

DREAM

# Danmarks fremtidige befolkning Befolkningsfremskrivning 2011

Marianne Frank Hansen & Peter Stephensen

## Indholdsfortegnelse

1	Indledning .....	6
1.1	Opbygningen af befolkningsmodellen .....	8
1.2	Vitale begivenheder .....	9
1.2.1	Dødsfald.....	9
1.2.2	Estimation af dødelighed.....	9
1.2.3	Fødsler .....	10
1.2.4	Estimation af fertilitet .....	10
1.2.5	Immigration .....	12
1.2.6	Emigration .....	13
1.2.7	Naturalisering .....	14
2	Befolkningsfremskrivningsmodellen .....	15
2.1	Befolkningen fordelt på alder, køn og oprindelse .....	15
2.2	Bestemmelse af risikogrupper og vitale begivenheder.....	18
2.2.1	Det generelle tilfælde.....	19
2.2.2	Nyfødte.....	21
2.2.3	Identisk dødelighed på tværs af oprindelse .....	24
2.2.4	Fødsler og fertilitet.....	25
2.2.5	Nyfødtes fordeling på køn og oprindelse .....	25
2.3	Eksogene parametre i fremskrivningsmodellen .....	27
3	Datagrundlaget for fremskrivningsmodellen.....	29
3.1	Dødelighedsdata .....	29
3.2	Rapportering af restlevetid .....	31
4	Dødelighed og forventet levetid .....	33
4.1	Lee-Carter metoden.....	34
4.2	Estimation af Lee-Carter modellen .....	36
4.3	Korrektioner af fremskrivningsmetoden.....	37
4.3.1	Bell's metode samt modifikation heraf.....	38
4.3.2	Udglætning af dødeligheder for forskellige aldre i fremskrivningen .	40
4.4	Valg af dataperiode.....	41
4.5	Estimationsresultater .....	42
4.6	Konstruktion af dødshyppigheder for de ældste ældre .....	45

4.7	Estimation og fremskrivning af dødsfald inden for de første 10 levedage	46
4.8	Resultater af fremskrivningen .....	47
4.9	Udviklingen i dødeligheden for de enkelte aldersgrupper.....	49
4.10	Definition af restlevetid .....	51
4.10.1	Sædvanlig fremgangsmåde for beregning af restlevetid i dødelighedstavler.....	51
4.10.2	Alternativ beregning af restlevetid .....	53
4.10.3	Intuitiv formulering af restlevetid.....	55
4.11	Restlevetidsudviklingen for mænd .....	57
4.12	Restlevetidsudviklingen for kvinder.....	59
4.13	Fremtidig udvikling i den aldersbetingede restlevetid.....	61
4.13.1	Fremtidig udvikling i middellevetiden .....	61
4.13.2	Fremtidig udvikling i restlevetiden for 60-årige.....	63
4.14	Sammenfatning af resultater om restlevetidsstigning.....	64
5	Fødsler og Fertilitet .....	65
5.1	Historisk udvikling i fødsler og fertilitet .....	65
5.2	Fremskrivning af aldersbetingede fertilitetskvotienter .....	68
5.3	Estimation af den aldersbetingede fertilitet for de store grupper .....	69
5.4	Bestemmelse af fertilitetskvotienter for de øvrige befolkningsgrupper	71
5.5	Konvergens af fertilitet på langt sigt.....	72
5.6	Fastlæggelse af det langsigtede fertilitetsniveau .....	74
5.7	Fremskrivningsresultater.....	75
5.8	Udviklingen i den aldersbetingede fertilitet.....	79
5.9	Nyfødtes fordeling på oprindelsesgrupper .....	80
6	Indvandring, udvandring og statsborgerskab .....	82
6.1	Lovgivning med betydning for indvandringen.....	82
6.2	Udviklingen i indvandringen 1992 - 2010.....	84
6.3	Fremskrivning af indvandringen .....	87
6.3.1	Førstegangsendvandring – indvandrere uden dansk statsborgerskab	87
6.3.2	Genindvandring.....	88
6.4	Udvandring .....	91

6.5	Statsborgerskabsskift.....	93
7	Resultater .....	95
7.1	Udviklingen i den samlede befolkning .....	95
7.2	Indvandring og udvandring .....	96
7.3	Befolkningsgrupperne.....	98
7.4	Aldersfordeling .....	100
8	Sammenligning med Befolkningsfremskrivning 2010 .....	106
8.1	Udviklingen i den samlede befolkning .....	106
8.2	Fødsler og fertilitet .....	108
8.3	Dødsfald og middellevetid.....	110
8.4	Vandringer.....	111
8.5	Udvikling i forsørger- og ældrekvote .....	113
9	Referencer .....	115

## Forord

4. maj 2011 udsendes den anden samordnede landsdækkende befolkningsfremskrivning mellem Danmarks Statistik (DST) og DREAM. Forskellen mellem sidste og dette års fremskrivning er primært forklaret af tilføjelsen af et nyt historisk år, idet opdateringen grundlæggende er baseret på samme metodiske tilgang som anvendt til 2010-fremskrivningen.

