Indeks (defineret i kapitel 2), som for indsatsgruppen i gennemsnit er tæt ved -5, jf. Figur 5.3a. For kontrolgruppen, som ved tilflytning blev placeret i et ikke-udsat boligområde, ses indekset i gennemsnit at være tæt ved nul, dvs. kontrolgruppen placeres i områder, som blot er lidt mere udsatte end landet som helhed³².

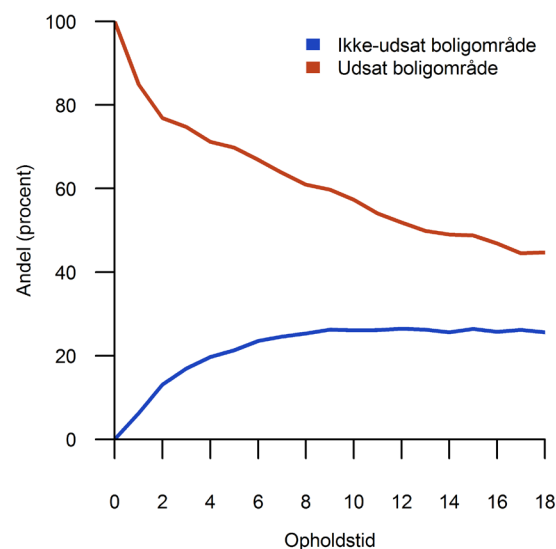
Over tid flytter flygtningene mellem områder, ligesom nabolag kan ændre status og overgå fra udsat til ikke-udsat (eller omvendt). Blandt flygtninge i indsatsgruppen aftager andelen bosat i udsatte områder således, at det efter 12 års ophold i Danmark blot er halvdelen, som fortsat bor i et udsat boligområde, jf. Figur 5.3b. Modsat stiger andelen af kontrolgruppen, som er bosat i et udsat område.

Figur 5.3. Placering af indsats- og kontrolgruppe i udsatte og ikke-udsatte boligområder.

a) Fordeling af nabolagsindeks ved indvandring



b) Andel i udsat boligområde efter opholdstid

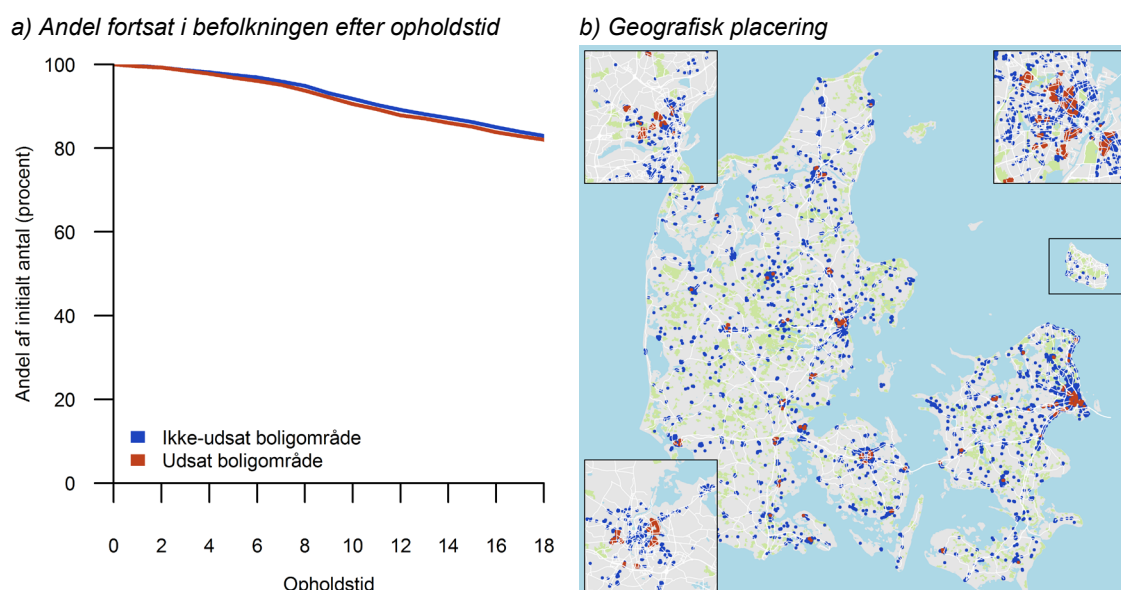


Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Personer udtræder af befolkningen ved dødsfald eller udvandring. Frafald fra indsats- og kontrolgruppen er stort set ens over tid, jf. Figur 5.4a. Efter 18 års ophold i Danmark er cirka 18 procent af den oprindelige flygtningebestand udvandret eller døde.

Udsatte boligområder er beliggende i de større byer, hvorfor indsatsgruppen ved indrejse er placeret i hovedsageligt København, Københavns Vestegn, Nordsjælland, Holbæk, Slagelse, Nykøbing Falster, Nakskov, Nyborg, Svendborg, Odense, Haderslev, Sønderborg, Esbjerg, Fredericia, Horsens, Kolding, Vejle, Herning, Randers, Viborg og Aalborg, jf. Figur 5.4b. Flygtninge i kontrolgruppen er derimod nogenlunde jævnt fordelt over de resterende dele af landet.

³² En mindre andel af kontrolgruppen bor i nabolag, der er socioøkonomisk udsatte på niveau med de mindst udsatte i indsatsgruppen. Disse områder er ikke defineret som udsatte, da områdernes størrelse ikke er tilstrækkelig stor.

Figur 5.4. Frafaldstilbøjelighed og geografisk placering af indsats- og kontrolgruppe.

Anm.: Af hensyn til datadiskretionskrav er flygtninge placeret i områder med få andre flygtninge udeladt af figur b), hvorfor enkelte markeringer i tyndt befolkede områder ikke er vist.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

5.3 Effekt af placering i udsat boligområde

I dette afsnit vurderes det, hvorvidt flygtninge placeret i udsatte boligområder (indsatsgruppen) har haft en anden adfærd end flygtninge placeret i ikke-udsatte boligområder. Der fokuseres på arbejdsmarkedsdeltagelse og kriminalitetsomfang.

Flygtninge der ankom i perioden 1986-1998 blev tildelt en bolig ud fra en fordelingspolitik, der havde til hensigt at placere dem spredt ud over landet. Som følge heraf blev flygtningene tilfældigt anvist til udsatte eller ikke-udsatte boligområder. På trods af tilfældig placering ses mindre forskelle mellem de to grupper. Derfor foretages balancering af data således, at der dannes en gruppe af personer placeret i ikke-udsatte boligområder, der på baggrundskarakteristika matcher gruppen anvist til udsatte boligområder, jf. afsnit 5.1 (gruppen kaldes kontrolgruppen).

Efter balancering af data haves en approksimation for et randomiseret eksperiment, hvor idéen er, at placeringen i enten et udsat eller i et ikke-udsat boligområde er den eneste væsentlige forskel mellem de to grupper af flygtninge.

For at vurdere, hvorvidt de to grupper af flygtninge har forskellig adfærd, anvendes en statistisk model, der tester, om der er statistisk signifikant forskel på erhvervsdeltagelse og kriminalitetsomfang i de to grupper. Den statistiske model beskrives nærmere i det følgende delafsnit.

Målt på erhvervsfrekvens – dvs. beskæftigede og ledige i alderen 15 til 64 år sat i forhold til den samlede befolkning i samme aldersgruppe – findes en signifikant negativ effekt af at være blevet placeret i et udsat boligområde i forhold til flygtninge anvist til ikke-udsatte boligområder. Der findes en større negativ effekt for kvinder end for mænd. For begge køn mindskes effekten ved længere ophold i landet, og for mænd kan der ikke findes nogen signifikant effekt efter 12 års ophold. Placeringen i udsatte boligområder har ikke en signifikant effekt for flygtninge, der ankom som børn, hvorimod der findes en negativ signifikant negativ effekt for flygtninge der ankom som voksne. Der findes ikke en signifikant forskel på erhvervsdeltagelsen for flygtninge, som kom til landet som uaglærte og faglærte.

Andelen af dømte er ikke signifikant forskelligt for flygtninge placeret i et udsat boligområde og kontrolgruppen. Der er således ikke statistisk belæg for at sige, at placeringen i et udsat boligområde har påvirket omfanget af kriminalitet. Der findes heller ingen signifikant effekt ved opsplitning efter køn eller opdeling efter, om man er barn eller voksen på indrejsetidspunktet.

Statistisk regressionsmodel

Der opstilles en statistisk model, som estimerer den kausale effekt af at blive placeret i et udsat boligområde som flygtning. Med kausalitet (eller årsagssammenhæng) menes, at placering i forskellige boligområder forårsager den socioøkonomiske effekt. Dette sikres, da den statistiske model estimeres på data, hvor flygtningene er tilfældigt placeret i de to grupper (data er en approksimation for et randomiseret eksperiment).

Den anvendte statistiske model undersøger, om flygtnings erhvervsfrekvens og kriminalitetsomfang er forskelligt i de to grupper, som initialt blev placeret i henholdsvis et udsat og et ikke-udsat boligområde. Modellen er nærmere beskrevet i Boks 5.1.

Boks 5.1. Regressionsmodel.

For hver flygtning i med opholdstid t ønskes det at bestemme, hvorvidt placeringen i et udsat boligområde har haft effekt på vedkommendes erhvervsfrekvens eller kriminalitetsomfang, $Y_{i,t}$. Dette undersøges ved følgende regressionsligning:

$$Y_{i,t} = \alpha_t + \beta_t T_i + \epsilon_{i,t}$$

hvor α_t er en konstant, T_i er en indikator for indsats, dvs. om personen initialt er allokeret i et udsat område (lig én) eller er i kontrolgruppen (lig nul), β_t er en koefficient hørende til indikatoren, og $\epsilon_{i,t}$ er et normalfordelt fejllid.

Det forventede udfald for en flygtning med opholdstid t placeret i et udsat område (en person i indsatsgruppen) er $E[Y_{i,t} | T_i = 1] = \alpha_t + \beta_t$, mens det forventede udfald for en flygtning i kontrolgruppen er $E[Y_{i,t} | T_i = 0] = \alpha_t$. Altså kan β_t tolkes som den gennemsnitlige effekt for en flygtning med opholdstid t ved at være placeret i et udsat boligområde i forhold til at være i kontrolgruppen.

Bemærk, at regressionsligningen ovenfor ikke indeholder individuelle karakteristika. Disse udelades, da data er balanceret³³. Regressionsmodellen kvantificeres med udgangspunkt i Rubins kausale model jf. appendiks G.

Kilde: Egen tilblivelse, se i øvrigt appendiks G.

Forskelle undersøges på kort sigt (mindre end 6 års ophold i Danmark), på mellemlangt sigt (6 til 12 års ophold) og langt sigt (mere end 12 års ophold). En eventuel forskel på adfærden mellem de to grupper af flygtninge placeret i henholdsvis udsatte og ikke-udsatte boligområder kaldes *gennemsnitsindsatseffekten* (*average treatment effect, ATE*).

³³ Regressionsligningen kan også opstilles i et "long" format: $Y_{i,t} = \alpha_t + \beta_t T_i + X_i' \gamma_t + \epsilon_{i,t}$, hvor X_i er en vektor af de individuelle karakteristika (kovariaterne), som data blev balanceret med hensyn til jf. afsnit 5.1, og γ_t er en vektor bestående af koefficienter hørende til hvert karakteristika. Da X_i for det balancerede data er ukorreleret med indsatsindikatoren T_i , har de ingen effekt på estimatet af β_t . Men inklusionen af baggrundskarakteristika kan give et meget præcist estimat af indsatseffekten, da de reducerer residualvariansen, som kan give en mindre standardfejl på β_t . Modellen er ligeledes estimeret i "long" format, hvilket giver marginale ændringer i de estimerede standardfejl. Resultaterne præsenteret i rapporten er ikke fra "long" formatet.

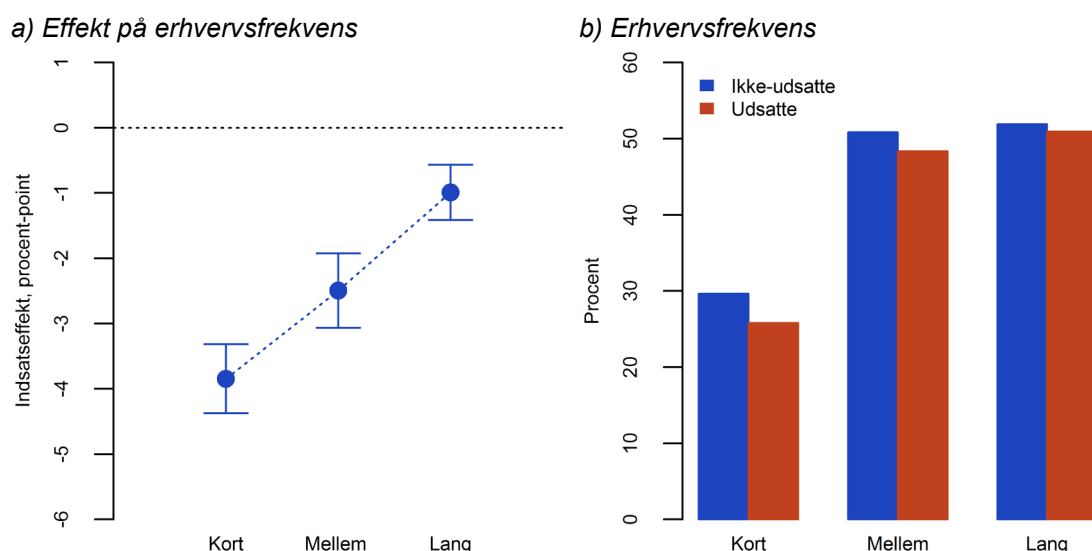
Indsatseffekten udtrykker den gennemsnitlige effekt ved at være blevet placeret i udsatte boligområder i forhold til placering i ikke-udsatte boligområder. For at lette forståelsen omtales gennemsnitsindsatseffekten i den resterende del af kapitlet typisk blot som "effekten".

Effekt af placering i udsat boligområde på erhvervsfrekvens

Erhvervsfrekvensen for flygtninge er stigende i opholdstid således, at erhvervsdeltagelsen er relativt lav i de første år efter ankomst til Danmark, hvorefter erhvervsfrekvensen over tid stiger således, at omkring halvdelen deltager i arbejdsstyrken, jf. Figur 5.5b. Dette gælder både flygtninge placeret i udsatte samt ikke-udsatte boligområder. Til sammenligning er den gennemsnitlige erhvervsfrekvens i hele Danmark i 2017 på 75 procent (15-64 årige), og i et udsat boligområde er det tilsvarende tal 54 procent. Desuden ses der i Figur 5.5b en tendens til, at erhvervsdeltagelsen konsekvent er lavere for flygtninge placeret i udsatte boligområder end for flygtninge anvist til ikke-udsatte områder.

Erhvervsfrekvensen for flygtninge placeret i udsatte områder er på både kort, mellemlangt og langt sigt signifikant lavere end erhvervsfrekvensen for personer i kontrolgruppen, jf. Figur 5.5a. Effekten er målt for flygtninge i den erhvervsduelige alder (15-64 år), og figuren viser effekten for flygtninge med kort, mellemlang og lang opholdstid i Danmark.

Figur 5.5. Effekt af placering i udsat boligområde på erhvervsfrekvens.



Anm.: Konfidensbånd er indtegnet omkring estimatet for indsatseffekten og viser, om estimatet er signifikant på et 5 procent signifikansniveau. Estimatet er signifikant forskelligt fra nul, hvis konfidensbåndet ikke overlapper andenaksen ved værdien nul (den stiplede linje).

På førsteaksen vises opholdstid, som er inddelt i tre grupper: Kort sigt (0-6 års ophold), mellemlangt sigt (6-12 års ophold) og langt sigt (mindst 12 års ophold).

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

På kort sigt er erhvervsfrekvensen 3,4 procentpoint lavere for flygtninge anvist til udsatte boligområder end for kontrolgruppen. Det ses, at forskellen på erhvervsfrekvensen mellem de to grupper af flygtninge mindskes, jo længere tid flygtningene har opholdt sig i landet. På mellemlangt og langt sigt er erhvervsfrekvensen henholdsvis 2,5 og 1,0 procentpoint lavere for flygtninge placeret i et udsat boligområde end for flygtninge i kontrolgruppen.

Det ses herved, at udsatte boligområder har en signifikant negativ effekt på flygtningenes erhvervsfrekvens uafhængigt af opholdstid. Det ses også, at størrelsen af effekten mindskes, jo længere tid flygtningene har opholdt sig i Danmark.

For at undersøge effekten af placering i udsatte boligområder for delgrupper af den samlede gruppe af flygtninge opdeles den samlede effekt på erhvervsdeltagelsen efter alder, køn, oprindelsesland og uddannelsesniveau på indvandringstidspunktet. Desuden betragtes effekten på erhvervsindkomst for dem, der er i arbejdsstyrken. Opdelingen ses i Tabel 5.3 og resultaterne uddybes i appendiks H.1.

Målt på erhvervsfrekvens findes en negativ effekt af placering i udsatte boligområder for begge køn. Der findes en større negativ effekt for kvinder end for mænd. For begge køn mindskes effekten ved længere tids ophold i landet, og for mænd kan der ikke findes nogen signifikant effekt efter 12 års ophold.

Placeringen i udsatte boligområder har større negativ effekt på erhvervsdeltagelsen for flygtninge kommet hertil som voksne, mens der ikke findes en signifikant effekt for flygtninge, der var yngre end 16 år ved ankomst. Der findes en negativ effekt af placering i udsatte boligområder uanset uddannelsesniveau, dog med en svag tendens til, at placeringen i udsatte boligområder har større negativ effekt på erhvervsdeltagelsen for faglærte personer.

Erhvervsdeltagelsen er i gennemsnit meget forskellig for flygtninge med forskelligt oprindelsesland. Generelt har flygtninge fra Afrika og Mellemøsten lavere erhvervsfrekvens end flygtninge fra Asien og Balkanhalvøen. Der findes en negativ effekt for alle grupper uanset oprindelsesland. Effekten er dog meget forskellig efter opholdstid. Det findes, at flygtninge fra Balkanhalvøen og Asien forklarer den negative totale effekt på kort sigt, mens der på langt sigt ikke findes nogen signifikant effekt. Effekten for flygtninge fra Afrika og Mellemøsten er mere stabil over tid, hvorfor flygtninge fra Afrika og Mellemøsten forklarer den negative totale effekt på langt sigt.

Tabel 5.3. Betinget effekt af placering i udsat boligområde på erhvervsfrekvens.

	Kort sigt	Mellemlangt sigt	Langt sigt	
Total gennemsnitlig indsats effekt (<i>procentpoint</i>)	-3,38 (0,27) **	-2,50 (0,29) **	-0,99 (0,22) **	
Betinget gennemsnitlig indsats effekt (<i>procentpoint</i>)				
Mænd	-3,68 (0,37) **	-2,23 (0,38) **	-0,22 (0,78)	
Kvinder	-4,21 (0,38) **	-3,05 (0,45) **	-2,13 (0,34) **	
Afrika	-2,52 (0,62) **	-3,06 (0,83) **	-3,06 (0,70) **	
Asien	-6,52 (0,89) **	-4,56 (0,90) **	0,22 (0,58)	
Balkanhalvøen	-6,09 (0,50) **	-2,07 (0,49) *	0,26 (0,41)	
Mellemøsten	-2,08 (0,38) **	-2,36 (0,44) **	-1,41 (0,30) **	
Voksen ved indvandring	-4,08 (0,28) **	-3,06 (0,33) **	-1,56 (0,26) **	
Barn ved indvandring	-1,28 (1,00)	-0,46 (0,63)	-0,19 (0,38)	
Ufaglært	-3,99 (0,36) **	-2,95 (0,44) **	-1,81 (0,34) **	
Faglært	-4,31 (0,44) **	-3,32 (0,50) **	-1,22 (0,40) **	
Erhvervsindkomst (<i>procentvis ændring</i>)	-2,80 (1,30) **	-1,03 (0,88)	-2,17 (0,63) **	

Anm.: Signifikansniveau: ***' 0,01 og **' 0,05.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Effekt af placering i udsat boligområde på kriminalitetsomfang

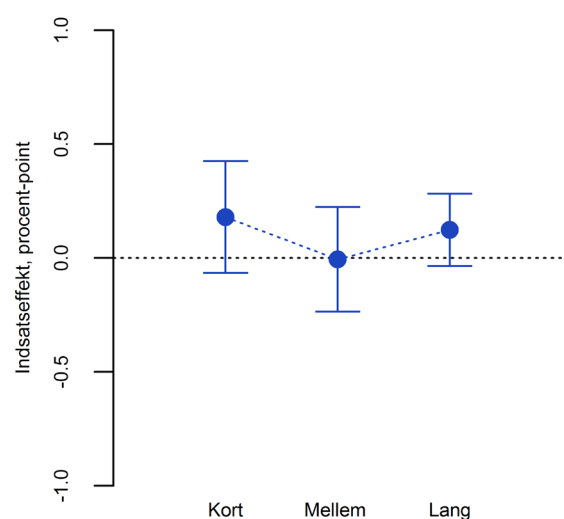
På tilsvarende vis måles effekten i forhold til, om flygtninge placeret i udsatte boligområder er mere kriminelle end flygtninge i kontrolgruppen. Omfanget af kriminalitet måles ud fra andelen af flygtninge i gruppen, som inden for det seneste kalenderår er dømt for overtrædelse af straffeloven, våbenloven eller loven om euforiserende stoffer.

Andelen af dømte varierer kun i mindre grad efter opholdstid i Danmark, jf. Tabel 5.6b. Dog ses en tendens til en svagt faldende andel af dømte efter lang opholdstid. Dette kan blandt andet hænge sammen med, at alderssammensætningen i grupperne ændres over tid, og at omfanget af kriminalitet generelt er lavere med alderen.

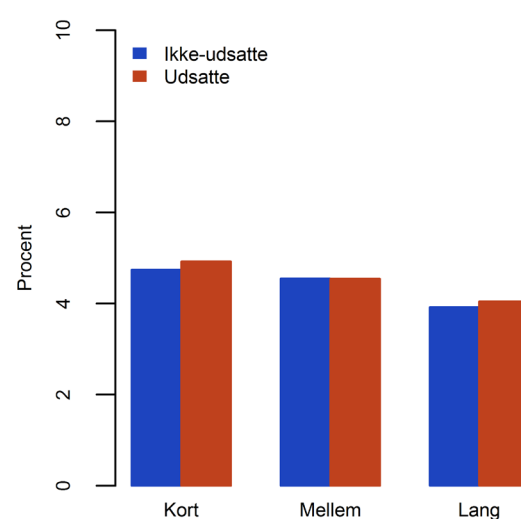
Andelen af dømte er ikke signifikant forskellig for flygtninge placeret i et udsat boligområde og kontrolgruppen, jf. Tabel 5.6a. Der er således ikke statistisk belæg for at sige, at placeringen i et udsat boligområde har påvirket omfanget af kriminalitet blandt personer i indsatsgruppen.

Figur 5.6. Effekt af placering i udsat boligområde på kriminalitetsomfang.

a) Effekt på andel dømte



b) Andel dømte



Anm.: Konfidensbånd er indtegnet omkring estimeret for indsats effekten. Se i øvrigt anm. til Figur 5.5.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

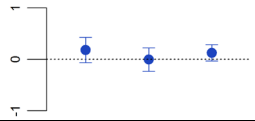
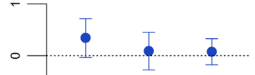
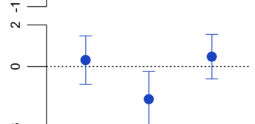
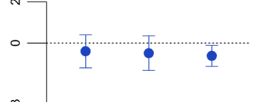
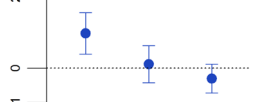
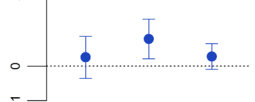
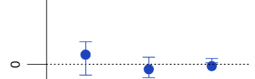
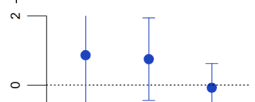
Effekten på andelen af dømte af placering i et udsat boligområde opdeles efter køn, oprindelsesland og om flygtningen indvandrede som voksen eller barn. Opdelingen ses for mænd i Tabel 5.4 og for kvinder i Tabel 5.5.

Det overordnede resultat er, at der ikke findes betingede effekter af placering i et udsat boligområde, som er signifikant forskellige fra nul. Der findes få tilfælde, hvor flygtningedelgruppen oplever en signifikant effekt efter et bestemt antal års ophold i Danmark. Et eksempel herpå er mandlige flygtninge fra Balkanhalvøen, der initialt blev placeret i et udsat boligområde i gennemsnit er mere kriminelle end mandlige flygtninge fra Balkanhalvøen placeret i et ikke-udsat boligområde.

Resultatet fra analysen peger herved på, at der stort set ikke kan findes en effekt på kriminalitetsomfanget af, at flygtninge bliver placeret i et udsat boligområde uanset køn, oprindelsesområde og alder ved ankomst. Det er dog vigtigt at understrege, at det kun er effekten, der ikke er forskellig. Kriminalitetsniveauet varierer meget mellem forskellige karakteristika. For eksempel er mænd i gennemsnit mere kriminelle end kvinder. Andelen af

dømte mænd er ca. 4 gange højere end kriminalitetsomfanget for kvinder, hvilket til sammenligning stort set er det samme som i den samlede befolkning og for udsatte boligområder i 2017.

Tabel 5.4. Betinget effekt af placering i udsat boligområde på kriminalitetsomfang, mænd.

	Kort sigt	Mellemlangt sigt	Langt sigt	
Total gennemsnitlig indsats effekt (<i>procentpoint</i>)	0,18 (0,13)	-0,01 (0,12)	0,12 (0,08)	
Betinget gennemsnitlig indsats effekt (<i>procentpoint</i>)				
Mænd	0,34 (0,19)	0,09 (0,19)	0,08 (0,13)	
Afrika, mænd	0,30 (0,60)	-1,59 (0,69) *	0,47 (0,86)	
Asien, mænd	-0,41 (0,40)	-0,49 (0,40)	-0,63 (0,23) *	
Balkanhalvøen, mænd	1,01 (0,31) **	0,11 (0,28)	-0,31 (0,21)	
Mellemøsten, mænd	0,25 (0,31)	0,79 (0,29) **	0,27 (0,19)	
Voksen ved indvandring, mænd	0,28 (0,20)	-0,14 (0,18)	-0,05 (0,11)	
Barn ved indvandring, mænd	0,86 (0,89)	0,75 (0,61)	-0,08 (0,36)	

Anm.: Signifikansniveau: *** 0,01 og ** 0,05.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Tabel 5.5. Betinget effekt af placering i udsat boligområde på kriminalitetsomfang, kvinder.

	Kort sigt	Mellemlangt sigt	Langt sigt
Total gennemsnitlig indsats effekt (procentpoint)	0,18 (0,13)	-0,01 (0,12)	0,12 (0,08)
<hr/>			
Betinget gennemsnitlig indsats effekt (procentpoint)			
Kvinder	-0,06 (0,35)	-0,17 (0,35)	0,19 (0,07) *
Afrika, kvinder	0,53 (0,35)	-0,02 (0,35)	0,14 (0,26)
Asien, kvinder	-0,91 (0,38)	-0,68 (0,37)	0,30 (0,23)
Balkanhalvøen, kvinder	0,09 (0,18)	-0,11 (0,15)	0,14 (0,12)
Mellemøsten, kvinder	-0,06 (0,26)	-0,15 (0,19)	0,20 (0,10)
Voksen ved indvandring, kvinder	-0,12 (0,14)	-0,22 (0,11)	-0,11 (0,08)
Barn ved indvandring, kvinder	0,61 (0,49)	0,02 (0,28)	0,41 (0,15) **

Anm.: Signifikansniveau: *** 0,01 og ** 0,05.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

5.4 Generalisering af resultatet af indsatsanalysen til andre befolkningsgrupper

Konklusionerne knyttet til analysen af indsats effekten ovenfor er i udgangspunktet begrænset til den befolkningsgruppe, der er eksponeret for det naturlige eksperiment, dvs. flygtninge eksklusiv familiesammenførte indvandret i perioden 1986 til 1998. I policy-sammenhæng er det imidlertid nyttigt at vide, hvorvidt resultaterne for analysegruppen kan generaliseres til en større befolkningsgruppe. Nærværende afsnit søger at afdække i hvilken udstrækning den gennemsnitlige effekt på erhvervsdeltagelsen og kriminalitetsomfanget kan generaliseres til en målgruppe bestående af både flygtninge og familiesammenførte til flygtninge indvandret i perioden 1986 til 1998. I forlængelse heraf afsøges endvidere, hvorvidt effekten kan generaliseres til henholdsvis en målgruppe bestående af alle flygtninge og familiesammenførte indvandret i 1986 eller senere, henholdsvis en målgruppe omfattende alle ikke-vestlige indvandrere indrejst i samme periode.

Det viser sig, at de gennemsnitlige effekter kan generaliseres til alle flygtninge og familiesammenførte uafhængig af indvandringsperiode, men ikke til ikke-vestlige indvandrere.

Generaliseringsspørgsmålet er adresseret inden for rammerne af en række centrale antagelser, hvorfor resultaterne skal fortolkes med tanke for disse forhold, jf. appendiks I.

Generalisering

Spørgsmålet om generalisering af resultater knyttet til randomiserede eksperimenter forekommer hyppigt inden for sundheds- og samfundsvidenskaben og er eksempelvis relevante i forbindelse med test af effekten af medicinsk behandling eller af implementering af et uddannelsesprogram, jf. Cole & Stuart (2010) og Stuart et. al (2015). Allerede når sådanne indsatseksperimenter designes sker det typisk med tanke på en senere generalisering og derfor under overvejelser om, at analysegruppens sammensætning på karakteristika så vidt muligt skal være repræsentativ for en større målgruppe.

I forbindelse med undersøgelse af effekten af boligplacering er analysegruppen, der består af indsats- og kontrolgruppen, ikke på forhånd afgrænset med tanke på at være repræsentativ, og boligplaceringen er ikke udført med tanke på indsatsevaluering. Undersøgelse af generaliseringsspørgsmålet er dermed ikke oplagt for et naturligt eksperiment i sammenligning med et kontrolleret eksperiment. Eftersom der i dag er et betydeligt fokus på de samfundsøkonomiske konsekvenser af befolkningens bosætningsmønstre, er det imidlertid oplagt at undersøge, om boligplacering ikke blot har betydning for analysegruppens erhvervsdeltagelse og kriminalitetsniveau, men også kan have signifikant indflydelse på de samme forhold for en større målgruppe. I fald dette kan etableres, kan resultatet opnå praktisk anvendelse og potentielt medvirke til at udstikke retningslinjerne for boligplacering i landets kommuner.

For at resultaterne er generaliserbare fordrer det, at analysegruppen har en sammensætning på personspecifikke karakteristika, der er repræsentativ for målgruppen. I praksis sikres overensstemmelse med de karakteristika, der er anvendt til balanceringen af det naturlige eksperiment i afsnit 2.1 og således ikke på karakteristika, der er underlagt endogenitetsproblemer. Er sammensætningen på disse karakteristika signifikant forskellig mellem analyse- og målgruppen, kan observationerne i analysegruppen vægtes, så strukturen i målgruppen i højere udstrækning gengives.

Opnås ensartethed med målgruppen er det meningsfuldt at estimere indsatseffekten på den vægtede analysegruppe. Det estimeres således, hvad den gennemsnitlige indsatseffekt af boligplacering i et udsat område ville være for målgruppen, når der tages højde for forskelle i de observerede karakteristika mellem analysegruppen og målgruppen. I forlængelse heraf er det vigtigt at bemærke, at tilgangen alene belyser den gennemsnitlige effekt for målgruppen, hvorfor indsatsen ikke nødvendigvis vil have tilsvarende effekt for delpopulationer i målgruppen afgrænset ved eksempelvis køn eller oprindelsesland. Som nævnt ovenfor er heterogenitet i effekten allerede påvist at være gældende for den oprindelige analysegruppe, jf. de betingede indsatseffekter.

I appendiks I beskrives, hvorledes *propensity scores* kan anvendes til beregning af vægte til brug for korrektion af kovariatstrukturen i analysegruppen og dermed til klarlægning af spørgsmålet om generalisering.

Målgrupperne

Det ønskes her at afdække, hvorvidt den gennemsnitlige indsatseffekt på erhvervsdeltagelsen og kriminalitetsomfanget kan generaliseres til tre forskellige målgrupper, der alle omfatter den oprindelige analysegruppe. Der betragtes to målgrupper bestående af flygtninge og familiesammenførte, hvor forskellen alene består i, at indvandringsåret for den ene målgruppe

er sammenfaldende med den periode, der afgrænsede personerne i analysegruppen, dvs. 1986-98, mens indvandringsåret for den anden målgruppe omfatter alle år fra og med 1986. Den tredje målgruppe omfatter alle ikke-vestlige indvandrere, der er indrejst i Danmark i 1986 eller senere. For samtlige målgrupper gælder, at deltagelse i grupperne ikke er betinget af, hvorvidt individerne på indvandringstidspunktet bosatte sig inden eller uden for et udsat område.

Forud for indsatsanalysen sigtes i forlængelse af ovenstående på at skabe ensartethed i kovariatstrukturen mellem analysegruppen og de respektive målgrupper gennem beregning af vægte betinget af udvalgte variable. Eftersom to af målgrupperne omfatter personer indvandret efter 1998 og dermed adskiller sig fra analysegruppen på dette område, indgår indvandringsåret ikke i modellen, der beregner de anvendte vægte, jf. appendiks I. Dette betyder, at observationer i analysegruppen med identiske kovariater vægtes på samme måde uafhængigt af indvandringsår, når den vægtede analysegruppe beregnes.

For hver målgruppe estimeres en deltagelsesmodel og den oprindelige analysegruppe vægtes efterfølgende med de inverterede propensity scores. Indsatsanalysen gennemføres kun såfremt det statistisk kan valideres, at kovariatstrukturen i den skalerede analysegruppe er i overensstemmelse med strukturen i målgruppen. Nærmere bestemt klarlægges det om forskellen i middelværdierne mellem hvert karakteristika i analyse- og målgruppen er insignifikant³⁴ såvel før som efter vægtning. I det følgende klarlægges dette for hver af de tre målgrupper, hvorefter indsatsanalysen gennemføres for de relevante tilfælde.

Flygtninge og familiesammenførte indvandret i perioden 1986-98

Gruppen omfatter 70.753 unikke personer, der har opholdt sig i Danmark i kortere eller længere tid i den betragtede periode. Heraf er 25.356 personer flygtninge ankommet i perioden 1986-1998 og udvalgt i indsatsanalysen i afsnit 5.3.

I Tabel 5.6 ses opsummerende karakteristika for målgruppen og for analysegruppen før og efter vægtning. Analysegruppens gennemsnitlige uddannelsesstruktur er før vægtning ikke signifikant forskellig fra strukturen i målgruppen. Ligeledes er der ensartethed i det gennemsnitlige antal børn. For de øvrige karakteristika er forskellen signifikant.

En mindre andel af analysegruppen stammer fra Asien og Afrika relativt til målgruppen, mens flere er fra Balkan og Mellempøsten. Analysegruppen er endvidere lidt ældre og omfatter flere enlige end målgruppen.

Det ses, at vægtningen afstedkommer ensartethed for flere beskrivende karakteristika mellem analysegruppen og målgruppen, idet forskellen mellem middelværdierne generelt mindskes. Trods den mindre forskel mellem de gennemsnitlige værdier er der dog fortsat forskelle, der vurderes signifikante eller lige på grænsen til at være det. Det kan endvidere bemærkes, at forskellen i det gennemsnitlige antal børn nu ikke længere er insignifikant.

Overordnet vurderes det dog, at den vægtede analysegruppe har en struktur, der er så ensartet med målgruppen, så det giver mening, at spørgsmålet om generalisering kan besvares gennem en indsatsanalyse på det vægtede grundlag.

³⁴ Hvis forskellen er insignifikant, godkendes nulhypotesen om, at forskellen ikke er signifikant forskellig fra nul.

Tabel 5.6. Sammenligning af analyse- og målgruppe bestående af alle flygtninge og familiesammenførte indvandret i perioden 1986 – 98, før og efter vægtning

	Gennemsnitlig andel			Forskel til målgruppe		z/t-test p-værdi*	
	Analyse	Analyse vægtet	Mål	Analyse	Analyse vægtet	Analyse	Analyse vægtet
Mænd	0,57	0,55	0,54	0,02	0,01	0,00	0,01
Alder*	26,31	25,17	25,09	1,22	0,08	0,00	0,36
Par	0,59	0,64	0,65	-0,06	-0,01	0,00	0,00
Antal børn*	1,64	1,65	1,63	0,01	0,02	0,37	0,01
Erhvervsfaglig	0,11	0,11	0,11	0,00	0,00	0,22	0,31
Gymnasial	0,07	0,07	0,07	0,00	0,00	0,60	0,13
KVU	0,03	0,03	0,03	0,00	0,00	0,71	0,53
MVU	0,04	0,04	0,04	0,00	0,00	0,05	0,50
LVU	0,03	0,03	0,03	0,00	0,00	0,54	0,79
Afrika	0,14	0,14	0,15	-0,01	0,00	0,03	0,09
Mellemøsten	0,42	0,39	0,39	0,03	0,00	0,00	0,70
Balkan	0,33	0,32	0,31	0,02	0,01	0,00	0,00
Asien	0,10	0,13	0,14	-0,04	-0,01	0,00	0,00
n	25.356	70.287	70.753				

Anm: * Alder og antal børn opgøres i absolutte enheder. Gennemsnit er derfor ikke angivet i andele og den gennemsnitlig ændring er ikke i andele. For kovariater angivet i andele er z-testet anvendt til vurdering af signifikansen, mens der for alder og antal børn er anvendt et t-test.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Flygtninge og familiesammenførte indvandret i 1986 eller senere

Denne målgruppe omfatter ikke blot analysegruppen, men vil også indeholde målgruppen analyseret ovenfor. Gruppen omfatter i alt 141.777 unikke personer, der har opholdt sig i Danmark i kortere eller længere tid i den betragtede periode. Opsummerende karakteristika for denne målgruppe og for analysegruppen før og efter vægtning fremgår af Tabel 5.7.

Målgruppen er for flertallet af kovariaternes vedkommende signifikant forskellig fra analysegruppen. Uddannelsesstrukturen vurderes dog relativt ensartet i de to grupper. En større andel af analysegruppen har dog en erhvervsfaglig eller kort videregående uddannelse som højst fuldførte uddannelse. En større andel af analysegruppen stammer desuden fra Asien eller Balkan, mens andelen fra Mellemøsten er betydeligt lavere end for målgruppens vedkommende. Gennemsnitsalderen er signifikant højere i analysegruppen, hvilket også gælder andelen af enlige familier.

Vægtningen giver også her anledning til en mindskelse af forskellen mellem middelværdierne i den skalerede analysegruppe og målgruppen og dermed en større grad af ensartethed mellem grupperne. Også her vil der dog være tilfælde, hvor insignifikante forskelle mellem den uvægtede analysegruppe og målgruppen er blevet signifikante efter skaleringen. Dette gælder specifikt for kønsfordelingen og antallet af børn. Til trods for dette bedømmes kovariatstrukturen i den vægtede analysegruppe dog overordnet at være så ensartet med målgruppens, at det er meningsfuldt at gennemføre indsatsanalysen på den vægtede analysegruppe.

Tabel 5.7. Sammenligning af analyse- og målgruppe bestående af alle flygtninge og familiesammenførte indvandret i 1986 eller senere, før og efter vægtning

	Gennemsnitlig andel			Forskel til målgruppe		z/t-test p-værdi*	
	Analyse	Analyse vægtet	Mål	Analyse	Analyse vægtet	Analyse	Analyse vægtet
Mænd	0,57	0,57	0,57	0,00	0,00	0,45	0,03
Alder*	26,31	24,46	24,51	1,80	-0,05	0,00	0,40
Par	0,59	0,60	0,60	-0,01	0,00	0,00	0,38
Antal børn*	1,64	1,67	1,66	-0,02	0,02	0,17	0,02
Erhvervsfaglig	0,11	0,08	0,07	0,04	0,00	0,00	0,00
Gymnasial	0,07	0,07	0,07	0,00	0,00	0,21	0,54
KVU	0,03	0,02	0,02	0,01	0,00	0,00	0,60
MVU	0,04	0,04	0,04	0,00	0,00	0,27	0,98
LVU	0,03	0,03	0,03	0,00	0,00	0,93	0,48
Afrika	0,14	0,15	0,16	-0,01	0,00	0,00	0,24
Mellemøsten	0,42	0,54	0,54	-0,12	0,00	0,00	0,62
Balkan	0,33	0,19	0,18	0,15	0,00	0,00	0,01
Asien	0,10	0,09	0,09	0,00	0,00	0,03	0,00
n	25.356	141.245	141.777				

Anm: * Alder og antal børn opgøres i absolutte enheder. Gennemsnit er derfor ikke angivet i andele og den gennemsnitlig ændring er ikke i andele. For kovariater angivet i andele er z-testet anvendt til vurdering af signifikansen, mens der for alder og antal børn er anvendt et t-test.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Ikke-vestlige indvandrere indvandret i 1986 eller senere

Målgruppen bestående af alle ikke-vestlige indvandrere, der er indvandret i 1986 eller senere omfatter 417.190 unikke personer, der har opholdt sig i Danmark i kortere eller længere tid i den betragtede periode. I tillæg til målgruppen i forrige afsnit indregnes nu også personer, hvis opholdsgrundlag ikke er funderet i asyl og familiesammenføring til flygtninge. Således vil vægtningen nu også afspejle den demografiske struktur og det uddannelsesniveau, der kendetegner indvandrere med ophold baseret på eksempelvis arbejde og uddannelse. På samme vis som i de ovenstående afsnit sammenfatter Tabel 5.8 karakteristika for målgruppen og for analysegruppen før og efter vægtning.

Med undtagelse af parstrukturen er målgruppen for samtlige karakteristika signifikant forskellig fra den oprindelige analysegruppe. Relativt til målgruppen stammer en større andel af analysegruppen fra Balkan eller Mellemøsten, mens andelen fra specielt Asien er væsentligt mindre. Endvidere gælder, at andelen af befolkningen med en længerevarende uddannelse er lidt højere i målgruppen. Det er kun en marginalt større andel af analysegruppen, der har en grundskoleuddannelse (residualberegnet), så forskellen i uddannelsesniveaet findes altså over dette niveau.

Den absolutte forskel mellem middelværdierne i analysegruppen og målgruppen mindskes ved vægtning, men forskellen er fortsat vurderet signifikant forskellig fra nul. Dog er andelen med enten en erhvervsfaglig eller kort videregående uddannelse vurderet ens.

Tabel 5.8. Sammenligning af analyse- og målgruppe bestående af alle ikke-vestlige indvandrere indvandret i 1986 eller senere, før og efter vægtning

	Gennemsnitlig andel			Forskel til målgruppe		z/t-test p-værdi*	
	Analyse	Analyse vægtet	Mål	Analyse	Analyse vægtet	Analyse	Analyse vægtet
Mænd	0,57	0,55	0,49	0,08	0,05	0,00	0,00
Alder*	26,31	24,53	24,83	1,48	-0,30	0,00	0,00
Par	0,59	0,56	0,58	0,01	-0,02	0,06	0,00
Antal børn*	1,64	1,36	1,06	0,58	0,30	0,00	0,00
Erhvervsfaglig	0,11	0,06	0,06	0,05	0,00	0,00	0,34
Gymnasial	0,07	0,08	0,09	-0,02	-0,01	0,00	0,00
KVU	0,03	0,02	0,02	0,01	0,00	0,00	0,54
MVU	0,04	0,05	0,06	-0,01	-0,01	0,00	0,00
LVU	0,03	0,05	0,07	-0,04	-0,01	0,00	0,00
Afrika	0,14	0,18	0,17	-0,02	0,01	0,00	0,00
Mellemøsten	0,42	0,30	0,26	0,16	0,04	0,00	0,00
Balkan	0,33	0,15	0,14	0,19	0,01	0,00	0,00
Asien	0,10	0,22	0,26	-0,16	-0,03	0,00	0,00
n	25.356	383.872	417.190				

Anm: * Alder og antal børn opgøres i absolutte enheder. Gennemsnit er derfor ikke angivet i andele og den gennemsnitlig ændring er ikke i andele. For kovariater angivet i andele er z-testet anvendt til vurdering af signifikansen, mens der for alder og antal børn er anvendt et t-test.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Det skal bemærkes, at analysegruppen efter vægtning ikke omfatter et antal personer svarende til den oprindelige målgruppe, men derimod godt 33.000 færre personer. At antallet af personer efter vægtning ikke svarer til målgruppen sås også i mindre udstrækning i Tabel 5.7, hvor differencen er godt 500 personer. Forskellene kan forklares med, at der i målgruppen forekommer personer med en kovariatstruktur, der ikke eksisterer i analysegruppen. Der kan derfor ikke opnås ensartethed med kovariatstrukturen i målgruppen ved vægtning. At der således mangler godt 33.000 personer i analysegruppen efter vægtning understøtter sammen med den signifikante forskel i middelværdierne mellem den vægtede analysegruppe og målgruppen, at kovariatstrukturen i målgruppen er væsentlig anderledes end i analysegruppen. Med andre ord er karakteristika for alle ikke-vestlige indvandrere indrejst i 1986 eller senere markant forskellige fra de karakteristika, der kendetegner flygtninge eksklusiv familiesammenførte indvandret i perioden 1986-98. Som konsekvens heraf vurderes det her ikke rimeligt at benytte den vægtede analysegruppe til vurdering af generaliseringsspørgsmålet.

Indsatseffekt på de vægtede analysegrupper

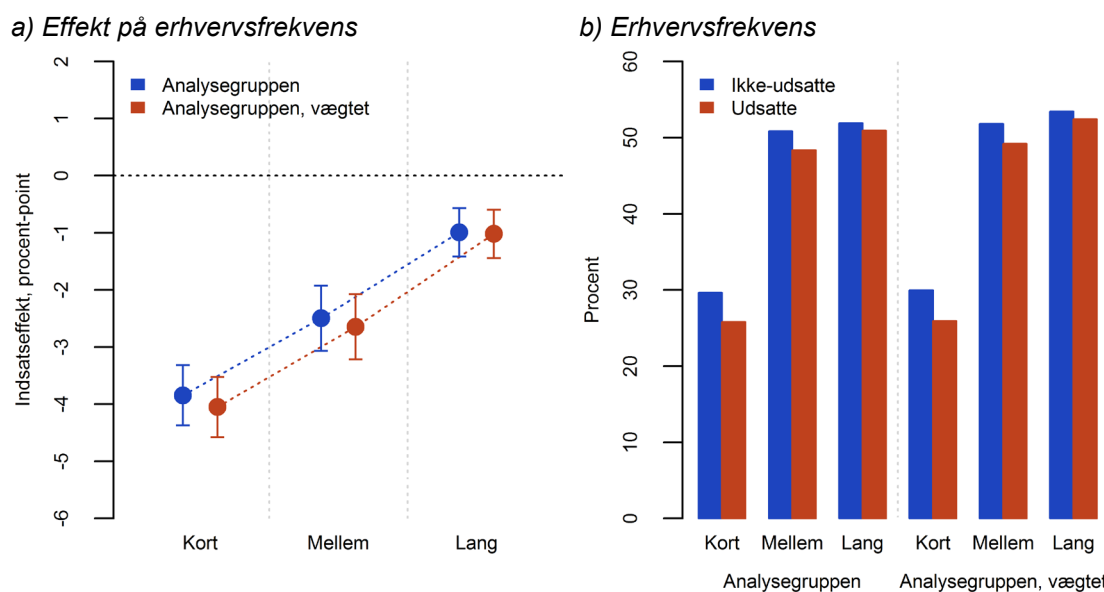
Sammenligningen af kovariatstrukturen mellem den vægtede analysegruppe og målgruppen i hvert af de tre ovenstående tilfælde motiverede, at indsatsanalysen med rimelighed kan gentages på de to vægtede analysegrupper, der gengiver kovariatstrukturen i målgrupperne omfattende flygtninge og familiesammenførte. Nedenfor gentages således indsatsanalysen fra afsnit 5.3 klarlæggende effekten af udsatte områder på hhv. erhvervsdeltagelsen og kriminalitetsomfanget på kort, mellem- og langt sigt for de to relevante vægtede analysegrupper. I modsætning til appendiks H dekomponeres effekten her ikke på de underliggende personkarakteristika. Resultaterne i dette afsnit udtaler sig således kun om

den gennemsnitlige effekt og det kan derfor ikke udelukkes, at indsatsen kan have heterogene effekter, som det var tilfældet for den oprindelige ikke-vægtede analysegruppe.