I nærværende publikation beskrives det datamæssige og modeltekniske grundlag for fremskrivningen samt de væsentligste resultater. Udgangspunktet for 2011-fremskrivningen er den registrerede befolkningsudvikling frem til primo 2011. Befolkningsudviklingen er – i overensstemmelse med tidligere år – fremskrevet indtil år 2100. Det skal understreges, at usikkerheden på befolkningsfremskrivningen er hastigt voksende i fremskrivningens længde. Idet der således er knyttet en betragtelig usikkerhed til især de langsigtede skøn, skal resultatet først og fremmest fortolkes som en illustration af befolkningsudviklingen og dens sammensætning for det anvendte sæt antagelser.

Fortolkninger og synspunkter udtrykt i nærværende publikation er DREAMs egne og ikke nødvendigvis sammenfaldende med Danmarks Statistiks.

# 1 Indledning

Befolkningsfremskrivninger har fået en stadigt mere fremtrædende placering ved fastlæggelse af den økonomiske politik og ved centrale beslutninger i f.eks. private pensionsselskaber. Derfor er der også i de seneste år kommet mere fokus på de anvendte metoder og modeller.

DREAM har udarbejdet befolkningsfremskrivninger siden 1999, og disse er løbende blevet dokumenteret i forbindelse med beskrivelse af resultatet af den aktuelle fremskrivning. Gradvist er der sket en udbygning af de anvendte metoder. I 2010 udsendte DREAM og Danmarks Statistik (DST) den første samordnede landsdækkende befolkningsfremskrivning. Formålet med samordningen var dels at eliminere den forvirring og manglende konsistens som eksistensen af to forskellige fremskrivninger hidtil har givet anledning til og dels på sigt at skabe et gunstigt miljø for videreudvikling af metoder til fastlæggelse af demografiske begivenheder og det bagvedliggende datagrundlag. Udfærdigelsen af den første fælles fremskrivning var primært funderet i de metoder, der hidtil har været anvendt af DREAM, mens datagrundlaget optimeredes gennem udtræk fra PersonStatistikDatabasen (PSD) leveret fra DST.

Relativt til DREAMs tidligere fremskrivninger var den mest centrale ændring i forbindelse med samordningen været et skift i valg af oprindelsesfordeling. Dette skift var nødvendigt for tilpasning til DSTs regionale fremskrivning og øvrig statistik, der er funderet i en opdeling af befolkningen i fem grupper bestående af personer af dansk oprindelse, indvandrere fra vestlige/ikke-vestlige lande samt efterkommere fra vestlige/ikke-vestlige lande. Tidligere har DREAM anvendt en opdeling på mere-/ og mindre udviklede lande følgende FNs definitioner<sup>1</sup>.

Formålet med denne publikation er at dokumentere de anvendte metoder i DREAMs befolkningsfremskrivningsmodel og det datagrundlag, der danner udgangspunkt for modellen. Modellens grundlæggende struktur, der præsenteres i kapitel 2, er med undtagelse af en mindre korrektion uændret relativt til sidste år. Datagrundlaget stammer som sidste år fra et udtræk fra, PersonStatistikDatabasen, hvilket indebærer, at historisk data er revideret i det omfang, der er indløbet nye oplysninger til databasen siden sidste år. I publikationen præsenteres endvidere resultaterne af årets fremskrivning og endelig sammenlignes dette års fremskrivning med fremskrivningen fra 2010 og forskellene dekomponeres i effekter hidrørende de respektive demografiske bevægelser.

Dokumentationen er organiseret på følgende måde: I dette kapitel gives det første overblik over modellen. I kapitel 2 opstilles den anvendte

---

<sup>1</sup> <http://esa.un.org/unpp/definition.html> "Vestlige lande" omfatter alle 27 EU-lande samt Andorra, Island, Liechtenstein, Monaco, Norge, San Marino, Schweiz, Vatikanstaten, Canada, USA, Australien og New Zealand. Alle øvrige lande udgør gruppen "ikke-vestlige lande".