Effekt på erhvervsfrekvensen

Efter vægtning er erhvervsfrekvensen for personer placeret i udsatte områder på både kort, mellemlangt og langt sigt fortsat signifikant lavere end erhvervsfrekvensen for personer i kontrolgruppen, jf. Figur 5.7a og Figur 5.8a. Uanset om analysegruppen er vægtet med henblik på tilpasning til kovariatstrukturen i målgruppen omfattende flygtninge og familiesammenførte indvandret i løbet af den korte eller lange periode, er effekten på erhvervsdeltagelsen dog ikke signifikant forskellig fra den effekt, der estimeredes for den oprindelige ikke-vægtede analysegruppe. Såfremt befolkningens sammensætning på eksogene variable i analysegruppen repræsenterer strukturen for flygtninge og familiesammenførte indvandret i enten perioden 1986-98 eller i 1986 og senere, kan det oprindelige resultat altså generaliseres. Konsistent hermed er forskellen i erhvervsdeltagelsen mellem indsats- og kontrolgruppen også stort set den samme som for den oprindelige analysegruppe, jf. Figur 5.7b og Figur 5.8b.

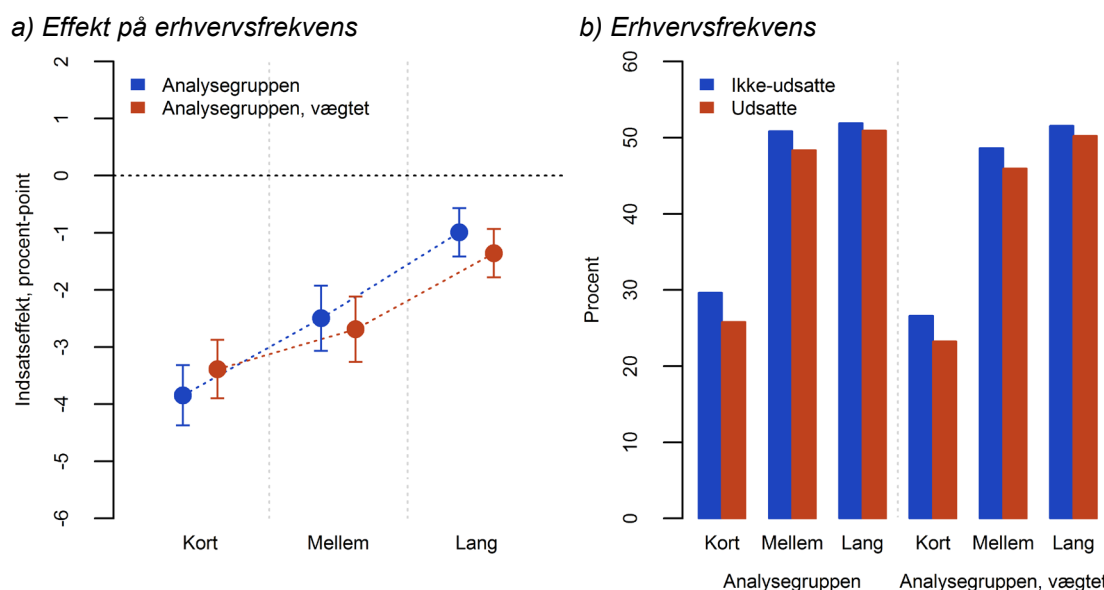
Figur 5.7. Effekt af placering i udsat boligområde på erhvervsfrekvensen for den oprindelige og den vægtede analysegruppe, kovariatstruktur tilpasset målgruppen omfattende flygtninge og familiesammenførte indvandret i perioden 1986-98



Anm.: Analysegruppen (blå kurve i figur a) svarer til Figur 5.5a. Konfidensbånd er indtegnet omkring estimatet for indsats-effekten. Se i øvrigt anm. til Figur 5.5.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 5.8. Effekt af placering i udsat boligområde på erhvervsfrekvensen for den oprindelige og den vægtede analysegruppe, kovariatstruktur tilpasset målgruppen omfattende flygtninge og familiesammenførte indvandret i 1986 eller senere



Anm.: *Analysegruppen (blå kurve i figur a) svarer til Figur 5.5a. Konfidensbånd er indtegnet omkring estimatet for indsatseffekten. Se i øvrigt anm. til Figur 5.5.*

Kilde: *Egne beregninger på baggrund af registerdata fra Danmarks Statistik.*

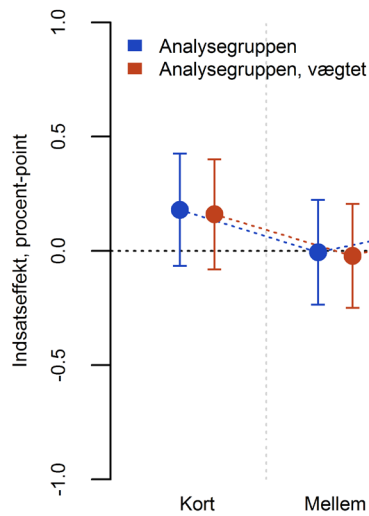
Effekt på kriminalitetsomfang

Konklusionerne vedrørende den gennemsnitlige indsatseffekt på analysegruppens kriminalitetsomfang er stort set de samme for de to vægtede analysegrupper som for den oprindelige analysegruppe. Placering i et udsat område har således ikke signifikant indflydelse på kriminalitetsomfanget på kort, mellemlangt og langt sigt, når samtlige flygtninge og familiesammenførte indvandret i perioden 1986-98 betragtes, jf. Figur 5.9a. Der er dog en tendens til, at udsatte boligområder kan have signifikant indflydelse på graden af kriminalitet på det lange sigt, når den vægtede analysegruppe gengiver kovariatstrukturen fra målgruppen omfattende alle flygtninge og familiesammenførte indvandret i 1986 eller senere, jf. Figur 5.10a.

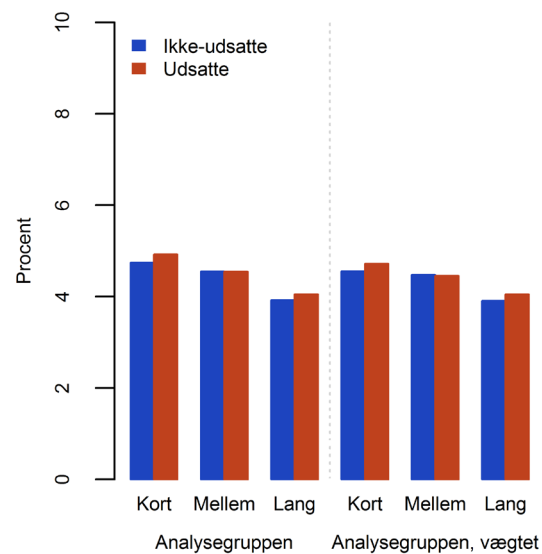
For analysegruppen tilpasset målgruppen med den lange indvandringsperiode medfører vægtningen, at andelen af dømte opjusteres en smule, men den overordnede forskel mellem udsatte og ikke-udsatte områder bevares for begge vægtede analysegrupper relativt til den oprindelige analysegruppe, jf. Figur 5.9b og Figur 5.10b.

Figur 5.9. Effekt af placering i udsat boligområde på kriminalitetsomfang for den oprindelige og den vægtede analysegruppe, kovariatstruktur tilpasset målgruppen omfattende flygtninge og familiesammenførte indvandret i perioden 1986-98

a) Effekt på andel dømte



b) Andel dømte

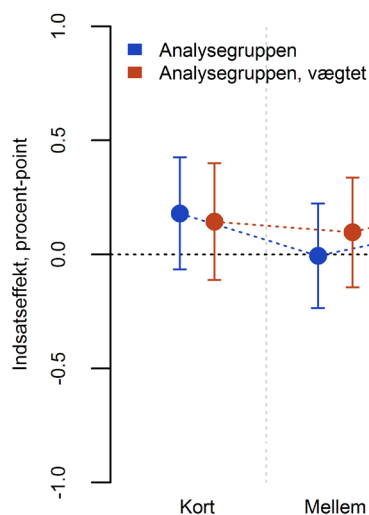


Anm.: Analysegruppen (blå kurve i figur a) svarer til Figur 5.6a. Konfidensbånd er indtegnet omkring estimatet for indsats effekten. Se i øvrigt anm. til Figur 5.5.

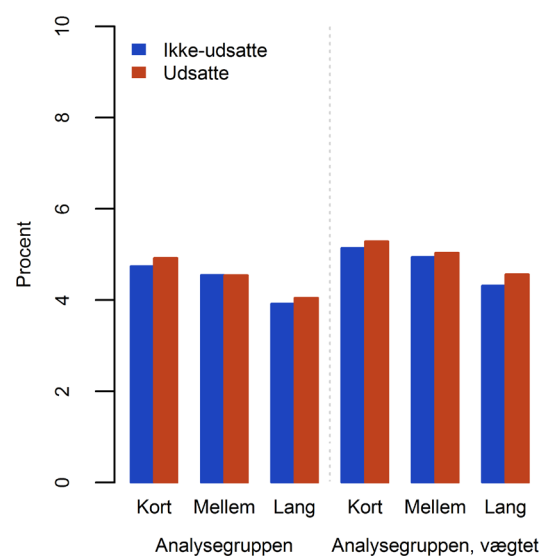
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 5.10. Effekt af placering i udsat boligområde på kriminalitetsomfang for den oprindelige og den vægtede analysegruppe, kovariatstruktur tilpasset målgruppen omfattende flygtninge og familiesammenførte indvandret i 1986 eller senere

a) Effekt på andel dømte



b) Andel dømte



Anm.: Analysegruppen (blå kurve i figur a) svarer til Figur 5.6a. Konfidensbånd er indtegnet omkring estimatet for indsats effekten. Se i øvrigt anm. til Figur 5.5.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

6. Clusted Nearest Neighbor algoritmen

Dette afsnit beskriver Clusted Nearest Neighbor algoritmen (CNN), som er en metode udviklet til at definere segregerede boligområder. Algoritmen tager udgangspunkt i unsupervised learning metoder som k -nearest neighbor algoritmen, principal komponent analyse og hierarkisk clustering analyse.

Algoritmen starter med at udregne afstanden mellem alle boliger/opgange i Danmark, hvorefter hver boligs nabolag bestemmes ud fra k -nearest neighbor algoritmen. Enhver bolig får tildelt et unikt nabolag, som består af de 1.000 nærmeste naboer. For ethvert af disse nabolag udregnes et gennemsnit for de socioøkonomiske variable, dvs. uddannelses- og arbejdsmarkedstilknytning, kriminalitet, uddannelsesniveau og indkomst. Med baggrund i disse socioøkonomiske variable bliver der udført en principal komponent analyse. Dette resulterer i et endimensionalt socioøkonomisk indeks, der rangerer ethvert nabolag fra socioøkonomisk udsatte til bedre stillede i forhold til alle andre nabolag. Indekset trunkeres således, at ejendomme med de mest udsatte nabolag vælges, hvor fem procent af befolkningen bor.

Med baggrund i de udvalgte ejendomme udføres en hierarkisk clustering analyse. Det vil sige, at ejendomme, hvis individuelle nabolag opfylder kravet om at tilhøre de fem procent mest udsatte nabolag, grupperes efter korteste afstand mellem hinanden. Alle klynger af ejendomme, der består af mere end 1.000 beboere, beholdes og benævnes et udsat boligområde.

De følgende afsnit beskriver hvert af skridtene i algoritmen mere detaljeret.

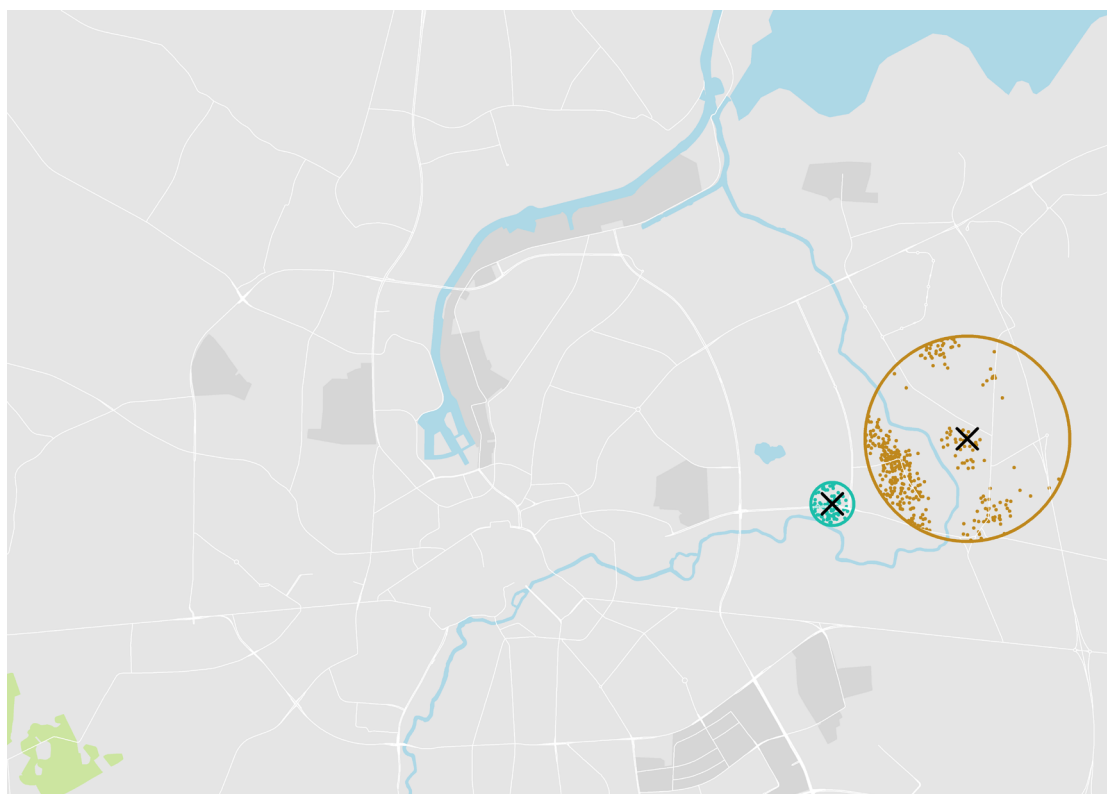
6.1 Nabolag og socioøkonomiske variable

Et nabolag er i denne analyse defineret ud fra k -nearest neighbor algoritmen (k -NN), jf. Fix and Hodges (1951), hvor k er antallet af naboer i nabolaget. k -NN er en ikke-parametrisk metode, hvilket vil sige, at der ikke er nogen underliggende antagelser omkring fordeling af data. Udgangspunktet for denne analyse er, at k er lig med 1.000. Dette vil sige, at k -NN algoritmen med udgangspunkt i afstanden mellem alle boliger bestemmer et unikt nabolag for hver bolig bestående af de 1.000 nærmeste naboer. Dette matcher minimumskravet til størrelsen af et udsat boligområde.

Figur 6.1 viser to tilfældigt udvalgte nabolag i Odense. Begge nabolag består af 1.000 personer. Det ses, at afstanden til den tusinde nabo i nabolaget varierer fra nabolag til nabolag. Afstanden afhænger af befolkningstætheden i det pågældende område. De to nabolag i Figur 6.1 har henholdsvis en radius på 200 meter og en på 1 kilometer.

To naboboliger, der er beliggende tæt ved hinanden, får to unikke nabolag, som dog stort set er ens. Måske vil de to nabolag kun have én bolig til forskel. k -NN-algoritmen sikrer herved geografisk spatial kontinuitet nabolagene imellem.

Figur 6.1. To tilfældigt udvalgte nabolag i Odense, 1.000 personer i nabolaget.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

For ethvert nabolag beregnes et gennemsnit for de anvendte socioøkonomiske variable, dvs. for hvert nabolag opgøres den gennemsnitlige uddannelses- og arbejdsmarkedstilknytning, kriminalitetsniveau, uddannelsesniveau og den gennemsnitlige indkomst.

- Uddannelses- og arbejdsmarkedstilknytning er defineret som andelen af beboere i et nabolag i den arbejdsdygtige alder (18-64 år) uden uddannelses- og arbejdsmarkedstilknytning.
- Kriminalitet er defineret som andelen af beboere i et nabolag på 15 år eller derover, som er dømt for overtrædelse af straffeloven, våbenloven eller loven om euforiserende stoffer inden for det seneste år.
- Uddannelsesniveau er defineret som andelen af beboere i et nabolag i alderen 30-59 år, som har grundskole som højst fuldførte uddannelsesniveau.
- Indkomst er defineret som nabolagets gennemsnitlige indkomst for beboere i den skattepligtige alder (15-64 år) eksklusiv uddannelsessøgende i forhold til samme gruppe for hele regionen.

Størrelsen af nabolag har en vis betydning for analysen. Vælges der et mindre k end 1.000, vil det resultere i nabolag, der er mindre udglattede og med højere kontrast. Hvilket i sidste ende kan resultere i, at områder ikke identificeres som udsatte områder i den efterfølgende analyse, pga. krav om minimum 1000 personer i et udsat område. Jo større k , der vælges, desto mere vil nabolaget minde om kommune-, regions- og landsgennemsnittet.

6.2 SNI via principalkomponentanalyse

Et endimensionalt Socioøkonomisk Nabolag-Indeks (SNI) konstrueres ud fra de socioøkonomiske variable, som vi ønsker at undersøge segregering i Danmark med hensyn

til. Dette indeks danner grundlag for den efterfølgende clustering analyse og bruges i afsnit 2.1 til at vurdere segregeringsniveauet. Indekset er konstrueret ud fra en principal komponent analyses første komponents scoreværdier.

Formålet med principalkomponentanalyse er at konstruere den bedste m -dimensionale approksimation, målt i euklidisk afstand, til enhver observation i et p -dimensionalt datasæt, hvor $p > m$, jf. James et al. (2013). I dette tilfælde er formålet med analysen at konstruere den bedst éndimensionale approksimation af de fire socioøkonomiske variable over nabolag.

Principal komponent analysen er udført på skaleret og centreret Yeo-Johnson (2000) power transformation af de socioøkonomiske parametre: Uddannelses- og arbejdsmarkedstilknytning, kriminalitet, uddannelsesniveau og indkomst. Transformationen udføres før selve principal komponent analysen for at sikre en standardiseret fordeling med forbedret normalfordelte egenskaber³⁵.

Den første principalkomponent Z_{1t} til tidspunkt t er defineret som den normaliserede lineære kombination af de p socioøkonomiske variable $X_{1t}, X_{2t}, \dots, X_{pt}$

$$Z_{1t} = \phi_{11t}X_{1t} + \phi_{21t}X_{2t} + \dots + \phi_{p1t}X_{pt}$$

som har den størst mulig varians til tidspunkt t under betingelse af normaliserede loadings, ϕ

$$\sum_{j=1}^p \phi_{j1t}^2 = 1 \quad \forall t.$$

Den første principalkomponents loadings til ethvert t findes ved at løse optimeringsproblemet for variansen³⁶:

$$\begin{aligned} \max_{\phi_{11t}, \phi_{21t}, \dots, \phi_{p1t}} \quad & \frac{1}{n} \sum_{i=1}^n \left(\sum_{j=1}^p \phi_{j1t} x_{ijt} \right)^2 \\ \text{st.} \quad & \sum_{j=1}^p \phi_{j1t}^2 = 1 \quad \forall t \end{aligned}$$

Hvor det er antaget, at vi har et $n \times p$ datasæt, X_t , hvorom det gælder, at $x_{ijt} \in X_t$, og hvor i er et af n nabolag, $i \in n$, og j er en af p skalerede og centrerede socioøkonomiske variable, $j \in p$.

Normaliseringen af loadingsvektoren sikrer en entydig løsning af optimeringsproblemet, således, at man ikke kan få en arbitrær stor varians blot ved at øge loadings-koefficienterne.

Figur 6.2a viser de optimale loadings, $\tilde{\phi}$, over det betragtede tidsinterval t . Det ses, at loadings er rimeligt stabile over tid. Der er dog en tendens til andelen af dømte i et nabolag

³⁵ Hvis analysen udføres på et ikke-skaleret datasæt, hvor de enkelte variable måles på forskellige skalaer, vil de tilhørende varianser også måles på forskellige skalaer. Dette ville medføre, at man i principalkomponentanalysen uden begrundelse vil give de enkelte variable forskellige vægte, jf. James et al. (2013).

³⁶ Da alle de socioøkonomiske variable er transformeret til at have middelværdi nul kan sample variansen for den første principalkomponent reduceres til:

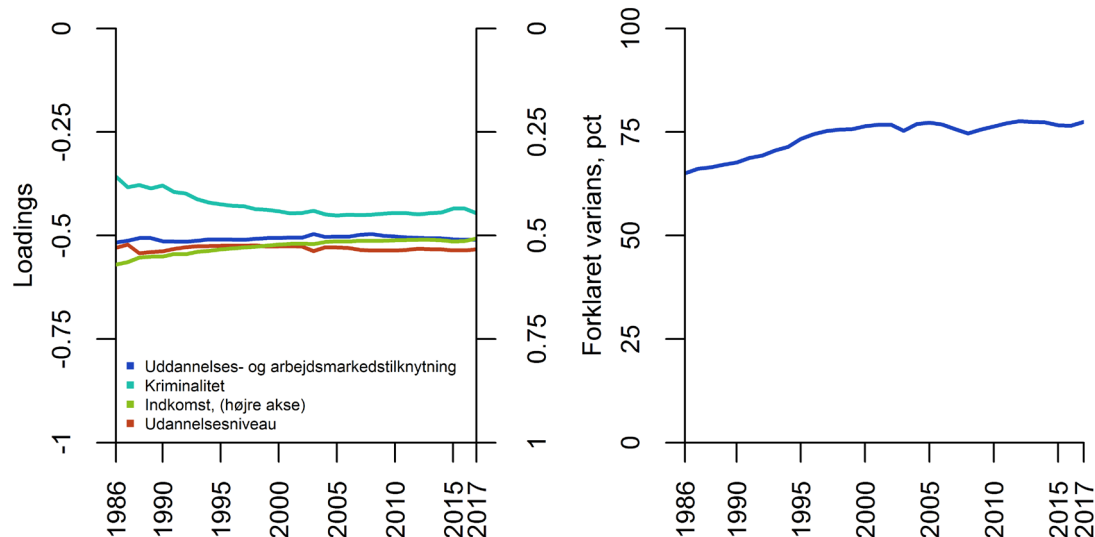
$$\text{var}(Z_{1t}(X_t)) = \frac{1}{n} \sum_{i=1}^n \left(\sum_{j=1}^p \phi_{j1t} x_{ijt} - \frac{1}{n} \sum_{i=1}^n x_{ijt} \right)^2 \Rightarrow \text{var}(Z_{1t}(X_t)) = \frac{1}{n} \sum_{i=1}^n \left(\sum_{j=1}^p \phi_{j1t} x_{ijt} \right)^2$$

var af mindre betydning i starten af perioden. Derudover ses det, at alle optimale loadings stort set findes til at have samme absolutte værdi.

Figur 6.2. Deskriptive resultater for principal komponent analysen.

a) Optimale loadings over tid

b) SNI forklaring af socioøkonomiske variables samlede varians over tid



Kilde: Egen beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 6.2b viser, hvor meget den første principalkomponent forklarer af de socioøkonomiske variables samlede varians over tid, PVE_t ³⁷:

$$PVE_t = \frac{\sum_{i=1}^n \left(\sum_{j=1}^p \phi_{j1t} x_{ijt} \right)^2}{\sum_{j=1}^p \sum_{i=1}^n x_{ijt}^2}$$

De efterfølgende principalkomponenter forklarer efter tur den øvrige del af variationen begrænset af at være perfekt ukorreleret med de foregående principalkomponenter. Per definition udspænder alle fire principalkomponenter hele datasættets varians. Andelen af den samlede varians, som den enkelte principalkomponent forklarer, aftager markant i antallet af komponenter. Dette giver mulighed for en lavdimensional repræsentation af de fire socioøkonomiske variable med blot én principalkomponent.

Op mod 80 procent af den samlede varians mellem de fire variable forklares af den første principalkomponent, jf. Figur 6.2b. Det konkluderes på denne baggrund, at første principalkomponent er en god endimensional repræsentation af de fire socioøkonomiske variable. Den første principalkomponent benævnes Socioøkonomisk Nabolag-Indeks (SNI).

Ud fra loadings'ene kan man finde scoreværdien, z_{i1} , i periode k til hvert enkelt nabolag som:

$$z_{i1t} = \tilde{\phi}_{11t} x_{i1t} + \tilde{\phi}_{21t} x_{i2t} + \dots + \tilde{\phi}_{p1t} x_{ipt},$$

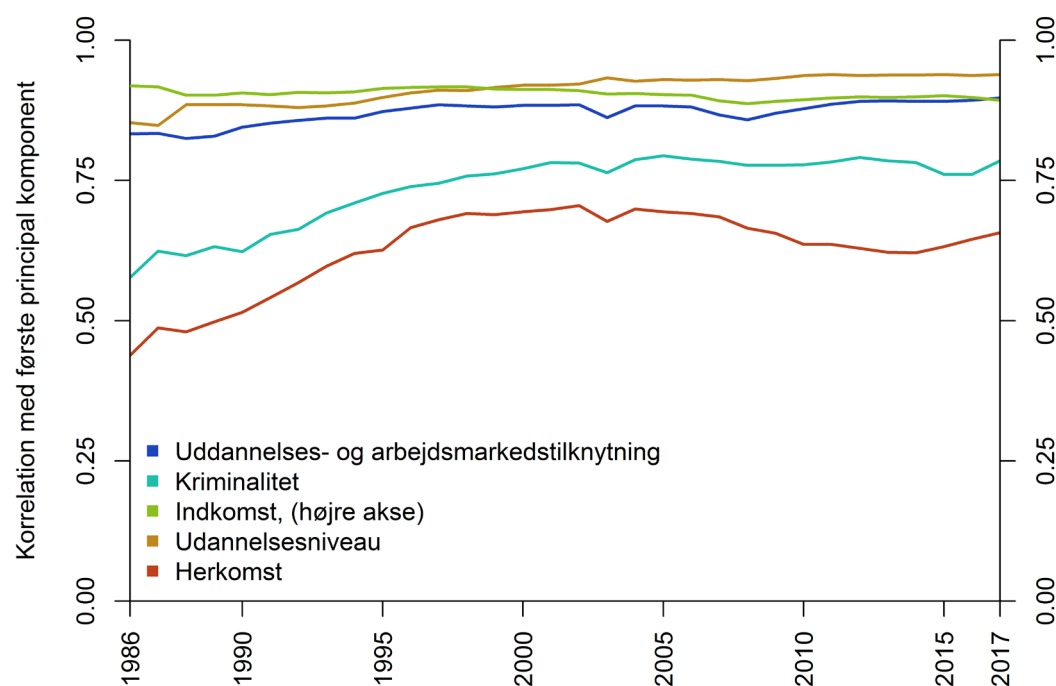
³⁷ Under antagelse af at de socioøkonomiske variable er centreret omkring middelværdi nul.

Fra Figur 6.2a ses det, at scoreværdien imputeres med stort set lige vægt på alle de socioøkonomiske variable, hvorfor alle af variablene stort set er lige vigtige, når de udsatte områder efterfølgende skal findes.

Til at fortolke scoreværdien kan det være nyttigt at kigge på korrelationen mellem principalkomponenten og de socioøkonomiske variable. Figur 6.3 viser korrelationen mellem den første principalkomponent og de socioøkonomiske variable over tid. Det ses, at uddannelses- og arbejdsmarkedstilknytning, kriminalitet og uddannelsesniveau er negativt korreleret med principalkomponenten, mens indkomst er positivt korreleret med den. Hvorfor udsatte nabolag har en høj negativ scoreværdi, mens bedrestillede nabolag har en høj positiv scoreværdi. Altså vil et nabolag med en stor andel af beboere, der er dømte, uden uddannelses- eller arbejdsmarkedstilknytning eller en stor andel med blot en grundskoleuddannelse, have en principalkomponent scoreværdi, som er negativ og vise versa. Mens et nabolag med en lav gennemsnitlig indkomst i forhold til hele regionen vil have en negativ principalkomponent scoreværdi og vice versa.

Bemærk at herkomst, defineret som andelen af ikke-vestlige indvandrere eller efterkommere i et nabolag, ikke er medtaget i principalkomponentanalysen, men trods det har en høj korrelation med det Socioøkonomiske Nabolag-Indeks.

Figur 6.3. SNI korrelation med socioøkonomiske variable og ikke-vestlig herkomst over tid.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Ud fra det Socioøkonomiske Nabolag-Indeks defineres hvilke ejendomme, der potentielt tilhører et udsat område. Der vælges ejendomme med de gennemsnitlige laveste scoreværdier, hvor 5 procent af befolkningen bor. 5 procent grænsen er valgt ud fra en kalibrering af CNN-algoritmen under kriteriet af, at CNN-algoritmen i 2017 udpeger 90 procent af TBBs udsatte boligområder. De valgte ejendomme med de mest udsatte nabolag grupperes til udsatte områder i den efterfølgende hierarkisk clustering analyse.

6.3 Hierarkisk clustering analyse

Ejendomme, der i den forgående principalkomponentanalyse blev fundet til at have de mest udsatte nabolag, grupperes efter korteste afstand mellem hinanden. Herved dannes samhörighed mellem ejendomme, der har de mest udsatte nabolag og hvis afstand er lille. Alle klynger af ejendomme, der består af mere end 1.000 beboere, benævnes et udsat boligområde.

Klyngerne af ejendomme findes via hierarkisk clustering analyse. Denne metode inddeler et dataset i homogene unikke undergrupper således at observationer, i dette tilfælde ejendomme, der er rimeligt ens er i samme klynge, mens de er forskellige fra ejendomme i andre klynger. I denne hierarkisk clustering analyse bruges den geografiske afstand mellem ejendomme til at afgøre om ejendommene skal grupperes.

Den hierarkiske clustering er en bottom-up proces, der starter med at alle ejendomme er i hver sin klynge. Sekventielt kombineres hver klynge til større og større klynger indtil alle ejendomme er i samme klynge. I hvert skridt kombineres de to klynger, der er adskilt med den korteste afstand. Med andre ord starter hver ejendom med at være hvert deres udsatte område, hvorefter ejendommene efter tur sammenlægges i udsatte områder, hvor afstanden mellem disse områder er den korteste. Specifikt er det algoritisk hierarkisk clustering med average-linkage, der bruges i denne analyse, jf. Sokal et al. (1958) og Sokal et al. (1963). Processen er eksemplificeret i appendiks D.

Udfordringen ved hierarkisk clustering analyse er for det første at bestemme, hvor mange klynger data skal underopdeles i, det vil sige, hvornår algoritmen skal stoppe. For det andet, hvordan metoden til at måle "den korteste afstand" mellem to klynger defineres, jf. appendiks D.

Hvis antallet af klynger er for stort, bliver de udsatte områder ofte mindre end 1.000 personer, og man vil se udsatte områder, der geografisk ligger tæt op af hinanden. Vælges et lille antal klynger, vil hvert område indeholde et stort antal ejendomme, som principal komponent analysen ikke vurderede til at have et udsat nabolag.

Antallet af klynger er valgt således, at den gennemsnitlige afstand mellem dem er mere end 600 meter. De 600 meter er fundet via en kalibrering af den korteste afstand, da det resulterede i de "pæneste" visuelle områder ud fra mange gennemkørsler af CNN algoritmen. CNN algoritmen er udført med Complete, Average, Single linkage. Det endelig valg af Average linkage blev foretaget med baggrund i en visuel betragtning af de udsatte områder, algoritmen resulterede i.

7. Konklusion

I analysen behandles bopælsmæssig socioøkonomisk segregering. Analysen består af fire dele, som behandler hver deres emne indenfor segregering.

Først identificeres udsatte boligområder i perioden 1986-2017 ved hjælp af en nyudviklet metode. Af metoden fremkommer et indeks, som siger noget om udviklingen i segregeringsniveauet gennem perioden. Til- og fraflytningsmønstre analyseres med udgangspunkt i de identificerede udsatte boligområder. Endelig relateres indkomstforløb til erfaring med bosættelse i udsatte boligområder samt andre personkarakteristika.

Metode til identifikation af udsatte boligområder

Analysen præsenterer en nyudviklet metode til at udpege segregerede boligområder. Metoden bruges i denne analyse til at identificere socioøkonomisk udsatte boligområder i perioden 1986 til 2017.

Transport-, Bygnings- og Boligministeriet har i 2008 udviklet en metode til at udpege udsatte almene boligområder. Denne metode bruges til at udarbejde regeringens årlige ghettoliste. Metoden præsenteret i denne analyse adskiller sig på flere punkter fra Transport-, Bygnings- og Boligministeriets metode. For det første er etnicitet ikke et kriterie, når segregerede boligområder identificeres. For det andet er metoden, som nævnt, udelukkende baseret på dataanalyse, således at der altid anvendes de samme kriterier. Som følge heraf er der ikke behov for at udføre områdespecifikke undersøgelser som for eksempel fysiske besøg af boligområder. Denne mere automatiserede tilgang gør det muligt at gennemføre analysen over en samlet historisk periode 1986-2017.

Udsatte boligområder identificeres til generelt at være placeret centralt i de større byer omkring større almene bebyggelser. Områderne har over den betragtede periode været tilnærmelsesvis stabile i udbredelsen og placeringen. Væsentlige undtagelser er dog København og Odense. Placeringen af udsatte boligområder har i København gennemgået en betydelig forandring, hvilket blandt andet forklares af store byfornyelsesprojekter. En anden undtagelse er det store almene boligområde Vollsmose i Odense, som i midt 80'erne ikke bliver identificeret som et udsat boligområde.

Kendetegnende for beboere i de udsatte områder er, at de har en lavere indkomst, som udgør cirka to tredjedele af indkomsten for personer bosat i ikke-udsatte områder. Desuden har beboerne i udsatte områder lavere arbejdsmarkedstilknytning, ringere uddannelsesniveau og er yngre end befolkningen generelt. Det trækker også mod en lavere gennemsnitlig indkomst. Andelen af beboere i udsatte boligområder, der har været dømt for forseelser udover færdselsloven, er omkring tre-fire gange så stor som i befolkningen som helhed. Andelen af ikke-vestlige indvandrere og efterkommere i udsatte boligområder har været stigende over perioden og tilvæksten overgår markant den generelle stigning på landsplan.

Socioøkonomisk segregering

Et interessant biprodukt af den udviklede metode til identificering af udsatte boligområder er, at vi kan udtale os om udviklingen i den socioøkonomiske segregering i Danmark. Segregering i denne analyse betegner en overrepræsentation eller koncentration af socioøkonomisk udsatte personer i bestemte områder. Vi finder, at de 5 procent mest udsatte nabolag i Danmark har en tendens til at være mere segregeret i forhold til den socioøkonomiske mellemklasse. Der argumenteres for, at den socioøkonomiske segregering for denne gruppe, over den betragtede periode, har været faldende i København, mens den

har været tilnærmelsesvis stabil i resten af landet. Med andre ord er sandsynligheden for at møde én person fra et udsat nabolag et vilkårligt sted i Københavns Kommune blevet større over den betragtede periode..

Flyttemønstre i varigt udsatte almene boligområder

Analysen viser, at udskiftningen af beboere i varigt udsatte almene boligområder er cirka 17 procent. Dette er på samme niveau som i øvrige almene boliger og lidt højere end i andels- og ejerlejligheder.

Flytninger kan over tid være med til at fastholde eller ændre strukturen i et boligområde. Analysen viser, at tilflyttere til varigt udsatte almene boligområder i gennemsnit har lavere indkomst og flere er uden uddannelse udover grundskolen, end det er tilfældet for beboere, som fraflytter området. Flyttemønstret er således medvirkende til at reducere det gennemsnitlige indkomstniveau i de varigt udsatte almene boligområder.

En stor andel af tilflytterne til de betragtede udsatte boligområder indvandrer fra udlandet i tilflytningsåret. Disse personer har i gennemsnit ringere arbejdsmarkedstilknytning, er dårligere uddannet og har lavere indkomstniveau end øvrige tilflyttere til områderne. Der ses en klar tendens til, at tilflytterne over få år nærmer sig indkomstniveauet for områdets øvrige beboere. Dette følger øjensynligt af, at der går noget tid før den relativt store andel af tilflyttere, som er nyindvandret, etablerer sig på arbejdsmarkedet.

Betragtes udelukkende tilflyttere til varigt udsatte almene boligområder, som flytter internt i landet, ligner disse tilflyttere i høj grad de personer, som fraflytter områderne.

Indkomstforløb og bosættelse i udsatte boligområder

For en udvalgt målgruppe gennemførtes en gruppering af individuelle indkomstforløb, der gav anledning til, at forløbene meningsfuldt kunne kategoriseres i fire forskellige grupper. Den første gruppe er karakteriseret ved at omfatte personer, der typisk opholder sig i de lavere indkomstdeciler gennem analyseperioden, hvor de pågældende har en relativt høj ledighed, en løs tilknytning til arbejdsmarkedet og en relativt høj tilbøjelighed til gradvist at overgå til førtidspension. Den anden og tredje gruppe omfatter personer, der har indkomstforløb, der typisk er karakteriseret af indkomster lidt over og under medianen. Begge grupper har en stabil tilknytning til arbejdsmarkedet. Forskellen mellem de to grupper skal primært findes i gruppernes uddannelsesstruktur, hvor den tredje gruppe er kendetegnet ved en lavere andel med grundskoleuddannelse og en større andel med en videregående uddannelse end den anden gruppe. Personer kategoriseret i den fjerde gruppe har indkomstforløb præget af ophold i de højeste indkomstdeciler, en stort set entydig tilknytning til arbejdsmarkedet og et højt uddannelsesniveau.

Personer, der i løbet af analyseperioden var bosat et eller flere år i et udsat område, placeredes fortrinsvis i grupperne med de laveste og næstlaveste indkomster, men var repræsenterede i alle fire grupper. Dette gav således ikke anledning til at tro, at personer med bopæls erfaring i et udsat område nødvendigvis skulle have indkomstforløb, der adskiller sig signifikant fra resten af befolkningens. Resultaterne fra en række statistiske analyser bekræftede, at der ikke var et entydigt signifikant forhold mellem gruppetilhørsforhold og bopæls erfaring fra et udsat område. Konklusionen afhang af modelformuleringen og resultaterne syntes frem for alt at indikere, at befolkningen tilordnedes indkomstgrupperne i kraft af øvrige socioøkonomiske karakteristika som uddannelsesniveau og arbejdsmarkedstilknytning. Resultaterne var funderet i en betragtning af befolkningen som helhed, hvorfor forskellige konklusioner om betydningen af områdestatus for tilhørsforholdet til

indkomstgrupperne potentielt kan opnås, hvis man betragter undergrupper af befolkningen afgrænset ved eksempelvis oprindelse, uddannelse eller arbejdsmarkedstilknytning.

Betragtes forholdet mellem personkarakteristika og forskellen i individuelle indkomstforløb, var betydningen af områdeerfaring konsekvent insignifikant, mens de socioøkonomiske og statiske demografiske karakteristika havde signifikant betydning. I forlængelse af de univariate og multivariate analyser tydeliggjorde regressionstræet, at betydningen af et karakteristikum for forskellen mellem indkomstforløb er knyttet til værdien af andre personspecifikke forhold. Dette motiverer, at bl.a. forholdet mellem områdeerfaring og indkomstforløb kan belyses nærmere for udvalgte undergrupper afgrænset af eksempelvis uddannelse, arbejdsmarkedsstatus eller oprindelse med henblik på afklaring af eventuelle heterogene effekter.

Effekt af bosætning i udsatte boligområder: Et quasi-randomiseret eksperiment

Analysen viser, at bosættelse i et udsat boligområde har en signifikant negativ effekt på flygtnings erhvervsfrekvens. Analysen finder således, at erhvervsfrekvensen for flygtninge allokere til et udsat boligområde på både kort og langt sigt er signifikant lavere end erhvervsfrekvensen for flygtninge, der initialt blev placeret i et ikke-udsat boligområde.

På kort og mellemlangt sigt har placeringen i et udsat boligområde for begge køn en negativ effekt på erhvervsfrekvensen. Det er dog kvinders arbejdsmarkedsadfærd, der påvirkes mest negativt. Effekten af initialt at blive placeret i et udsat boligområde findes på langt sigt at være forskellig for mænd og kvinder. På langt sigt finder analysen ingen signifikant forskel i erhvervsfrekvensen for mandlige flygtninge, der er allokere til et udsat boligområde i forhold til mandlige flygtninge, der blev allokere til et ikke-udsat boligområde. Modsat mænd har kvinder på langt sigt en signifikant lavere erhvervsfrekvens som en konsekvens af den initiale allokering til et udsat boligområde.

Effekten på erhvervsfrekvensen varierer desuden på tværs af oprindelsesland. På kort sigt oplever flygtninge fra Balkanhalvøen og Asien en stor negativ erhvervsfrekvenseffekt, mens der på langt sigt ikke findes nogen signifikant effekt. Flygtninge fra Afrika og Mellemøsten oplever både på kort, mellemlangt og langt sigt en stabil negativ effekt på arbejdsmarkedstilknytningen af initialt at blive placeret i et udsat boligområde.

Der findes ingen samlet signifikant effekt på kriminalitet af at bo i et udsat boligområde. Denne manglende effekt findes stort set også uanset køn, flygtnings oprindelsesområde og hvorvidt vedkommende var voksen eller barn på indvandringstidspunktet.

Analysen peger på, at effekten kan generaliseres til at omfatte alle flygtninge og familiesammenførte indvandrede op til i dag, men derimod ikke til alle ikke-vestlige indvandrere.

8. Referencer

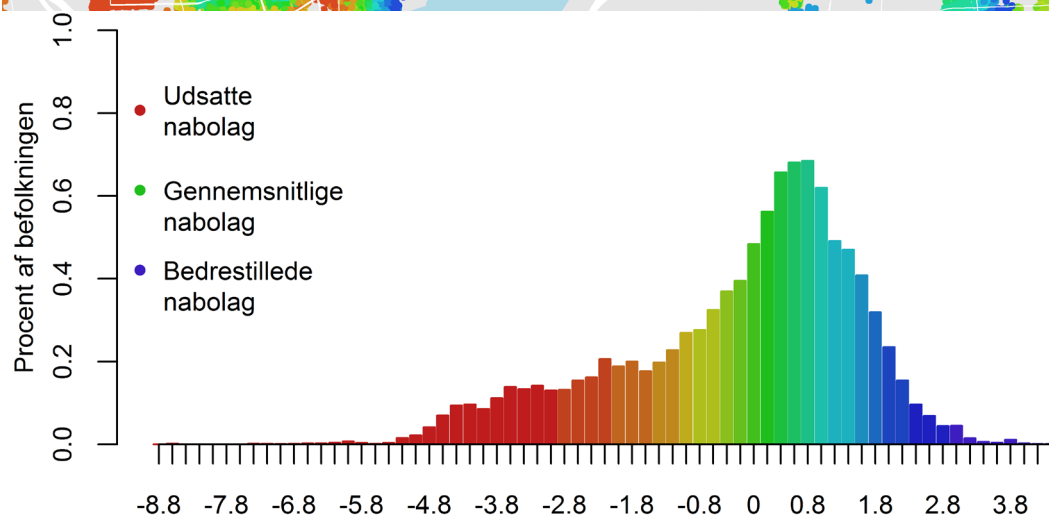
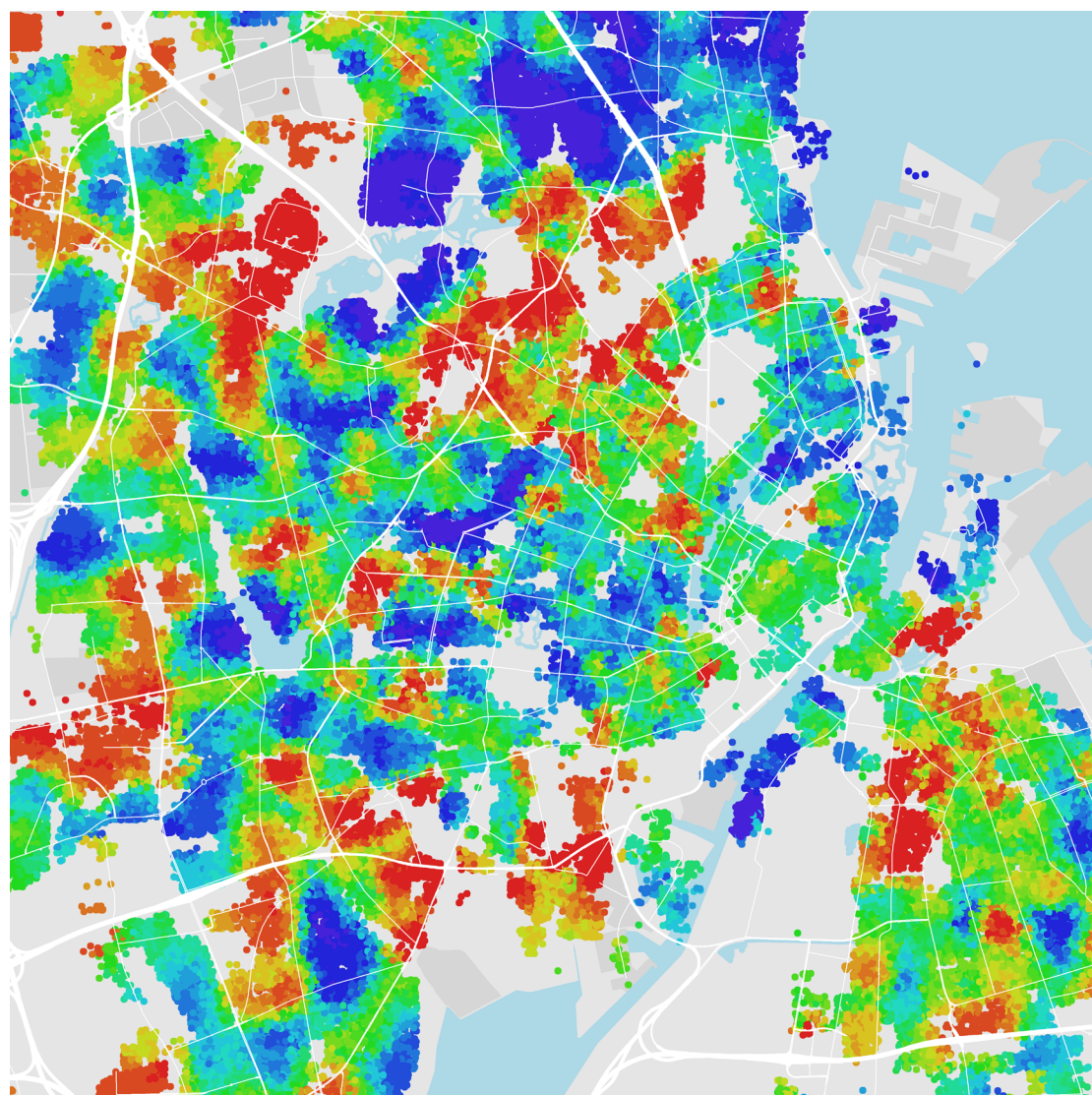
- Akerlof, G. A. (1997): *Social Distance and Social Decisions*. *Econometrica*, 65(5), 1005–1027.
- Anderson, M. J. (2001): *A new method for non-parametric multivariate analysis of variance*. *Austral Ecology*, 26, 32–46.
- Antonakis, J.; Bendahan, S.; Jacquart, P. & Lalive, R. (2010): *On making causal claims: A review and recommendations*. *The Leadership Quarterly* 21 (6), 1086–1120.
- Banerjee, A. V. (1992): *A Simple Model of Herd Behavior*. *Quarterly Journal of Economics*, 107, 797–818.
- Benedetti, J. K. (1977): *On the Nonparametric Estimation of Regression Functions*. *Journal of the Royal Statistical Society, Ser. B*, 39, 248–253.
- Bruch, E. E. & Mare, R. D. (2009): *Segregation Processes*. Side 269-293 i Bearman, P. & Hedström, P. (eds.): *Oxford Handbook of Analytical Sociology*. Oxford University Press.
- Caliendo, M. & Kopeining, S. (2005): *Some Practical Guidance for the Implementation of Propensity Score Matching*. Discussion Paper No. 1588.
- Cole, S. R. & Stuart, E. A. (2010): *Generalizing Evidence from Randomized Clinical Trials To Target Populations, The ACTG 320 Trial*. *American Journal of Epidemiology*, vol. 172, no. 1.
- Defays D. (1977): *An efficient algorithm for a complete link method*. *Computer Journal*, 20, 364–366.
- Damm, A. P. (2005): *The Danish Dispersal Policy on Refugee Immigrants 1986-1998: A Natural Experiment?* Aarhus School of Business, Department of Economics, WP 05-3.
- Damm, A. P. (2009a): *Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence*. *Journal of Labor Economics*, 27(2), 281–314
- Damm, A. P. (2009b): *Determinants of Recent Immigrants' Location Choices: Quasi-Experimental Evidence*. *Journal of Population Economics*, 22(1), 145–174
- Damm, A. P. & Rosholm, M. (2010): *Employment Effects of Spatial Dispersal of Refugees*. *Review of Economics of the Household*, 8(1), 105–146
- Damm, A. P. (2014): *Neighborhood Quality and Labor Market Outcomes: Evidence from QuasiRandom Neighborhood Assignment of Immigrants*. *Journal of Urban Economics*, 79, 139–166.
- Damm, A. P. & Dustmann, C. (2014): *Does Growing Up in a High Crime Neighborhood affect Youth Criminal Behavior?* *American Economic Review*, 104(6), 1806–1832.
- Danmarks Statistik (2018): *Indvandrere i Danmark 2018*. København.
- Dustmann, C., Vasiljeva, K. & Damm, A. P. (2016): *Refugee Migration and Electoral Outcomes*. Kommende publicering i *Review of Economic Studies*.
- Enos, Ryan D. (2018): *The Space between Us: Social Geography and Politics*. Cambridge University Press.
- Fix, E. & Hodges, J. L. (1951): *Discriminatory analysis, nonparametric discrimination, consistency properties*. *Randolph Field*, 21-49-004, No. 4.
- Florek, K.; Lukaszewics, J.; Perkal, J; Steinhaus, H. & Zubrzycki, S. (1951): *Sur la liaison et la division des points d'un ensemble fini*. *Colloquium Mathematicum*, 2, 282–285.