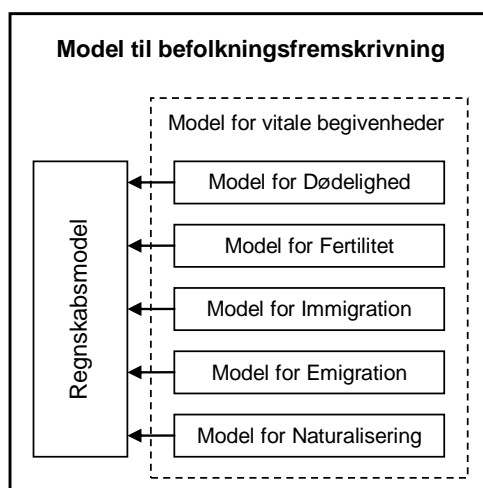
befolkningsfremskrivningsmodel og der redegøres for dimensionaliteten samt de eksogene variable, der fastlægges ved estimation baseret på den historiske udvikling. I kapitel 3 præsenteres de anvendte datakilder og centrale demografiske begreber defineres. Afsnit 4 omfatter en beskrivelse af udviklingen af dødeligheden i Danmark gennem det seneste århundrede og af Lee-Carter metoden, der i modificeret form anvendes til estimation af den fremtidige udvikling i dødeligheder og dermed den aldersbetingede restlevetid. Endvidere rummer kapitlet en beskrivelse af forskellige indbyrdes konsistente tilgange til beregning af forventet levetid. Tilsvarende indeholder kapitel 5 en beskrivelse af udviklingen i antallet af fødsler og i fertiliteten gennem det seneste århundrede og på denne baggrund præsenteres fremskrivningen af den samlede fertilitet. Fremskrivningen er baseret på en anvendelse af Cubic Spline Smoothing teknikken, som ligeledes beskrives i kapitlet. Kapitel 6 omfatter en gennemgang af den historiske udvikling i indvandringen og de vigtigste hertil knyttede politiktiltag. Yderligere beskrives den anvendte tilgang til fremskrivning af henholdsvis indvandring, udvandring og statsborgerskabsskifte. Kapitel 7 præsenterer resultaterne af den samlede befolkningsfremskrivning, mens fremskrivningen endeligt sammenlignes med den forrige i kapitel 8.

## 1.1 Opbygningen af befolkningsmodellen

Overordnet består en befolkningsmodel af et sæt ligninger, der holder regnskab med udviklingen i befolkningen fra år til år. Demografiske regnskabsmodeller blev første gang benyttet af Edwin Cannan (1861-1935) til at lave befolkningsfremskrivninger for England og Wales, mens amerikaneren Pascal K. Whelpton i et antal artikler fra omkring 2. verdenskrig opstillede en samlet metode til befolkningsfremskrivning, jf. Alho og Spencer (2005). Siden 1963 har Danmarks Statistik produceret befolkningsfremskrivninger, der ligeledes er baseret på en regnskabsmodel for befolkningen, jf. Danmarks Statistik (1978). Således har samordningen af fremskrivningerne mellem DREAM og DST altså heller ikke medført væsentlige ændringer af den anvendte grundmodel set fra DSTs synspunkt.

Den grundlæggende metode bag demografiske regnskabsmodeller består i, at en given udgangsbefolkning opdeles på et antal undergrupper (f.eks. køn, alder og oprindelse), der fremskrives ét år ad gangen på baggrund af vitale demografiske begivenheder: fødsler, dødsfald, indvandring, udvandring og naturalisering/statsborgerskabskifte. Figur 1 viser opbygningen af DREAMs model til befolkningsfremskrivning.

**Figur 1. En model til befolkningsfremskrivning er sammensat af en række delmodeller for de vitale begivenheder, og regnskabsmodellen holder regnskab med udviklingen i befolkningen på baggrund af udviklingen i de enkelte vitale begivenheder.**



Det helt centrale i enhver befolkningsfremskrivning er den fremtidige udvikling i fertilitet, dødelighed, indvandring, udvandring og naturalisering. I det følgende gennemgås de antagelser og metoder, der benyttes i befolkningsfremskrivningen.