- Fulda, B. E. (2016): *The diversity in Longitudinal partnership trajectories during the transition to adulthood: How is it related to individual characteristics and regional living conditions?* Demographic Research, 35, 1101–1134.
- Gabandinho, A.; Ritschard, G.; Studer, M. & Müller, N. S. (2011a): *Mining sequence data in R with the TraMineR package: A user's guide*. Department of Econometrics and Laboratory of Demography, University of Geneva, Switzerland.
- Gabandinho, A.; Ritschard, G.; Müller, N.S. & Studer, M. (2011b): *Analyzing and visualizing state sequences in R with TraMineR*, Journal of Statistical Software, 40 (4), 1–37.
- Gabandinho, A.; Ritschard, G.; Studer, M. & Müller, N. S. (2011c): *Discrepancy Analysis of State Sequences*. Sociological Methods and Research, 40 (3), 471–510.
- Graham, B.S. (2018): *Identifying and Estimating Neighborhood Effects*. Journal of Economic Literature, 56(2), 450–500.
- Holland, P.W. (1986): *Statistics and Causal Inference*. Journal of the American Statistical Association, 81(396), 945–960.
- James, G.; Witten, D.; Tibshirani, R. & Hastie, T. (2013): *An Introduction to Statistical Learning with Applications in R*. New York, Springer Texts in Statistics.
- Manly, B. F. J. (2007): *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman & Hall, 3. udgave.
- Manski, C.F. (1993): *Identification of Endogenous Social Effects: The Reflection Problem*. The Review of Economic Studies Vol. 60, No. 3, 531–542.
- Meyer, B.D. (1995): *Natural and quasi-experiments in economics*. Journal of Business & Economic Statistics 13 (2), 151–161.
- Ministeriet for By, Bolig og Landdistrikter (2014): *Segregering i de fire største danske byområder*. København.
- McQuitty, L. L. (1960): *Hierarchical linkage analysis for the isolation of types*. Educational and Psychological Measurement, 20, 55–67.
- Rosenbaum, P.R. & Rubin, D.B. (1983): *The Central Role of the Propensity Score in Observational Studies for Causal Effects*. Biometrika, 70(1), 41–55.
- Rubin, D.B. (1974): *Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies*. Journal of Educational Psychology, 66, 688–701.
- Rubin, D.B. (1978): *Bayesian Inference for Causal Effects: The Role of Randomization*. The Annals of Statistics, 6(1), 34–58.
- Rubin, D.B. (1990): *Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies*. Statistical Science, 5(4), 472–480.
- Sibson, R. (1973): *An optimally efficient algorithm for the single-link cluster method*. Computer Journal, 16, 30–34.
- Smith, H. (1997): *Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies*. Sociological Methodology, 27, 325–353.
- Sokal, R.R. & Michener, C.D. (1958): *A statistical method for evaluating systematic relationships*. University of Kansas Science Bulletin, 38, 1409–1438.
- Sokal, R.R. & Sneath, P.H.A. (1963): *Principles of Numerical Taxonomy*. San Francisco.
- Splawa-Neyman, J. (1923): *On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9*. Statistical Science, 5(4), 465–472.

- Stuart, E.A.; Bradshaw, C.P. & Leaf, P.J. (2015): *Assessing the Generalizability of Randomized Trial Results to Target Populations*. *Prevention Science*, vol. 16, no. 3.
- Sørensen, T. J. (1948): *A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons*. *Biologiske Skrifter, Kongelige Danske Videnskabernes Selskab*.
- Transport-, Bygnings- og Boligministeriet (2018): *Liste over ghettoområder pr. 1. december 2018*. København
- Vasiljeva, K. & Mairey, I. (2016): *Klarer beboere i ghettoer sig bedre eller dårligere? KRAKA analyse*.
- Weatherall, C. D.; Boje-Kovacs, B. & Egsgaard-Pedersen, A. (2016): *Et historisk tilbageblik på de særligt udsatte boligområder udpeget i 2014 - Udviklingen i tilflyttere, fraflyttere og fastboende*. Kraks Fond Byforskning.
- Yeo, In-K. & Johnson, R. (2000): *A new family of power transformations to improve normality or symmetry*. *Biometrika*, 87, 954–959.
- Zwiers, M.; Kleinhans, R. & Van Ham, M. (2017): *The Path-Dependency of Low-Income Neighbourhood Trajectories: An Approach for Analysing Neighbourhood Change*. *Applied Spatial Analysis*, 10, 363–380.

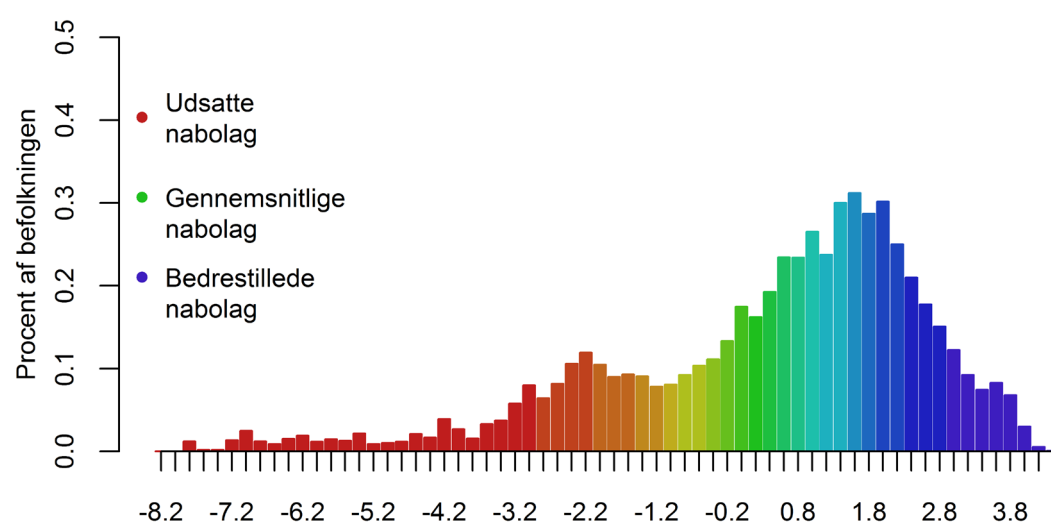
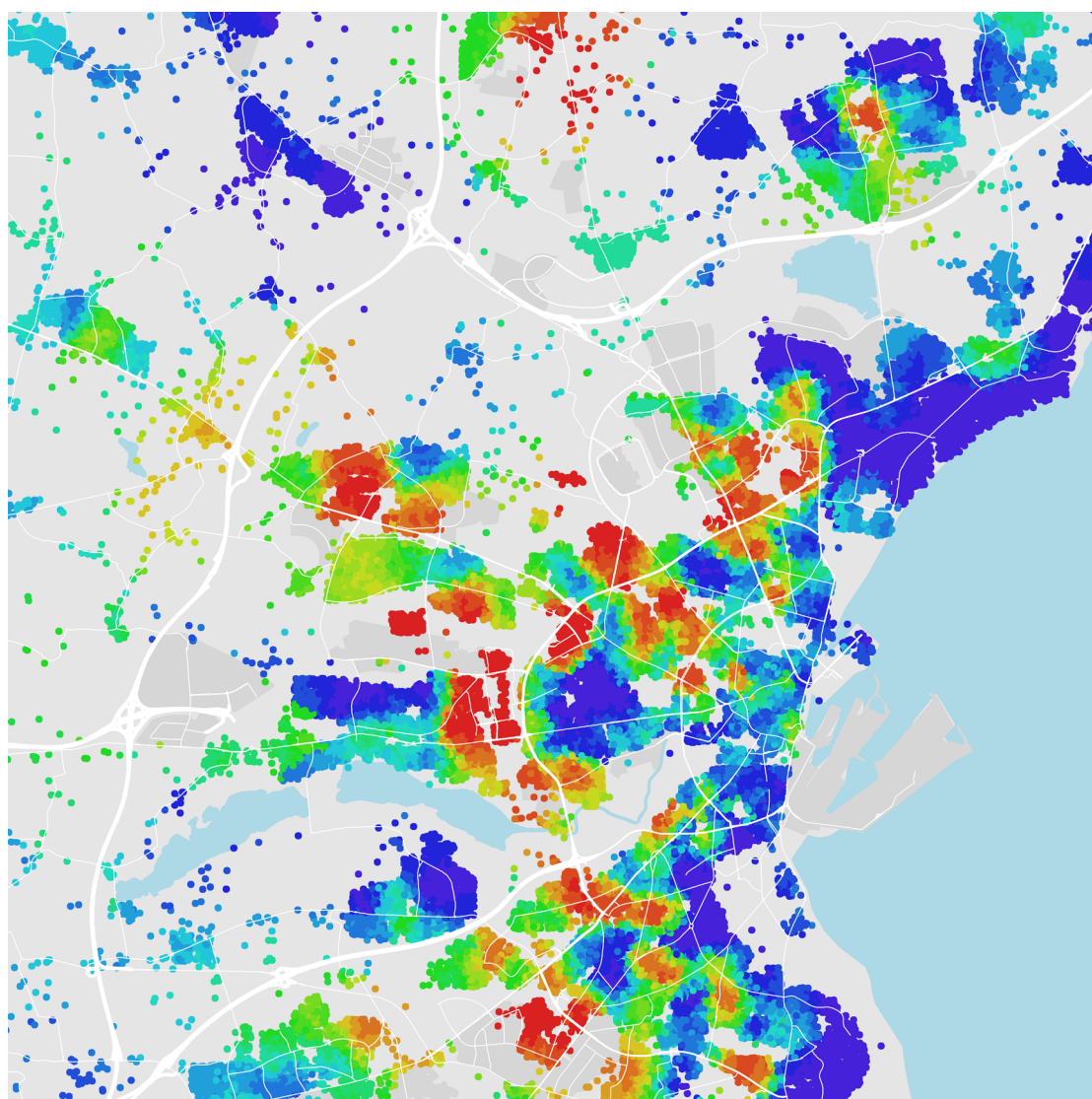
A. SNI for udvalgte kommuner

Figur A.1. Det Socioøkonomiske Nabolag-Indeks for København.



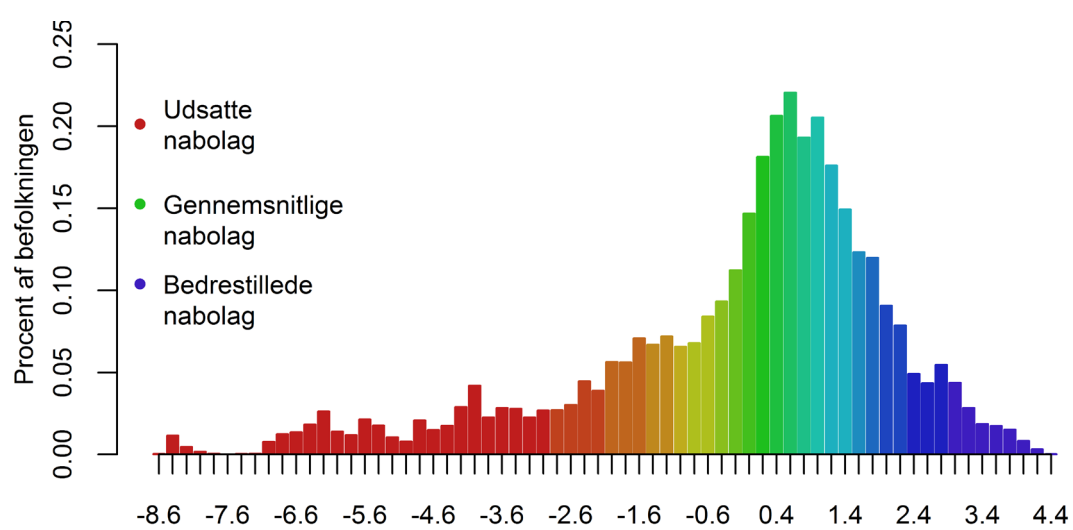
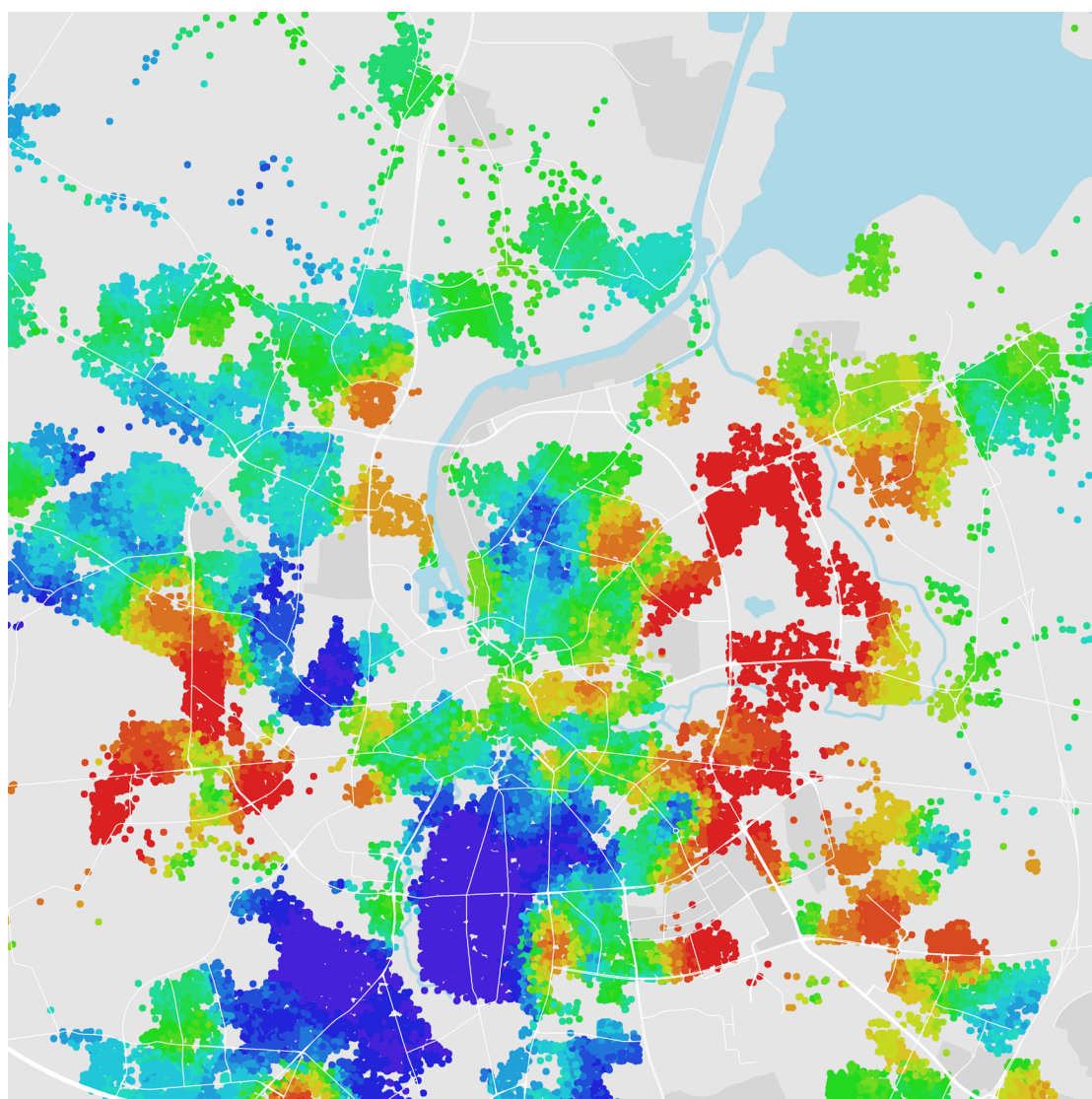
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur A.2. Det Socioøkonomiske Nabolag-Indeks for Århus.



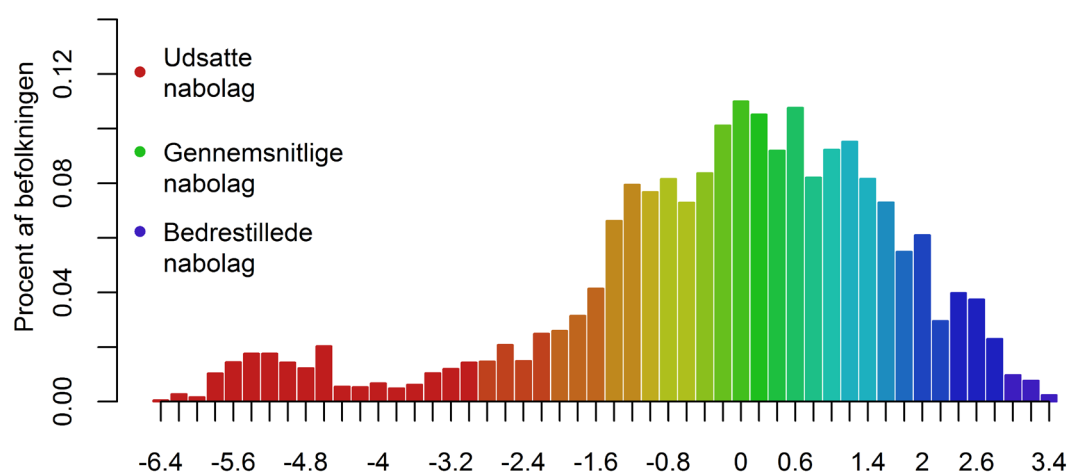
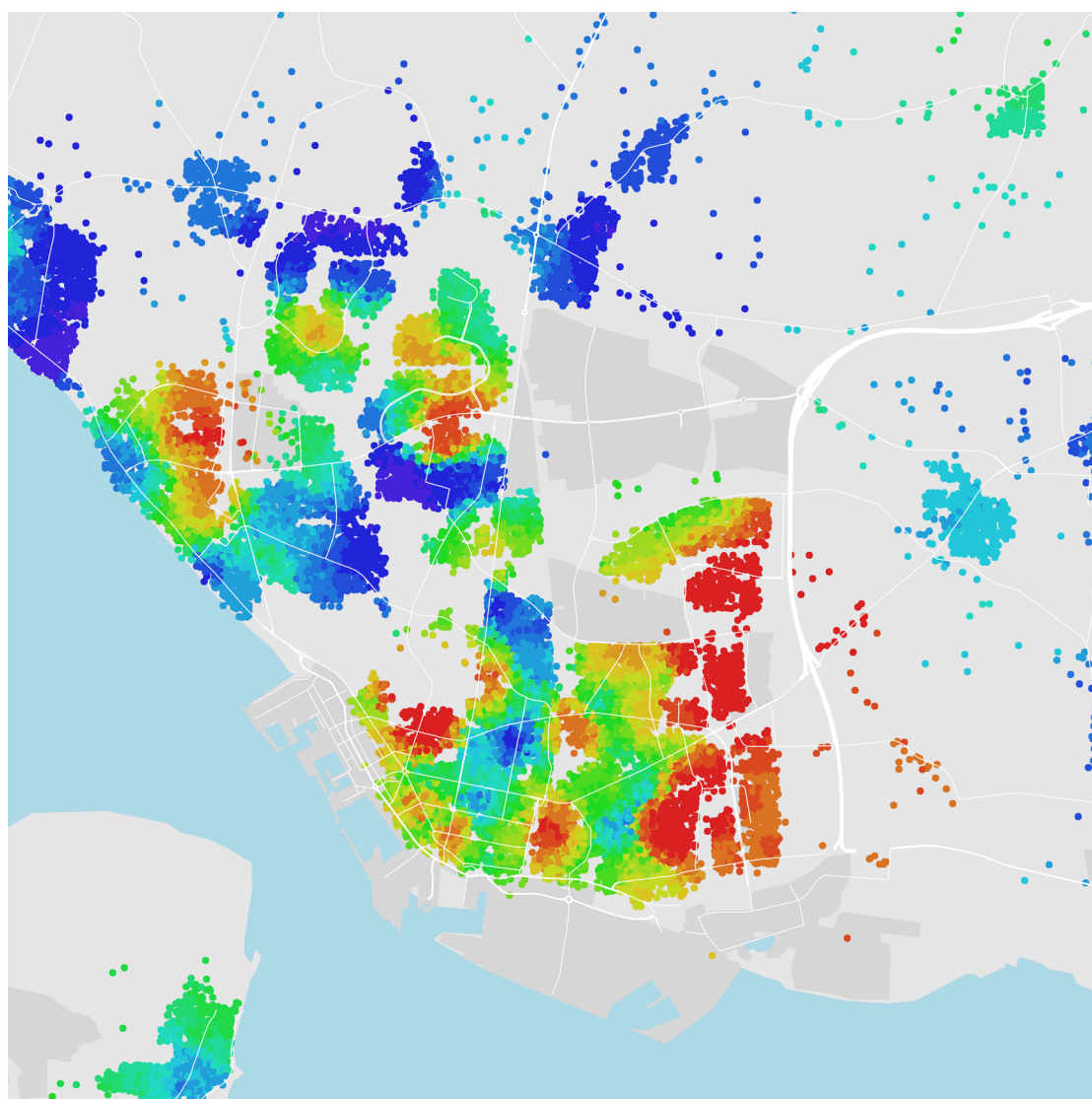
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur A.3. Det Socioøkonomiske Nabolag-Indeks for Odense.