## 1.2 Vitale begivenheder

Regnskabsmodellen for befolkningen holder styr på ændringer i primobefolkningen mellem to år. Mekanismerne, der angiver udviklingen af Danmarks befolkning er relativt få. Befolkningen vokser på grund af fødsler og indvandring og formindskes på grund af dødsfald og udvandring, mens sammensætningen af befolkningen på oprindelsesgrupper ændres gennem naturalisering, dvs. opnåelse af dansk statsborgerskab.

### 1.2.1 Dødsfald

Antallet af døde et givet år i fremskrivningen opnås ved at multiplicere fremskrevne køns- og aldersfordelte dødshyppigheder med antallet af personer, der er eksponeret for dødsrisiko. Der antages således identiske dødshyppigheder på tværs af oprindelse. Antagelsen skal tilskrives, at der blandt efterkommere på nuværende tidspunkt er relativt få personer i de ældre aldersgrupper, hvor dødeligheden er højst. Estimationsgrundlaget er således for tyndt til udarbejdelse af en fremskrivning af dødelighed for hver enkelt oprindelsesgruppe. For indvandrere er datamaterialet dog ved at opnå en tæthed, der retfærdiggør rimeligheden i inden for nærmeste fremtid at undersøge muligheden for særskilt fremskrivning for disse befolkningsgrupper. Således viser en deskriptiv analyse sammenlignende dødeligheden for personer med dansk oprindelse og indvandrere for perioden 2005-2009 viser, at indvandrere fra både vestlige og ikke-vestlige lande i stort set alle aldersgrupper har en lavere dødelighed end personer af dansk oprindelse<sup>2</sup>.

Som indført i 2010-fremskrivningen skelnes desuden mellem spædbørnsdødelighed indtrædende hhv. de første 10 dage efter fødslen og den resterende del af første leveår. Idet dødssandsynligheden blandt 0-årige er specielt høj kort tid efter fødslen og 0-årige indvandrere primært kommer til landet efter denne periode, opnås ved denne differentiering en mere realistisk vurdering af dødsfald blandt indvandrede spædbørn. Ved kun at betragte en enkeltstående dødelighed blandt 0-årige, vil der grundet det høje bidrag fra helt nyfødte alt andet være tendens til at overvurdere dødsfald blandt indvandrede spædbørn.

### 1.2.2 Estimation af dødelighed

For given alder og køn er dødeligheden i en given periode defineret som antallet af dødsfald divideret med antallet af personer, der er udsat for dødsrisiko. Til estimation af fremtidige dødeligheder benyttes en modificeret version af Lee-Carter metoden, der er en ekstrapolativ baseret metode, jf. Lee & Carter (1992a, 1992b). Den grundlæggende antagelse i Lee-Carter metoden er, at den fremtidige udvikling i en given aldersbetinget dødelighed kan beskrives ud fra den historiske udvikling i de aldersbetingede dødeligheder. For at reducere

---

<sup>2</sup> Danmarks Statistik (2010a): Tabel 1.21.

dimensionen af problemet antages det, at der er en betydelig regularitet i udviklingen af de aldersspecifikke dødeligheder, således at de kan beskrives af tre elementer: En aldersspecifik del, der er uafhængig af tidspunktet, et mortalitetsindeks, der er et tidsafhængigt mål for den samlede dødelighedsudvikling og en tidsafhængig aldersprofil, der vægter gennemslaget af udviklingen i mortalitetsindekset på dødeligheden i de enkelte aldersklasser. Metoden indebærer, at forholdet mellem vækstraterne i dødeligheden for to forskellige aldersgrupper er konstant. Den fremtidige udvikling i dødelighederne bestemmes alene af udviklingen i mortalitetsindekset, der estimeres ved standardmetoder fra tidsrækkeanalyse, jf. f.eks. Chatfield (2004). Til estimation af dødeligheder anvendes til 2011-fremskrivningen data fra perioden 1990-2010. I kapitel 4 vil den anvendte Lee-Carter metode samt metoden til ekstrapolation af dødeligheder for nyfødte blive gennemgået.

### 1.2.3 Fødsler

Det samlede antal fødsler findes som summen af børn født af mødre i de enkelte befolkningsgrupper. Nyfødtes oprindelsesgruppe afhænger, som følge af definitionerne af oprindelsesgrupperne, både af faderens og moderens oprindelse. Der konstrueres på basis af de historiske erfaringer gennem de sidste 3 år (2008-2010 i 2011-fremskrivningen) en sandsynlighedsfordeling for barnets oprindelse givet moderens. Denne fordeling anvendes i fremskrivningen til at fordele børnene på oprindelsesgrupper. I afsnit 2.2.5 beskrives metoden til at tildele børn oprindelse på baggrund af moderens oprindelse.

Fødslerne i en given befolkningsgruppe kan beregnes som den aldersbetingede fertilitet multipliceret med kvinder i den givne befolkningsgruppe og med den givne alder. Fødsler opdeles på køn ved at antage, at en konstant andel af alle nyfødte er drenge. Andelen af drengebørn beregnes som gennemsnittet over hele dataperioden.

### 1.2.4 Estimation af fertilitet

Den aldersfordelte fertilitet er defineret som antallet af fødsler blandt mødre med en given alder divideret med antallet af potentielle mødre, dvs. kvinder i den pågældende aldersklasse. Fremskrivningen af fertilitet foretages ved at fremskrive den aldersbetingede fertilitet for hver oprindelsesgruppe. I modsætning til estimationen af dødeligheder estimeres den enkelte aldersgruppes fertilitet uafhængigt af den samlede fertilitet og vil derfor også være uafhængig af de øvrige aldersgruppes fertilitet. Udviklingen i den samlede fertilitet i fremskrivningen fremkommer således som en sum af udviklingen i de enkelte aldersgruppes fertilitet. Der er historisk set større udsving i fertilitet end i dødelighed, hvorfor udviklingen på kort sigt bestemmes på baggrund af den nuværende trend i fertiliteten, mens fertiliteten på langt sigt konvergerer mod et

eksogent ligevægtsniveau. På kort sigt kan der således være stor forskel i udviklingen i fertilitet for befolkningsfremskrivninger med forskelligt basisår.

Det er udelukkende for de tre største befolkningsgrupper, at der er et tilstrækkeligt antal observationer til at kunne estimere den aldersbetingede fertilitet med et-års alderstrin. De tre grupper er: Personer med dansk oprindelse, indvandrere fra ikke-vestlige lande uden dansk statsborgerskab og indvandrere fra vestlige lande uden dansk statsborgerskab. For disse grupper uddrages for hver alder mellem 15 og 49 år trenden af den historiske udvikling i fertiliteten og den aldersbetingede fertilitet fremskrives derefter ved hjælp af disse trends. Trenden i den aldersbetingede fertilitet udledes ved anvendelse af en metode benævnt *Cubic Spline Smoothing* (CSS), se Hyndman, King og Billah (2002). Der estimeres til 2011-fremskrivningen på baggrund af data fra perioden 1992-2010. Metoden medfører en udglattet konvergens mod et ligevægtsniveau defineret på baggrund af den historiske udvikling. I afsnit 5.3 gennemgås Cubic Spline Smoothing metoden til estimation af fertilitet.

De resterende befolkningsgrupper er små. Det er problematisk at fremskrive fertiliteten for disse befolkningsgrupper af to grunde. For det første medfører gruppernes størrelse en betydelig statistisk usikkerhed. For det andet er de kvindelige medlemmer af efterkommergrupperne hovedsageligt under 25 år, hvilket betyder, at det er særdeles vanskeligt at skønne over den aldersbetingede fertilitet for den ældste halvdel af den fødedygtige alder på grund af manglende data. Fremskrivningen konstrueres derfor ved at sammenligne den aldersbetingede fertilitetsprofil for hver enkelt af de små befolkningsgrupper med den tilsvarende aldersbetingede fertilitetsprofil for de tre store grupper. Ved hjælp af lineær regressionsanalyse på observationer for de seneste 5 år estimeres de små befolkningsgruppers aldersbetingede fertilitet som et vejet gennemsnit af de tre store gruppers aldersbetingede fertilitet. Herefter fremskrives den aldersbetingede fertilitet for de små befolkningsgrupper på grundlag af fremskrivningerne af de tre store gruppers aldersbetingede fertilitet. I afsnit 5.4 beskrives metoden til estimation af fertilitet for de små befolkningsgrupper på baggrund af de tre store befolkningsgrupper.

Når den kortsigtede udvikling i fertiliteten er bestemt for samtlige befolkningsgrupper, bestemmes konvergensens mod langsigtudviklingen. Til dette benyttes en *Richardskurve*, der i løbet af en given tidsperiode vægter mellem den kortsigtede og den langsigtede udvikling i fertiliteten, hvor sidstnævnte er eksogent bestemt. I afsnit 5.5 gennemgås denne metode, der sikrer, at langsigtudviklingen ikke ændres voldsomt ved indførelsen af et nyt dataår. Dette ville være u hensigtsmæssigt, idet tilføjelsen af et enkelt dataår ikke indeholder information, der gør os mere vidende om det mellemlange og lange sigt. Såfremt der over en længere årrække observeres tendenser, der trækker i en entydig retning kan man tage langsigtetsniveauet op til overvejelse.