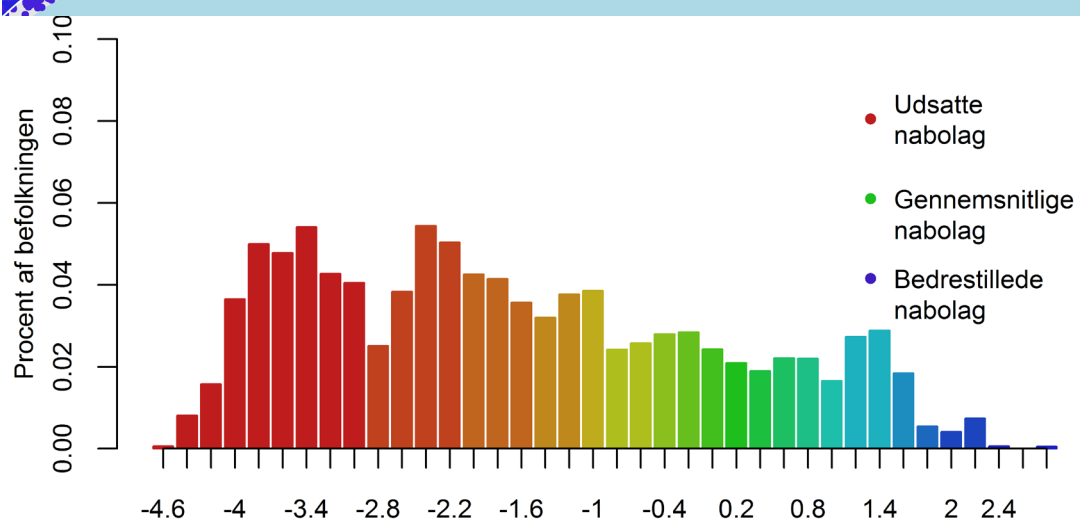
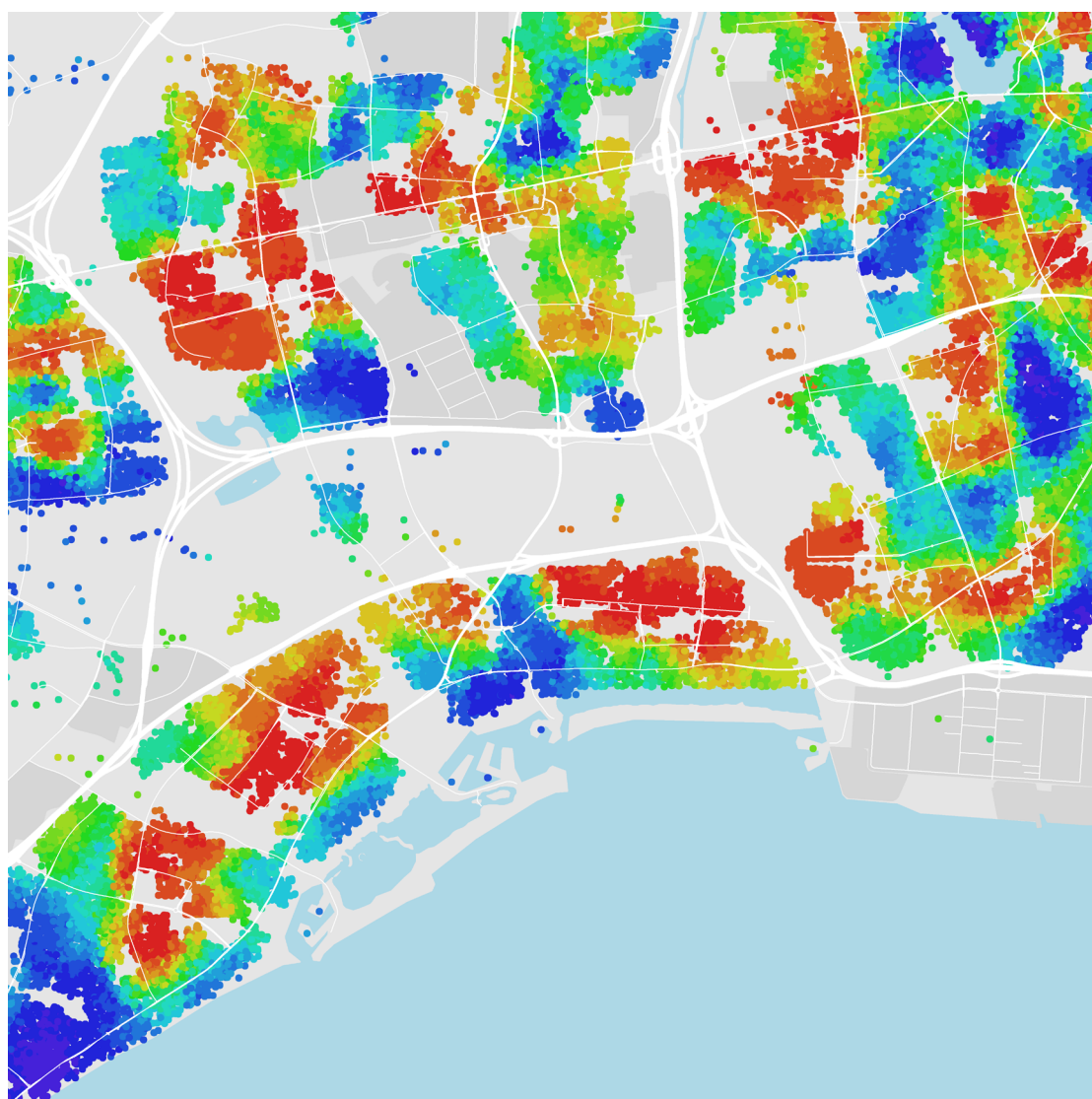
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur A.4. Det Socioøkonomiske Nabolag-Indeks for Esbjerg.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur A.5. Det Socioøkonomiske Nabolag-Indeks for Brøndby og Ishøj.



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

B. Kullback–Leibler divergensindekset

Afsnittet beskriver divergensindekset, som kan bruges til kvantificere hvor sammenklynget nabolagene er i det Socioøkonomiske Nabolag-Indeks. Afsnittet følger Roberto (2016), Stephensen (2016). Indekset er baseret på relativ entropi og er et mål, der sammenligner to diskrete sandsynlighedsfordelinger med hinanden, jf. Cover et al. (2006). Relativ entropi er også kendt som Kullback–Leibler divergence, jf. Kullback (1987). Indekset er i denne rapport konstrueret ved at sammenligne den diskrete faktiske spatiale sandsynlighedsfordeling af nabolag for en givet ventil af det Socioøkonomiske Nabolag-Indeks, med en teoretisk fuldstændig tilfældig spatial, CSR, sandsynlighedsfordeling.

Antag at der er N eksklusive områder der er konstante over t . Områderne er i denne analyse hektarceller. Lad $x_{i,t,v}$ være antallet af punkter (unikke nabolag) i område i til tidspunkt t , der tilhører SNI ventil v , og a_i være arealet af område i . Definer

$$X_{t,v} = \sum_i x_{i,t,v} \text{ og } A = \sum_i a_i$$

Ud fra data defineres den faktiske diskrete sandsynlighedsfordeling for område i til tidspunkt t for enhver SNI ventil v .

$$p_{i,t,v} = \frac{x_{i,t,v}}{X_{t,v}},$$

Dette er sandsynligheden for et bestemt nabolag fra SNI ventil v er beliggende i område i til tidspunkt t .

For punkterne fra CSR sandsynlighedsfordelingen gælder homogenitet og uafhængighed. Hvilket vil sige, at et intet punkt har en nogen præference for en hvilken som helst spatial placering, og at informationen om omkring udfaldet i et område ikke påvirker udfaldet i et andet område. Det gælder derfor at:

$$p_i^{CSR} = \frac{a_i}{A}.$$

Bemærk at denne sandsynlighedsfordeling er konstant over tid.

For diskrete sandsynlighedsfordelinger er Kullback–Leibler divergensen fra den faktiske observerede til den teoretiske sandsynlighedsfordeling defineret som³⁸

$$\begin{aligned} DI_{t,v} &= \sum_i p_{i,t,v} \log\left(\frac{p_{i,t,v}}{p_i^{CSR}}\right) \\ &= \frac{1}{X_{t,v}} \sum_i x_{i,t,v} \log\left(\frac{x_{i,t,v}}{X_{t,v}}\right) - \frac{1}{X_{t,v}} \sum_i x_{i,t,v} \log\left(\frac{a_i}{A}\right) \end{aligned}$$

Hvis $DI_{t,v} = 0$ er de to sandsynlighedsfordelinger ens. Hvilket betyder at nabolagene i den SNI ventil v til tidspunkt t er spatialt tilfældigt fordelt ud over hele Danmark.

Hvis $DI_{t,v=i} = DI_{t,v=j}$, $i \neq j$ betyder det, at nabolagene for den i 'te og den j 'te SNI ventil er sammenklynget i samme grad. Hvis derimod $DI_{t,v=i} > DI_{t,v=j}$, $i \neq j$ betyder det, at den i 'te SNI ventil er mere sammenklynget end den j 'te SNI ventil til tidspunkt t . Dette kan tolkes som, at den i 'te ventil af SNI er mere segregeret i forhold til den j 'te ventil af SNI.

³⁸ Hvor vi definerer, at $0 \log 0 = 0$, da $\lim_{x \rightarrow 0} (x \log x) = 0$.

C. Et eksempel på hierarkisk clustering processen

Dette afsnit beskriver og illustrerer hierarkisk clustering metoden, der anvendes i CNN algoritmen. Vi betragter et dataset bestående af seks punkter, et rødt, blåt, grønt, orange, pink og et sort punkt. Placeringen af punkterne ses i Figur C.1. Figuren viser også den trinvis clustings-proces og den indbyrdes euklidiske afstand mellem punkter og klynger for ethvert trin i denne proces.

Vi ønsker at inddele datasættet i homogene unikke undergrupper. I denne hierarkisk clustering analyse er det den geografiske afstand mellem punkterne og klynger af punkter, der afgør om punkterne skal grupperes til en ny større klynge.

Den hierarkiske clustering algoritme kører som en bottom-op proces, som starter med, at alle punkterne er i hver sin klynge. Dette er illustreret som step 1 i Figur C.1. Sekventielt kombineres hver klynge til større og større klynger indtil, at alle punkterne er i samme klynge. Dette illustreret som step 6 i Figur C.1. I hvert skridt kombineres de to klynger, der er adskilt med den korteste afstand. Fra step 1 til step 2 er dette afstanden mellem det grønne og det sorte punkt jf. afstandsmatricen i step 1.

I step 2 er det afstanden mellem det pink og det blå punkt, der er den korteste afstand. Men for kunne bestemme dette, har algoritmen udregnet en ny afstandsmatrice, hvor det i stedet for afstanden mellem alle punkter nu er afstanden mellem punkterne, rød, blå, orange og pink og klyngen bestående af det grønne og det sorte punkt.

I hierarkiske clustering analyse er der en del forskellige måder at måle afstanden af forbindelsen (linkage) mellem et punkt og en klynge på. De mest almindelig anvendte linkage er: Complete, Single og Average. I dette eksempel og i CNN-algoritmen er forbindelsen valgt til Average-linkage.

- Average-linkage, Sokal et al. (1958) og Sokal et al. (1963), med denne forbindelse udregnes alle afstande mellem alle punkter i en klynge til alle andre punkter i en anden klynge, og forbindelses-afstanden er så gennemsnittet af dem:

$$D(A, B) = \frac{1}{|A| \cdot |B|} \sum_{a \in A} \sum_{b \in B} d(a, b)$$

- Single linkage, Florek et al (1951) og Sibson (1973), her er forbindelse mellem to klynger den korteste afstand mellem to punkter i hver deres klynge:

$$D(A, B) = \min_{a \in A, b \in B} d(a, b)$$

- Complete linkage, Sørensen (1948), McQuitty (1960) og Defays (1977), her er forbindelsen mellem to klynger den største afstand mellem to punkter i hver deres klynge:

$$D(A, B) = \max_{a \in A, b \in B} d(a, b)$$

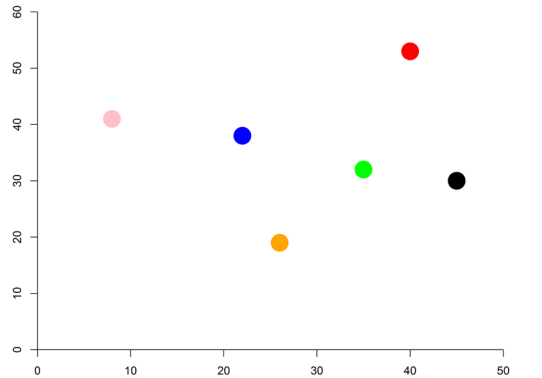
Hvor $D(A, B)$ er afstanden mellem klyngerne A og B og hvor $d(a, b)$ er afstanden mellem punkterne $a \in A$ og $b \in B$.

Det er vigtigt at understrege, at valget af linkage har betydning for, hvordan de enkelte punkter grupperes, jf. Figur C.1. Figur C.2 viser dendrogrammer for complete, single og average linkage. Et dendrogram opsummerer clustering processen, nederst i dendrogrammet er alle 6 punkter i hver deres klynge øverst er alle punkter i samme klynge.

På y-aksen aflæses minimumsafstanden mellem punkterne og klynger for at danne en ny klynge.

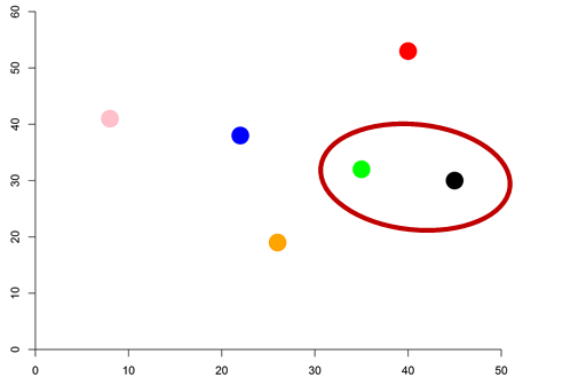
Figur C.1. Hierarkisk clustering algoritme med average linkage – bottom-up.

Step 1) mindste afstand er 10



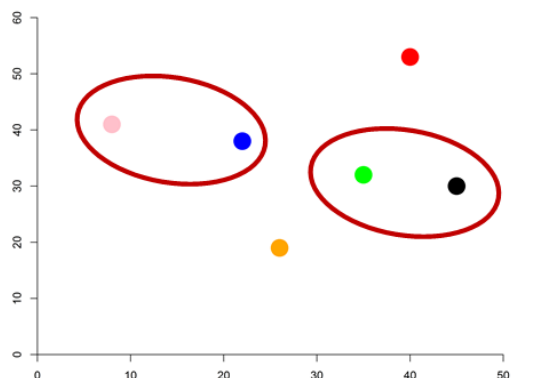
	Rød	Blå	Grøn	Orange	Pink	Sort
Rød	-	23.4	21.6	36.8	34.2	23.5
Blå	23.4	-	14.3	19.4	14.3	24.4
Grøn	21.6	14.3	-	15.8	28.5	10.2
Orange	36.8	19.4	15.8	-	28.4	22
Pink	34.2	14.3	28.5	28.4	-	38.6
Sort	23.5	24.4	10.2	22	38.6	-

Step 2) mindste afstand er 14



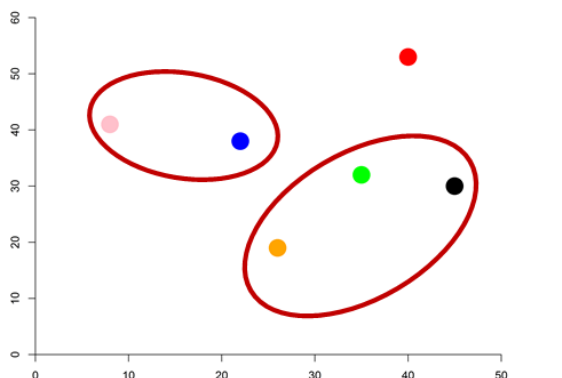
	Rød	Blå	Orange	Pink	Grøn-Sort
Rød	-	23.4	36.8	34.2	22.6
Blå	23.4	-	19.4	14.3	19.4
Orange	36.8	19.4	-	28.5	18.9
Pink	34.2	14.3	28.5	-	33.6
Grøn-Sort	22.6	19.4	18.9	33.6	-

Step 3) mindste afstand er 19



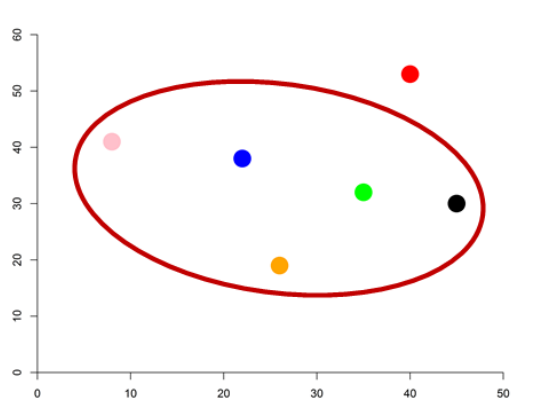
	Rød	Orange	Blå-Pink	Grøn-Sort
Rød	-	36.8	28.8	22.6
Orange	36.8	-	24.0	18.9
Blå-Pink	28.8	24.0	-	26.5
Grøn-Sort	22.6	18.9	26.5	-

Step 4) mindste afstand er 26

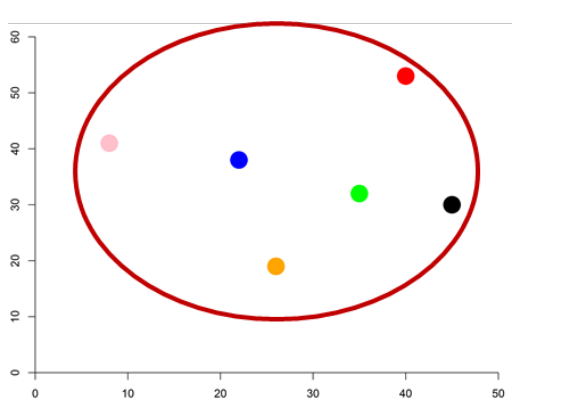


	Rød	Blå-Pink	Grøn-Sort-Orange
Rød	-	28.8	27.3
Blå-Pink	28.8	-	25.6
Grøn-Sort-Orange	27.3	25.6	-

Step 5) mindste afstand 38

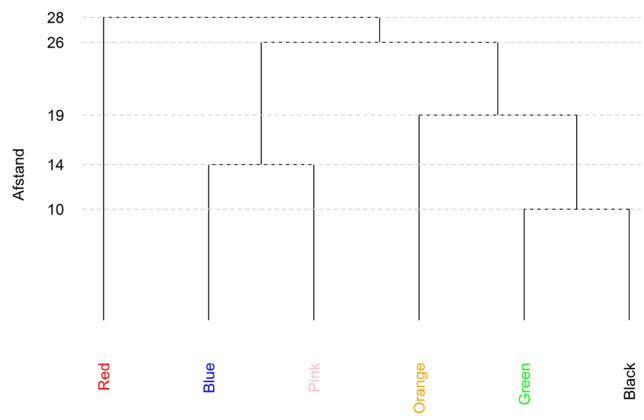
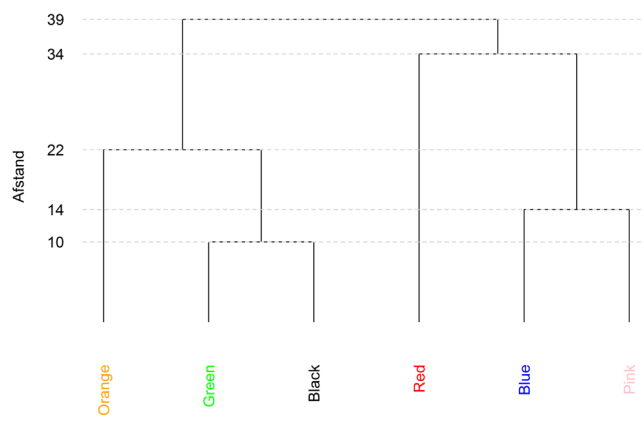
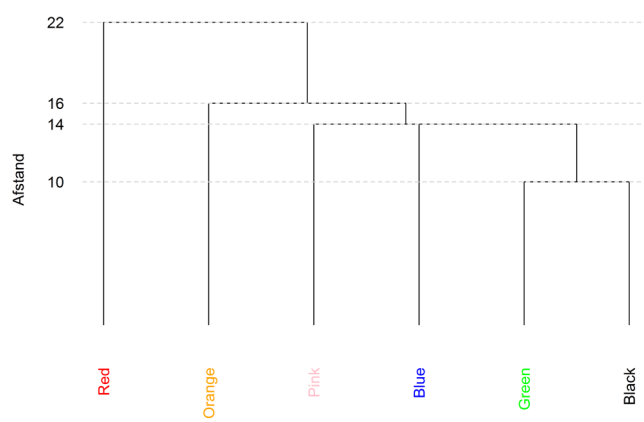


Step 6)



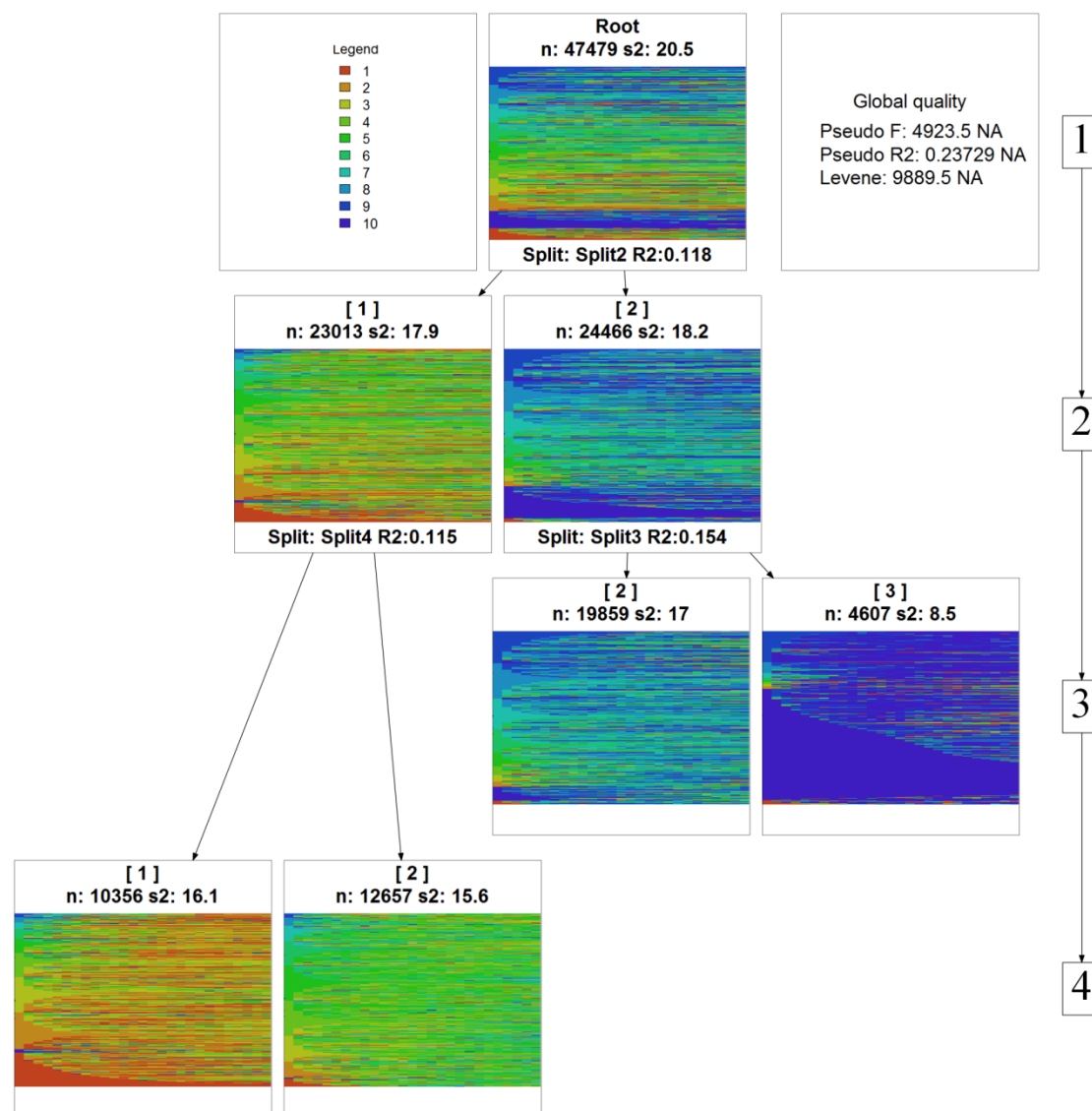
Kilde: Egne beregninger.