### 1.2.5 Immigration

Personer, der indvandrer til en given befolkningsgruppe, er opdelt i to forskellige typer indvandring. For indvandrere uden dansk statsborgerskab er der en eksogen tilstrømning af indvandrere, mens der for alle andre befolkningsgrupper er genindvandring, der beregnes ud fra indvandringsfrekvensen.

For indvandrere uden dansk statsborgerskab er den fremtidige indvandring antaget eksogent givet. Som udgangspunkt antages således, at indvandringen fra hhv. vestlige og ikke-vestlige lande fordelt på køn og alder i fremtiden bliver på et givet niveau. Udviklingen i de senere års indvandring må formodes i stor udstrækning at afspejle de konjunkturmæssige forhold, men vil også være præget af den gradvise lempelse af kravene i Østaftalen. For første gang siden 2003 kunne for begge typer af indvandrere uden dansk statsborgerskab noteres et fald i indvandringen fra 2008 til 2009. Fra 2009 til 2010 opfølges dette dog af en mindre stigning. Kombinationen af lempelser i Østaftalen og de seneste års højkonjunktur gør det vanskeligt at vurdere hvor stor en del af de seneste års udvikling, der måtte indikere en permanent ændring i indvandringstilbøjeligheden. Således er det valgt at skønne relativt forsigtigt over det fremtidige indvandringsniveau og dermed ikke lade de seneste års tendens have for stor indflydelse.

Den samlede indvandring til grupperne bestående af indvandrere uden dansk statsborgerskab fra vestlige og ikke-vestlige lande er i 2010-fremskrivningen fastlagt til hhv. 20.000 og 11.000 personer. For indvandrere fra vestlige lande er der relativt til 2010-fremskrivningen tale om en opjustering af det årlige eksogene vandringsomfang med 3.000 personer. Beslutningen er funderet i, at denne justering sammen med antagelserne anvendt om udvandringstilbøjeligheden vil give anledning til en nettoindvandring, der er mere konsistent med den historisk observerede udvikling. Indvandringen fordeles på køn og alder i henhold til, hvad der gennemsnitligt har været gældende for indvandringen til de respektive grupper i perioden 2008-2010.

Indvandring til øvrige befolkningsgrupper med dansk statsborgerskab, beregnes ud fra indvandringsfrekvenser fordelt på alder, køn og oprindelse. Frekvenserne beregnes som et gennemsnit over de sidste tre dataår, senest 2008-2010. En mekanisk fremgangsmåde som denne indebærer en betydelig risiko for, at konjunktur- og midlertidige politikeffekter videreføres gennem fremskrivningen. Selv om genindvandring må betegnes som mindre konjunkturfølsomt end indvandring af indvandrere uden dansk statsborgerskab, bør beregningen suppleres med en vurdering af realismen af de langsigtede effekter. Fordelen ved at anvende et gennemsnit, der gradvist ændres med hver fremskrivning, er på den anden side, at man ved permanent ændret adfærd opnår en gradvis tilpasning til den nye tendens.

Med udgangspunkt i data fra PSD er det blevet klarlagt, at indvandringen opgjort efter registrerede hændelser er en anelse undervurderet relativt til den faktiske indvandring. Dette skyldes forsinkede indberetninger. Indvandring hørende til år  $t$  opgøres sædvanligvis 1. februar år  $t+1$ . Revideres opgørelsen for år  $t$  i år  $t+2$  ses en stigning i niveauet relativt til den første opgørelse, idet yderligere indberetninger er indløbet. Yderligere revideringer i de efterfølgende år, dvs.  $t+3$ ,  $t+4$  osv. indikerer, at der ikke indløber ekstra indberetninger af betydning efter år  $t+2$ . Ud fra historisk data for indvandring opgjort i hhv. år  $t+1$  og  $t+2$  bestemmes køns- og aldersgruppeafhængige skaleringsfaktorer, der efterfølgende pålægges de fundne indvandringsfrekvenser for at korrigere for forsinkelser. De eksogene indvandringsniveauer ikke underlagt korrektion for forsinkelser, da disse skøn i forvejen må betegnes som relativt usikre.

I kapitel 6 er uddybet de overvejelser, der ligger til grund for bestemmelse af den fremtidige indvandring.

### 1.2.6 Emigration

Antallet af personer, der i et givet år udvandrer fra en given befolkningsgruppe, beregnes ud fra udvandringsfrekvensen fordelt på alder, køn og oprindelse. Frekvenserne beregnes som et gennemsnit over de sidste tre dataår, senest 2008-2010. Sammenholdes denne antagelse med antagelsen om genindvandring fås, at netto-indvandringsfrekvensen til befolkningsgrupper med dansk statsborgerskab er konstant, hvilket dog ikke gælder for netto-indvandringen til befolkningsgrupper uden dansk statsborgerskab.

Specielt for udvandring kan det konkluderes, at forsinkede indberetninger relativt til de registrerede hændelser spiller en central rolle. Opgørelsen af udvandring for år  $t$ , der som udgangspunkt er opgjort i år  $t+1$ , vil først udvise tendens til at stabilisere sig fra og med år  $t+4$ . Udvandring hørende til året 2005 opgøres således 1. februar 2009, 2006 opgøres i 2010 og 2007 opgøres i 2011 og der beregnes igen en gennemsnitlig køns- og aldersgruppeafhængig skaleringsfaktor, som de fundne udvandringshyppigheder justeres med.

Til sine tidligere befolkningsfremskrivninger, dvs. til og med år 2009, har DREAM selv udtrukket relevant data for vandring. En person blev ved denne lejlighed betegnet som værende udvandret, såfremt vedkommende var til stede primo året, ikke død i løbet af året og fraværende ultimo året. Denne opgørelsesmetode implicerer, at en vandring kan tilknyttes et år, hvor den faktiske hændelse ikke har fundet sted, men blot er registreret. Idet metoden dog er anvendt konsekvent for alle dataår, får man dog alligevel i et vist omfang inddraget de forsinkede indberetninger, men altså ikke i de korrekte år. Muligheden for at skelne mellem hændelser og registreringer er først opstået med anvendelsen af PSD, men grundet den tidligere anvendte metode til konstruktion af data, vurderes emnet altså ikke at have haft nogen signifikant betydning for DREAMs tidligere fremskrivninger.

For såvel ind- som udvandring gælder, at aldersprofilen for hyppighederne efter behov tættes ved skalering af den tilsvarende profil for personer af dansk

oprindelse. Dette vil afstedkomme en mere realistisk profil for aldersgrupper, hvor data er tyndt og der tidligere blot har været antaget konstant ind- og udvandringstilbøjelighed for personer ældre end en given alder.

### **1.2.7 Naturalisering**

Antallet af personer fra befolkningsgrupper uden dansk statsborgerskab, der i et givet år skifter statsborgerskab, beregnes ud fra frekvensen for statsborgerskabsskift fordelt på alder, køn og oprindelse. Frekvensen beregnes som et gennemsnit over de sidste tre dataår, senest 2008-2010.

Der er kun set på skift fra udenlandsk statsborgerskab til dansk, da den modsatte bevægelse er så lille, at man med rimelighed kan se bort fra den.

I afsnit 2.3 samt kapitel 6 beskrives estimation af den eksogene indvandring samt indvandrings-, udvandrings- og naturaliseringsfrekvenserne.

## 2 Befolkningsfremskrivningsmodellen

I dette afsnit vil der blive redegjort for de regnskabsmæssige sammenhænge i modellen til fremskrivning af befolkningen. Som beskrevet i forrige afsnit er de regnskabsmæssige sammenhænge mellem vitale begivenheder universelle for befolkningsmodeller, og de store forskelle mellem befolkningsmodeller ligger i modellering af vitale begivenheder, som vil blive behandlet i de kommende afsnit.

Udviklingen i den samlede befolknings størrelse fra et år til det næste afhænger af antal fødte, antal døde og antallet af ind- og udvandrede.

$$N_{t+1} = N_t + F_t - D_t + I_t - U_t \quad (1)$$

hvor  $N_t$  er befolkningen primo år  $t$ ,  $F_t$  er antal fødte i år  $t$ ,  $D_t$  er antal døde i år  $t$ ,  $I_t$  er antal indvandrede i år  $t$  og  $U_t$  er antal udvandrede i år  $t$ .