Figur C.2. Dendrogrammer for Average, Complete, Single og Ward-linkage.

a) *Average-linkage*b) *Complete-linkage*c) *Single-linkage*Kilde: *Egne beregninger.*

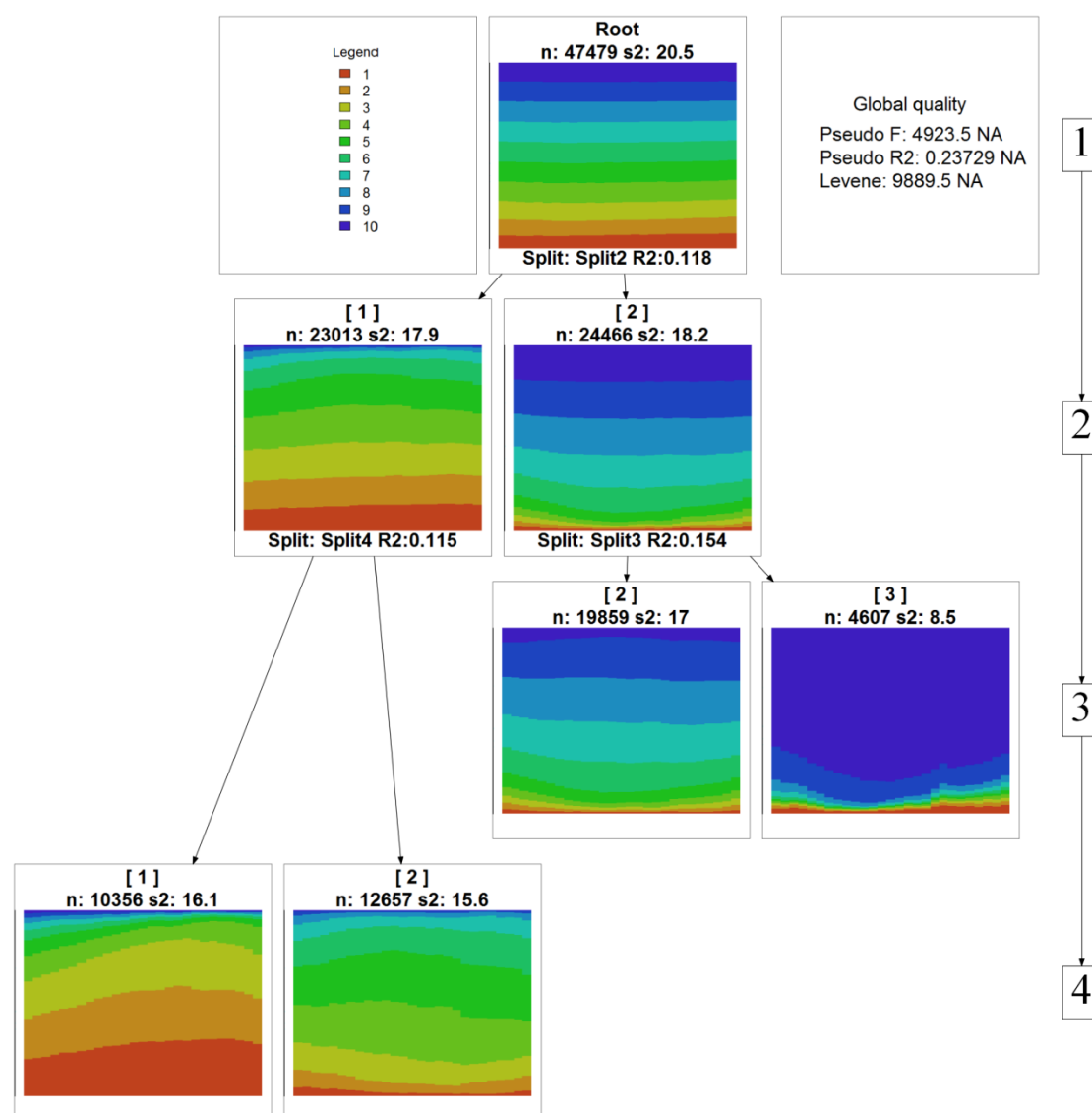
D. Gruppering af indkomstforløb og analyse af individuelle indkomstforløb

Figur D.1. Trinvis binær opdeling af målgruppens indkomstforløb på fire grupper ved hierarkisk clustering og average linkage, individuelle forløb



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik ved anvendelse af Traminer og Graphviz til R.

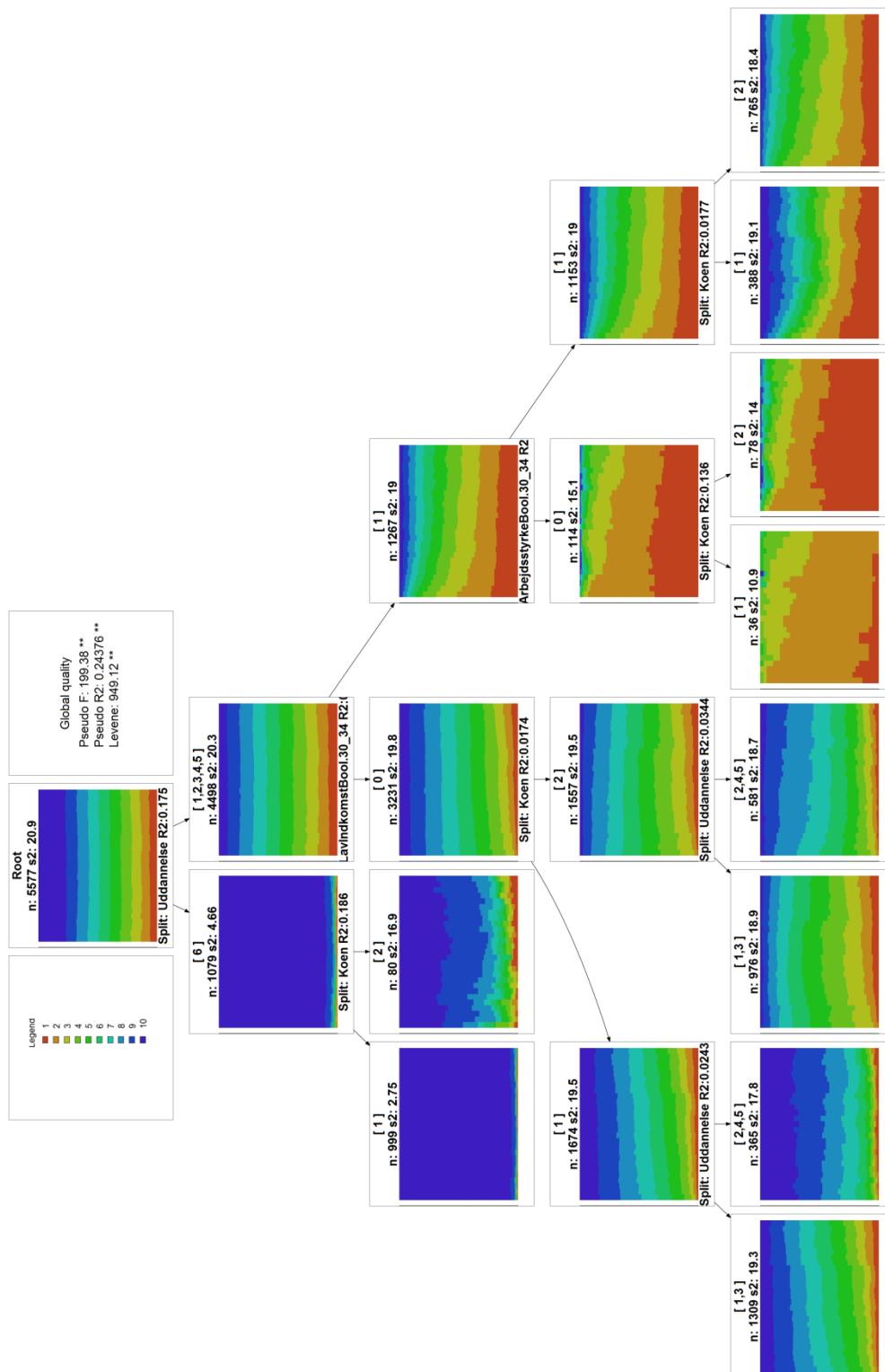
Figur D.2. Trinvis binær opdeling af målgruppens indkomstforløb på fire grupper ved hierarkisk clustering og average linkage, fordeling



Anm. Indkomstfordelingen i slutgrupperne er gengivet i Figur 4.13.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik ved anvendelse af Traminer og Graphviz til R.

Figur D.3. Regression Tree for indkomstforløbenes tilknytning til køn, oprindelse, uddannelsesniveau som 35-årig samt mindst et års ophold i de to nederste indkomstdeciler i alderen 30-34 år og mindst et års tilknytning til arbejdsmarkedet i samme aldersgruppe



Anm. Baseret på 10 procent stikprøve af målgruppen. Køn: 1=Mænd. Uddannelse: 1=Grundskole, 2=Gymnasial, 3=Erhvervsfaglig, 4=KVU, 5=MVU, 6=LVU.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

E. Statistisk analyse af gruppetilhørsforhold

Tabel E.1. Multinomial logistisk regression af individkarakteristika på gruppetilhørsforhold. Log odds. Model uden interaktion.

	Gruppe 2 vs. gruppe 1	Gruppe 3 vs. gruppe 1	Gruppe 4 vs. gruppe 1
Indvandrere, ikke-vestlig (ref. dansk)	-0,846*** (-0,104)	-1,006*** (-0,122)	-0,760*** (-0,286)
Indvandrere, vestlig	-0,078 (-0,11)	0,121 (-0,115)	0,608*** (-0,186)
Efterkommer, ikke-vestlig	-0,07 (-1,433)	1,3 (-1,237)	-7,267*** (-0,001)
Efterkommer, vestlig	0,075 (-0,396)	0,499 (-0,374)	0,674 (-0,502)
Kvinde	0,041 (-0,031)	-0,964*** (-0,03)	-2,383*** (-0,056)
Gymnasial (ref. grundskole)	0,256*** (-0,07)	1,465*** (-0,066)	2,443*** (-0,095)
Erhvervsfaglig	0,270*** (-0,032)	0,803*** (-0,033)	0,793*** (-0,063)
KVU	0,398*** (-0,093)	1,640*** (-0,087)	1,904*** (-0,128)
MVU	0,552*** (-0,062)	2,529*** (-0,057)	3,112*** (-0,083)
LVU	-1,309*** (-0,213)	1,926*** (-0,125)	4,436*** (-0,135)
Mindst 1 år i de to laveste indkomstdeciler, 30-34 år	-1,060*** (-0,029)	-1,794*** (-0,032)	-1,674*** (-0,061)
Mindst 1 år i almen bolig, 30-34 år	-0,230*** (-0,037)	-0,362*** (-0,038)	-0,894*** (-0,076)
Mindst 1 år i et parforhold, 30-34 år	0,427*** (-0,041)	0,702*** (-0,041)	1,175*** (-0,077)
(a) Mindst 1 år i arbejdsstyrken, 30-34 år	1,073*** (-0,093)	1,445*** (-0,142)	1,476*** (-0,396)
(b) Mindst 1 år i et udsat område, 30-34 år	-0,121** (-0,058)	-0,164*** (-0,062)	-0,595*** (-0,134)
Konstant	-0,783*** (-0,102)	-0,888*** (-0,147)	-3,124*** (-0,403)

Anm.: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standardafvigelse er angivet i parentes. Forløb med ukendt boligtype (48 stk.) er frasorteret i estimationen.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

F. Balancering af data ved propensity score matching

I undersøgelsen foretages balancering af kontrolgruppen ved *propensity score matching*. Dette beskrives i afsnit 2.1, mens dette appendiks indeholder estimationsresultater ved beregning af sandsynlighedsværdier samt en beskrivelse af den anvendte matching metode.

F.1 Propensity score

Sandsynligheden for at blive bosat i et udsat boligområde, propensity scoren, estimeres ved en logit-model. Tabel F.1 viser estimationsresultatet.

Tabel F.1. Estimation af sandsynlighed for som flygtning at blive placeret i et udsat boligområde ved logit-model.

	Før matching			Efter matching	
	Estimat	Std. fejl		Estimat	Std. fejl
Konstant	-1.42	0.11	***	0.09	0.13
Køn, ref. kvinder	-0.05	0.02	*	0.02	0.03
Familietype, ref. enlig	0.17	0.03	***	0.03	0.03
Antal børn	0.03	0.01	***	-0.01	0.01
Alder	0.00	0.00	**	0.00	0.00
<i>Indvandrings år, ref. 1998</i>					
1986	-0.35	0.06	***	-0.07	0.08
1987	-0.31	0.07	***	-0.11	0.09
1988	-0.27	0.07	***	-0.13	0.09
1989	-0.19	0.07	**	-0.10	0.09
1990	-0.18	0.08	*	-0.03	0.09
1991	0.21	0.07	**	-0.02	0.08
1992	0.44	0.06	***	-0.13	0.08
1993	0.24	0.07	***	-0.05	0.08
1994	0.44	0.07	***	-0.15	0.08
1995	0.26	0.06	***	-0.03	0.07
1996	-0.23	0.06	***	-0.11	0.08
1997	0.02	0.07		-0.06	0.08
<i>Uddannelsesniveau, ref. grundskole inkl. uoplyst uddannelse</i>					
Erhvervsfaglig	0.11	0.03	**	0.02	0.04
Gymnasial	0.05	0.04		-0.01	0.05
Kort videregående	0.01	0.06		-0.05	0.08
Mellemlang videregående	0.07	0.05		0.00	0.06
Lang videregående (inkl. Ph.d.)	-0.01	0.06		0.01	0.08
<i>Oprindelsesland</i>					
Afrika	0.11	0.09		-0.02	0.12
Mellemøsten	0.19	0.09	*	-0.04	0.11
Balkanhalvøen	-0.36	0.09	***	-0.07	0.12
Asien	-0.07	0.10		-0.01	0.12

Anm.: Signifikansniveau: **** 0,001, *** 0,01 og ** 0,05.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Før matching ses det, at stort set alle kovariaterne er signifikant bestemt. I det ubalancerede datasæt har flygtninge fra Afrika for eksempel en signifikant højere sandsynlighed for på ankomsttidspunktet at blive placeret i et udsat boligområde.

Sandsynligheden for at blive placeret i et udsat boligområde kan nu estimeres for hver enkelt flygtning ved anvendelse af den estimerede logit-model. På baggrund af disse sandsynligheder dannes en kontrolgruppe, som ud fra de anvendte baggrundskarakteristika er lig indsatsgruppen (sker ved matching, se følgende afsnit F.2).

Tabel F.1 viser desuden resultatet af den samme logit-model estimeret på det balancerede datasæt. Her ses, at alle kovariater er insignifikante. Dette indikerer, at data efter matching er balanceret.

F.2 Valg af matching-algoritme

I denne analyse er det valgt at matche med nearest neighbour algoritmen (NN). Dette er den mest almindeligt brugte algoritme. Ved *Nearest Neighbour Matching* findes et match ved den korteste afstand af propensity score mellem personer i indsats- og kontrolgruppen. Der er flere forskellige versioner af Nearest Neighbour Matching, for eksempel kan metoden anvendes med eller uden tilbagelægning af personer i kontrolgruppen og med ét eller flere match i kontrolgruppen.

Hvis matching sker med tilbagelægning, kan én person fra kontrolgruppen, der er underrepræsenteret i forhold til indsatsgruppen medtages flere gange. I vort tilfælde vil en flygtning, der er allokeret til et ikke-udsat boligområde og hvis karakteristika ved indvandring er underrepræsenteret, kunne matches flere gange med personer fra et udsat boligområde med samme karakteristika. Det er en afvejning mellem bias og varians, der afgør, om matching skal ske med eller uden tilbagelægning. Tillades tilbagelægning, bliver matchet bedre, men estimatets varians øges. Derfor vælges Nearest Neighbour Matching med tilbagelægning ofte, hvis der er stor forskel mellem fordelingen af propensity score i kontrolgruppen og i indsatsgruppen, eller hvis kontrolgruppen er meget lille. Dette er ikke tilfældet for dette eksperiment, hvorfor Nearest Neighbour Matching sker uden tilbagelægning (jf. Figur 5.2 er der common support før matching).

Nearest Neighbour Matching kan også ske, hvor der findes mere end ét match i kontrolgruppen. Her foretages der ligeledes en afvejning mellem bias og varians, idet mere end ét match øger bias, mens variansen reduceres, jf. Smidt (1997). Da indsatsgruppen i dette eksperiment er forholdsvis stor, vælges der at lave et én-til-én match.

Nearest Neighbour Matching kan resultere i et relativt dårligt match, hvis afstanden målt ved propensity score til nærmeste nabo er stor. Dette kan løses med valget af en alternativ algoritme som for eksempel Caliper matching, der dog reducerer antallet af mulige match. Det vil sige, at det igen foretages en afvejning mellem bias og usikkerhed på estimatet. Caliper matching er ikke valgt i denne analyse som følge af det store overlap af propensity score mellem kontrolgruppen og indsatsgruppen, jf. Figur 5.2.

G. Rubins kausale model

Rubins kausale model er baseret på idéen om potentielle udfald. Modellen beskriver den kausale effekt mellem potentielle udfald under indsats (behandling) og kontrol (ingen behandling), men hvor kun et af udfaldene er observeret for hver person, jf. Holland (1986), Splawa-Neyman (1923) og Rubin (1974, 1978, 1990). Herved er indsatseffekten defineret som forskellen mellem et observeret udfald og det ikke-observerede kontrafaktiske udfald.

Lad $Y_{i,1}$ være det potentielle udfald for person i , hvis personen er i indsatsgruppen, og lad $Y_{i,0}$ være det potentielle udfald for den samme person, hvis personen er i kontrolgruppen. Vi kan nu definere indsatseffekten som:

$$\tau_i = Y_{i,1} - Y_{i,0}$$

For at beregne indsatseffekten opstår et manglende dataproblem, da $Y_{i,1}$ og $Y_{i,0}$ aldrig er observeret på samme tid for samme person i . Person i er enten i indsatsgruppen eller i kontrolgruppen. Dette betyder, at vi ikke kan måle indsatseffekten på personniveau, men kun på populationsniveau. Dette sker ved at sammenligne det gennemsnitlige udfald i indsatsgruppen med det gennemsnitlige udfald i kontrolgruppen og herved finde den gennemsnitlige indsatseffekt.

Lad T_i være en indikator for indsats, som har værdien 1, hvis person i er i indsatsgruppen og værdien 0, hvis personen er i kontrolgruppen. Det observerede udfald for enhver person i kan nu opstilles som:

$$Y_i = T_i Y_{i,1} + (1 - T_i) Y_{i,0}$$

Rubins kausale model viser, at hvis allokeringen til indsatsgruppen og kontrolgruppen er tilfældig (dvs. eksperimentet er randomiseret), i så fald er den gennemsnitlige indsatseffekt lig med forskellen mellem det gennemsnitlige udfald for personer i indsatsgruppen og det gennemsnitlige udfald for personer i kontrolgruppen:

$$ATE = E[Y_{i,1} - Y_{i,0}] = E[Y_{i,1} | T_i = 1] - E[Y_{i,0} | T_i = 0]$$

Dette resultat gælder kun, når der er tale om et randomiseret eksperiment, hvor allokeringen til hver af grupperne er tilfældig og uafhængig af de potentielle udfald $\{Y_{i,1}, Y_{i,0} \perp T_i\}$. Ellers kan der opstå selektionsbias. Ved at balancere data i afsnit 2.1 har vi konstrueret en approksimation for et randomiseret eksperiment, og ovenstående resultater fra Rubins kausale model antages at være gældende for undersøgelsens estimationsresultater.

H. Betinget effekt på erhvervsfrekvens af placering i udsat område

For at undersøge effekten af placering i udsatte boligområder for delgrupper af den samlede gruppe af flygtninge opdeles den samlede effekt præsenteret i afsnit 5.3 efter alder, køn, oprindelsesland og uddannelsesniveau på indvandringstidspunktet.

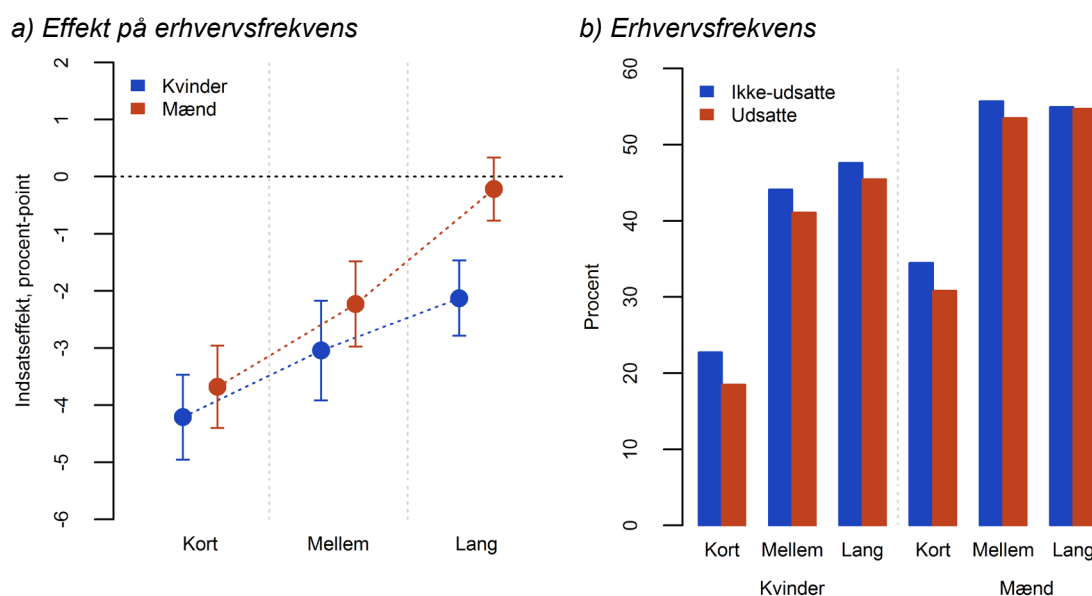
Målt på erhvervsfrekvens blev der overordnet set påvist en statistisk signifikant negativ effekt af at blive placeret i et udsat boligområde, dvs. flygtninge placeret i et udsat boligområde har lavere erhvervsdeltagelse end flygtninge i kontrolgruppen, jf. afsnit 5.3.

I dette delafsnit opdeles den totale effekt efter køn, alder ved ankomst (barn eller voksen), oprindelsesland og uddannelsesniveau ved indrejse.

Effekt på erhvervsfrekvens efter køn

I de betragtede grupper har mænd generelt højere erhvervsdeltagelse end kvinder, jf. Figur H.1b. For begge køn stiger erhvervsdeltagelsen i takt med, at flygtningene opholder sig længere tid i landet. Kort efter indrejse er kvindernes erhvervsfrekvens omkring 20 procent, mens den for mænd er omkring 30 procent. På mellemlangt sigt er erhvervsfrekvensen for henholdsvis kvindelige og mandlige flygtninge omkring 40 procent og 55 procent. På langt sigt stiger kvindernes erhvervsfrekvens til lige under 50 procent, mens mandlige flygtninges erhvervsfrekvens forbliver stabilt på omkring 55 procent. Til sammenligning er erhvervsfrekvensen for den samlede befolkning i 2017 på 76 og 73 procent for henholdsvis mænd og kvinder, mens de tilsvarende tal for beboere i et udsat boligområde er 59 og 50 procent.

Figur H.1. Effekt af placering i udsat boligområde på erhvervsfrekvens efter køn.



Anm.: Konfidensbånd er indtegnet omkring estimatet for indsatseffekten og viser, om estimatet er signifikant på et 5 procent signifikansniveau. Estimatet er signifikant forskelligt fra nul, hvis konfidensbåndet ikke overlapper andenaksen ved værdien nul (den stiplede linje).

På førsteaksel vises opholdstid, som er inddelt i tre grupper: Kort sigt (0-6 års ophold), mellemlangt sigt (6-12 års ophold) og langt sigt (mindst 12 års ophold).

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Som det var tilfældet for den totale flygtningebestand ses en tendens til, at erhvervsfrekvensen for begge køn er højere for kontrolgruppen end for flygtninge placeret i udsatte boligområder, jf. Figur H.1b. Dette gælder ligeledes statistisk set, idet der estimeres en signifikant negativ effekt for flygtninge af begge køn og uanset opholdstid, jf. Figur H.1a. Dog kan man på langt sigt ikke konkludere, at erhvervsfrekvensen for mænd er forskellig for flygtninge placeret i et udsat og et ikke-udsat område. Det ses herved, at den langsigtede signifikante negative totale effekt i Figur H.1a alene skal tilskrives den langsigtede effekt på kvinders erhvervsfrekvens.

Generelt har det at blive placeret i et udsat boligområde en større negativ effekt på erhvervsdeltagelsen for kvinder end for mænd, dog overlapper konfidensbåndene for de to køn på kort og mellemlangt sigt. På langt sigt har kvinder en signifikant lavere erhvervsfrekvens som en konsekvens af den initiale placering i et udsat boligområde.

Effekt på erhvervsfrekvens efter oprindelsesland

Der er i gennemsnit forskel på erhvervsdeltagelsen for de flygtninge med forskelligt oprindelsesland, der i Figur H.2 er inddelt i verdensdele. Flygtninge fra Afrika og Mellemøsten har lavere erhvervsfrekvens end flygtninge fra Asien og Balkanhalvøen. Efter ophold i Danmark i mindst 12 år er der i størrelsesordenen 20 procentpoint forskel i erhvervsfrekvensen, jf. Figur H.2b samt d.

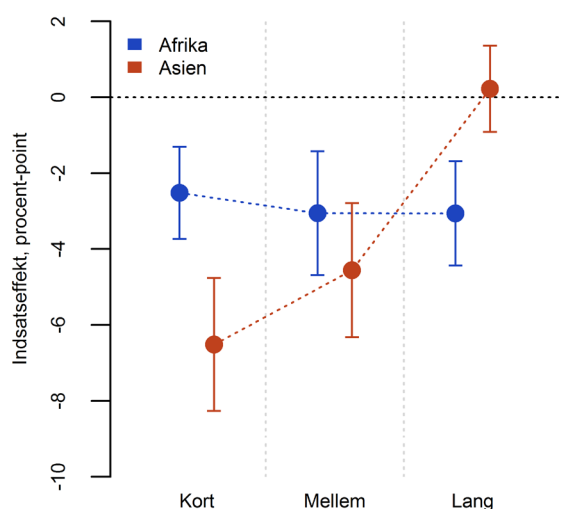
For landegrupperne med høj erhvervsfrekvens – flygtninge fra Balkan og Asien – findes på kort sigt en stor negativ effekt af initialt at blive placeret i et udsat boligområde, jf. Figur H.2a og c. For asiatiske flygtninge er effekten på -6,5 procentpoint og flygtninge fra Balkanhalvøens erhvervsfrekvens i indsatsgruppen er -6,0 procentpoint lavere end kontrolgruppen. På mellemlangt mindskes den negative effekt, og effekten er ikke signifikant forskellig fra nul på langt sigt..

For de to øvrige grupper med relativt lav erhvervsdeltagelse – flygtninge fra Afrika og Mellemøsten – findes effekten mere konstant over opholdstid. For grupperne gælder, at flygtninge placeret i udsatte boligområder har en erhvervsfrekvens, der er cirka 2-3 procentpoint lavere end erhvervsfrekvensen for flygtninge i kontrolgruppen.

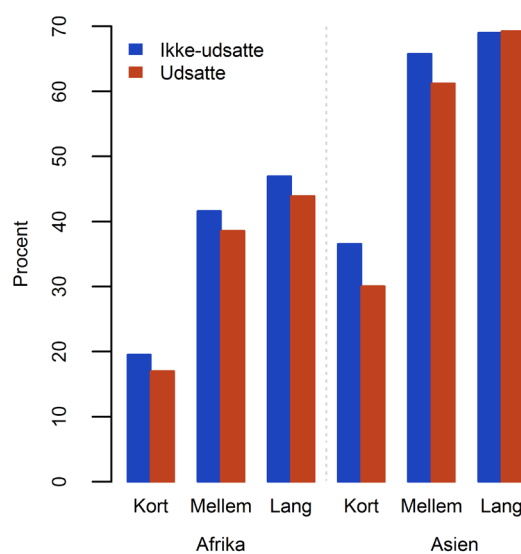
Overordnet set var effekten af placering i et udsat boligområde på kort sigt negativ, jf. Figur H.1a. Den negative effekt på kort sigt forklares af et stort negativt bidrag fra flygtninge fra Balkanhalvøen og Asien. På langt sigt er det derimod i overvejende grad flygtninge fra Afrika, der påvirker den overordnede effekt mest i negativ retning.

Figur H.2. Effekt af placering i udsat boligområde på erhvervsfrekvens efter oprindelsesland.

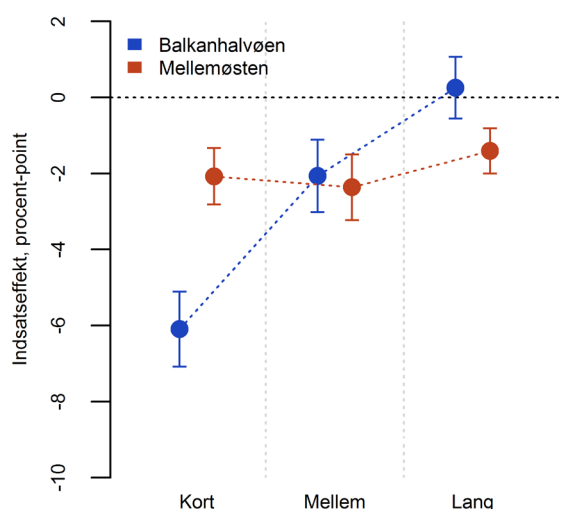
a) Effekt på erhvervsfrekvens, Afrika og Asien



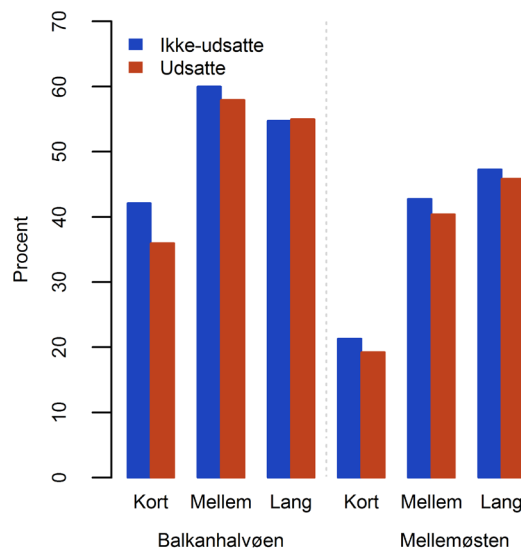
b) Erhvervsfrekvens, Afrika og Asien



c) Effekt på erhvervsfrekvens, Balkan og Mellemøsten



d) Erhvervsfrekvens, Balkan og Mellemøsten



Anm.: Konfidensbånd er indtegnet omkring estimatet for indsatsseffekten. Se i øvrigt anm. til Figur H.1.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Effekt på erhvervsfrekvens efter voksen eller barn ved indvandring

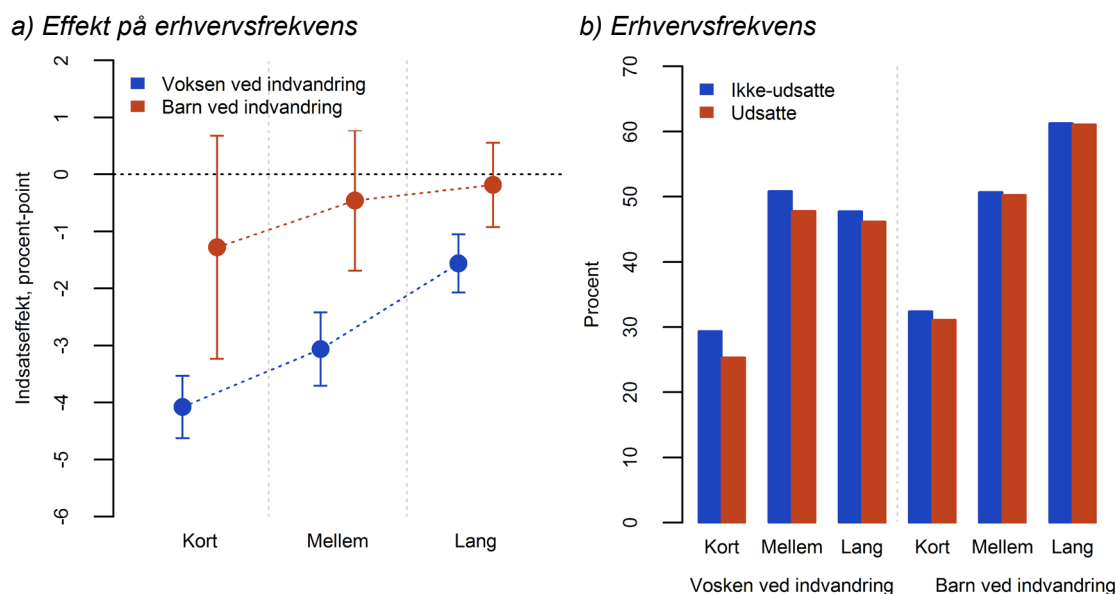
Flytningene opdeles efter, om de ved ankomst til Danmark var børn (0-15 år gamle) eller voksne (16 år og opefter). For disse to grupper betragtes udviklingen i erhvervsdeltagelsen efter opholdstid i landet, jf. Figur H.3b. Der ses en tendens til større erhvervsfrekvens blandt flygtninge i kontrolgruppen i forhold til flygtninge placeret i udsatte boligområder.

Der findes en statistisk signifikant negativ effekt for flygtninge, som kom hertil som voksne, jf. Figur H.3a. Effekten er størst mens personerne har opholdt sig kort tid i landet, hvorefter effekten mindskes, i takt med opholdstiden øges.

For flygtninge, som ankom som børn, ses ingen signifikant effekt af initialt at blive allokeret til et udsat boligområde. Og flygtninge, der indvandrer som børn, har en erhvervsfrekvens, der på langt sigt er cirka 10 procentpoint højere end flygtninge indvandret som voksne.

Det skal bemærkes, at kun et lille antal flygtninge, der indvandrer som børn, når at komme i den erhvervsduelige alder på kort sigt. For eksempel er det kun flygtninge på 15 år, der efter et års ophold i Danmark er i den arbejdsduelige alder. Dette afspejles i usikkerheden på estimatet af effekten, jf. Figur H.3a. Estimerne på kort og mellemlangt sigt skal derfor fortolkes med forsigtighed.

Figur H.3. Effekt af placering i udsat boligområde på erhvervsfrekvens efter alder.



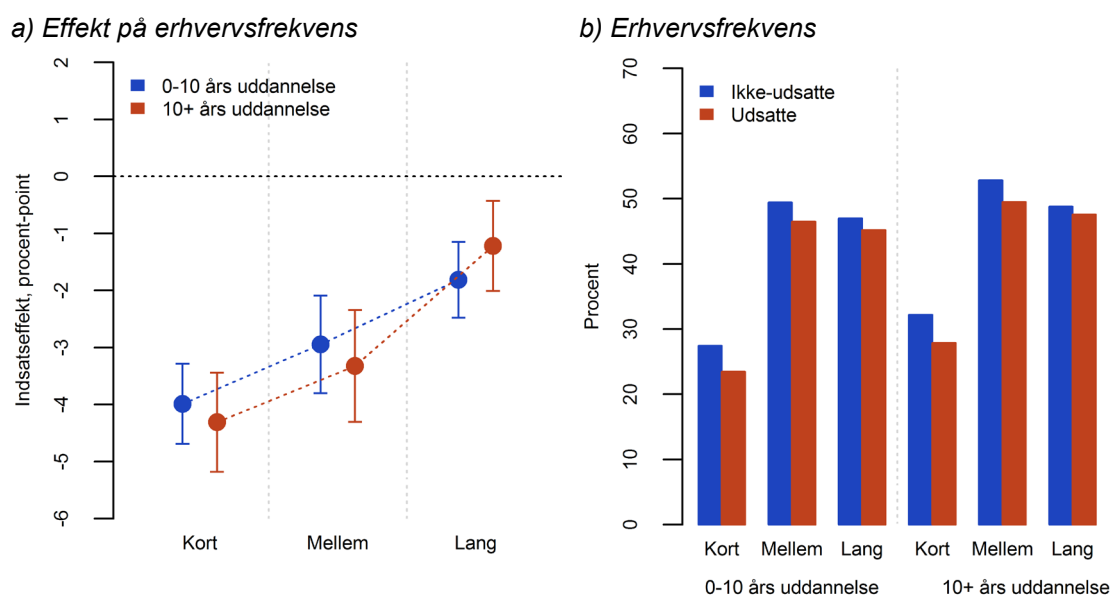
Anm.: Konfidensbånd er indtegnet omkring estimatet for indsatseffekten. Se i øvrigt anm. til Figur H.1.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Effekt på erhvervsfrekvens efter uddannelsesniveau på indvandringstidspunkt

Flygtnings uddannelsesniveau opgøres ved indvandringstidspunktet og inddeles i to grupper efter ufaglærte og faglærte. Førstnævnte gruppe har grundskolen som højest fuldførte uddannelse eller uoplyst uddannelsesniveau (dvs. gruppen har fuldført 0-10 års skolegang). Faglærte er personer, som ved indrejse har fuldført en uddannelse udover grundskolen (dvs. gruppen har mindst 10 års uddannelse).

Der findes ikke nogen signifikant forskellig effekt for flygtninge med forskelligt uddannelsesniveau, jf. Figur H.4a. Det vil sige, at både ufaglærte og faglærte påvirkes negativt i samme omfang af initialt at blive placeret i et udsat boligområde. Dog ses en svag tendens til, at flygtninge, der ankommer med mere end 10 års uddannelse, oplever en mere negativ effekt på erhvervsfrekvensen på kort sigt og mellemlangt sigt, mens effekten er mindre på langt sigt.

Figur H.4. Effekt af placering i udsat boligområde på erhvervsfrekvens efter uddannelsesniveau.

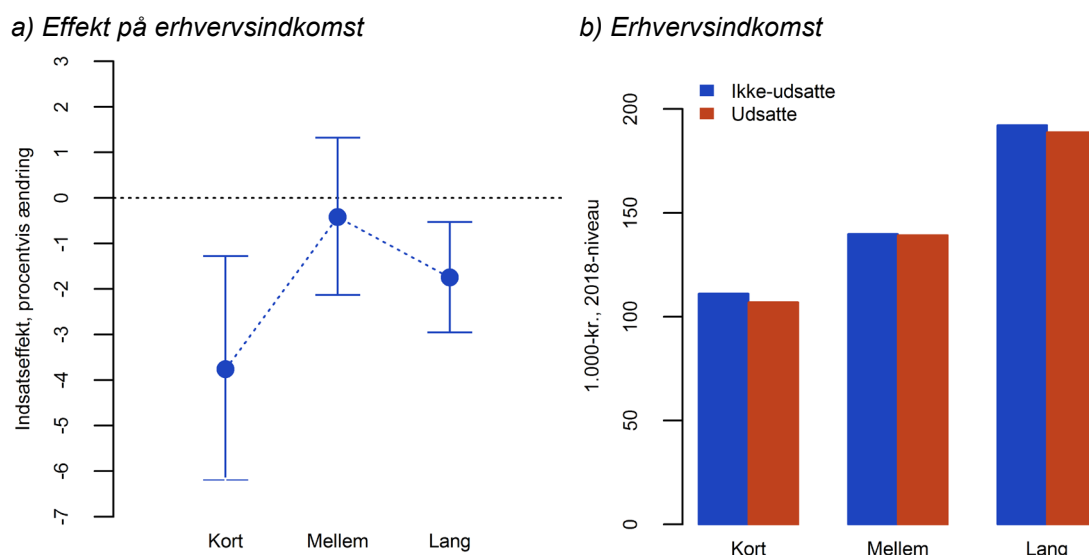
Anm.: Konfidensbånd er indtegnet omkring estimatet for indsatseseffekten. Se i øvrigt anm. til Figur H.1.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Effekt på erhvervsindkomst

Til at supplere analysen af effekten på erhvervsfrekvens fokuseres der i det følgende på effekten på erhvervsindkomst for dem, som er i beskæftigelse.

Det ses, at erhvervsindkomsten er stigende i opholdstid, jf. Figur H.5. På langt sigt ses en negativ forskel for flygtninge, der er placeret i et udsat boligområde, på 2,2 procentpoint, hvilket svarer til ca. 4.000 kr. om året i 2018-niveau.

Figur H.5. Effekt af placering i udsat boligområde på erhvervsindkomst.

Anm.: Konfidensbånd er indtegnet omkring estimatet for indsatseseffekten. Se i øvrigt anm. til Figur H.1.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

I. Anvendelse af propensity scores til fastlæggelse af gennemsnitlig indsats effekt for målgruppen

Typisk interesserer man sig for, hvorvidt resultatet fra en indsatsanalyse på analysegruppen S kan generaliseres til en større population afgrænset ved målgruppen P . Analysegruppen, der består af indsats- og kontrolgruppen, er typisk en delmængde af målgruppen, men dette behøver ikke at være tilfældet. Konkret ønskes det at kunne estimere den gennemsnitlige indsats effekt for ikke blot analysegruppe S , men for hele målgruppe P . Denne gennemsnitlige indsats effekt for analysegruppen er beskrevet i appendiks G, men kan alternativt skrives som

$$SATE = E[Y_{i,1} - Y_{i,0} | S_i = 1] = \frac{1}{N_S} \sum_{i=1}^{N_P} (Y_{i,1} - Y_{i,0} | S_i = 1)$$

hvor $S_i = 1$ angiver, at individ i tilhører analysegruppen. $SATE$ svarer således til ATE .

Ved generalisering ønskes det at fastlægge $PATE$, der på samme vis som ovenfor kan skrives som

$$PATE = E[Y_{i,1} - Y_{i,0} | P_i = 1] = \frac{1}{N_P} \sum_{i=1}^{N_P} (Y_{i,1} - Y_{i,0} | P_i = 1)$$

Hvis indsats effekten er konstant, vil $PATE = SATE$.

Forskelle mellem propensity scores er i afsnit 2.1 anvendt til vurdering af forskelligheden mellem indsats- og kontrolgruppen og til etablering af et balanceret datasæt til brug estimation af den gennemsnitlige indsats effekt for analysegruppen. Propensity scores kan imidlertid også finde anvendelse i forbindelse med estimation af $PATE$. Mens formålet med estimation af propensity scores i afsnit 2.1 var en opnåelse af sammenlignelighed i kovariatstrukturen mellem indsats- og kontrolgruppen, sigtes der her mod sammenlignelighed mellem analyse- og målgruppen.

Tilgangen, der følger Stuart et al. (2015) indebærer, at der for målgruppen indledningsvist estimeres en logit-model for deltagelse i analysegruppen betinget af en række kovariater. På baggrund af de estimerede sandsynligheder for det binære udfald, dvs. propensity scores p_i , beregnes efterfølgende individuelle vægte $w_i = 1/p_i$, der varierer på tværs af kovariaterne. Disse vægte benyttes efterfølgende til at vægte samtlige individer i analysegruppen, således at kovariatstrukturen blandt såvel kontrol- som indsatsgruppen afspejler strukturen i målgruppen. For en person med identiske karakteristika anvendes samme vægt i både indsats- og kontrolgruppen, hvorfor balanceringen er upåvirket. $PATE$ fås som den indsats effekt, der estimeres ved at gentage indsatsanalysen på den vægtede analysegruppe.

Det er en central forudsætning, at modellen, der beskriver deltagelse i analysegruppen omfatter de samme variable, som er benyttet til estimation af indsats effekten. Det antages dermed, at de pågældende variable er tilstrækkelige til at generalisere indsats effekten fra analyse- til målgruppen.

Vurderingen af i hvilken udstrækning indsats effekten kan generaliseres til en større målgruppe, er dermed alene knyttet til udvalgte målbare karakteristika for befolkningen og resultaterne skal derfor fortolkes med dette forbehold.

J. Undersøgelsens datagrundlag

Analysens datagrundlag er en række administrative registre, som indeholder en række oplysninger på befolkningen og boligerne i Danmark. Registerne er tilgængelige for DREAM under Danmarks Statistiks forskerserviceordning. Transport-, Bygnings- og Boligministeriet har leveret data til identifikation af boliger på ministeriets ghettoliste.

Udgangspunktet er befolkningsstatistikken, der indeholder den samlede danske befolkning i perioden 1986-2017. Til hver person er tilknyttet en bopæls- og opgangsadresse. Til hver opgangsadresse hører en unik geografisk placering.

Til dannelse af nabolag har Danmarks Statistik dannet en afstandsmatrice, hvori afstanden for hver opgangsadresse til de 2.500 nærmeste naboadresser i 2017 kan aflæses. Dette gør det muligt at danne nabolag indeholdende de 1.000 personer, som bor nærmest den betragtede bolig. Desuden kan socioøkonomiske karakteristika for personer i nabolaget udregnes.

Adresser kan over tid ændres, og nogle adresser udgår ved nedrivning af boliger. Dette medfører, at nogle adresser ikke findes i den af Danmarks Statistiks dannede nøgle. Dette problem løses ved en randomiseret placering af de manglende opgangsadresser på den rigtige vej ved boliger af samme type.

For at bestemme andelen af beboere i alderen 18-64 år, der er uden tilknytning til arbejdsmarkedet eller uddannelse, er der taget udgangspunkt i den registerbaserede arbejdsstyrkestatistik (RAS) og uddannelsesstatistikken (UDDA). Et individ er betragtet som at være uden tilknytning til arbejdsmarkedet, hvis personen hverken er beskæftiget eller ledig ifølge RAS. Bemærk, at det er individets primære arbejdsmarkedsstatus ultimo november, der tages udgangspunkt i. For primobefolkningen et givent år anvendes tilknytning til arbejdsmarkedet ultimo november det foregående år. For at bestemme om et individ er i gang med en uddannelse bruges UDDA-datasættet, der informerer om individet er i gang med en uddannelse pr. 1. oktober.

For i et givent år at bestemme andelen af beboere i alderen 30-59 år, som har grundskole som højst fuldførte uddannelse inkl. uoplyst uddannelse, bruges uddannelsesstatistikken (UDDA) fra oktober året før. Der ses bort fra informationer om indvandreres medbragte uddannelser, som enten er fremkommet ved imputation, eller ved at individet selv har besvaret spørgsmål om uddannelsesniveaut. Individet for hvem dette gælder, betragtes som havende en ukendt højst fuldført uddannelse.

Andelen af beboere over 15 år, der er dømt for overtrædelse af straffeloven, våbenloven eller loven om euforiserende stoffer opgøres ud fra kriminalstatistikken indeholdende afgørelser (KRAF). Det opgøres, om en person i primobefolkningen har modtaget dom i løbet af det foregående kalenderår.

Bruttoindkomsten for 15-64 årige eksklusiv uddannelsessøgende bestemmes ved datasættet for personindkomster (IND). Her benyttes indkomstbegrebet personindkomst i alt ekskl. beregnet lejeværdi af egen bolig og før fradrag af renteudgifter (PERINDKIALT_13). For en person i primobefolkningen opgøres indkomsten i det foregående kalenderår.

I rapportens kapitel 5 identificeres flygtninge ud fra indvandreres opholdsgrundlag. På individniveau findes data om opholdsgrundlag dog kun for personer indvandret fra og med 1997. For personer indvandret før 1997 vurderes en persons opholdsgrundlag ud fra historisk viden om indvandringsmønstre kombineret med viden om statsborgerskab ved indvandring og første indvandringsår, jf. Danmarks Statistik (2018).

DREAM

Amaliegade 44
DK-1256 København K

info@dreammodel.dk
www.dreammodel.dk

ISBN: 978-87-93360-20-4