Der eksisterer ikke andre faktorer, der ændrer den samlede befolkning og derfor består en befolkningsfremskrivning generelt af ændringer i disse fire typer vitale begivenheder i hvert af de fremtidige år. For at kunne fastlægge disse størrelser opdeles befolkningen på et givet tidspunkt efter dimensionerne alder, køn og oprindelse, hvor oprindelse i nærværende sammenhæng anvendes som fælles betegnelse for kombinationen af oprindelsestype, oprindelsesland og statsborgerskabsstatus<sup>3</sup>.

### 2.1 Befolkningen fordelt på alder, køn og oprindelse

Kombinationen af oprindelsestype, oprindelsesland og statsborgerskabsstatus giver anledning til, at befolkningen opdeles i ti oprindelsesgrupper:

- Resterende befolkning (personer af dansk oprindelse) uden dansk statsborgerskab, *dan*.<sup>4</sup>
- Resterende befolkning (personer af dansk oprindelse) med dansk statsborgerskab, *dad*.

<sup>3</sup> *Oprindelsestype* omfatter indvandrere, efterkommere og den resterende befolkning, jf. afsnit 2.2. *Oprindelsesland* angiver for indvandrere hvorvidt vedkommende stammer fra et vestligt eller et ikke-vestligt land, mens der for efterkommere henvises til arten af moderens oprindelsesland. Mht. til statsborgerskabsstatus skelnes mellem hvorvidt man har dansk eller udenlandsk statsborgerskab.

<sup>4</sup> Gruppen af personer af dansk oprindelse indeholder et antal personer uden dansk statsborgerskab. Børn født i Danmark af disse personer kan derfor principielt bliver karakteriseret som efterkommere, jf. definitionen af oprindelsestype. Da antallet er meget begrænset er det valgt at se bort fra denne mulighed. Generelt behandles befolkningsgruppen *dan* adfærdsmæssigt som den store gruppe *dad*. I sådanne tilfælde anvendes den fælles oprindelsesbetegnelse *da*.

- Indvandrere fra ikke-vestlige lande uden dansk statsborgerskab, *ixn*
- Indvandrere fra ikke-vestlige lande med dansk statsborgerskab, *ixd*
- Indvandrere fra vestlige lande uden dansk statsborgerskab, *iwn*
- Indvandrere fra vestlige lande med dansk statsborgerskab, *iwd*
- Efterkommere uden dansk statsborgerskab hvis mor kommer fra et ikke-vestligt land, *dxn*
- Efterkommere med dansk statsborgerskab hvis mor kommer fra et ikke-vestligt land, *dxd*
- Efterkommere uden dansk statsborgerskab hvis mor kommer fra et vestligt land, *dwn*
- Efterkommere med dansk statsborgerskab hvis mor kommer fra et vestligt land, *dwd*

Fastlæggelse af oprindelsestypestatus, dvs. om et individ er indvandrer, efterkommer eller af dansk oprindelse følger Danmarks Statistiks definition og er givet ved:

**Indvandrere:** Personer, der er født i udlandet af forældre, der begge er udenlandske statsborgere eller er født i udlandet. Hvis der kun foreligger oplysninger om den ene forælder, defineres personen som indvandrer, hvis vedkommende er født i udlandet, og forælderen er udenlandsk statsborger eller født i udlandet. Hvis der ikke findes oplysninger om nogen af forældrene, og personen er født i udlandet, defineres personen ligeledes som indvandrer.

**Efterkommere:** Personer, der er født i Danmark af forældre, hvoraf ingen er både dansk statsborger og født i Danmark. Hvis der ikke findes oplysninger om nogen af forældrene, og personen er udenlandsk statsborger født i Danmark, betragtes vedkommende også som efterkommer.

**Dansk oprindelse:** Personer, hvoraf mindst en af forældrene er dansk statsborger og født i Danmark, uanset personens eget fødeland og statsborgerskab. Hvis der ikke findes oplysninger om nogen af forældrene, betragtes vedkommende som værende af dansk oprindelse, hvis vedkommende er dansk statsborger født i Danmark.

Disse gruppedefinitioner er udtømmende, dvs. alle personer i befolkningen tilhører en af de tre grupper. Definitionerne er endvidere entydige, dvs. en given person placeres i én og kun én af de tre grupper.

Underopdelingen af indvandrere og efterkommere efter oprindelsesland følger som nævnt tidligere ikke længere FNs fordeling på mere og mindre udviklede lande, men er erstattet med opdelingen på henholdsvis vestlige og ikke-vestlige lande, jf. fodnote 1. Den primære forskel mellem oprindelseslandsdefinitionerne er, at Balkanstaterne, Rusland og Japan regnes som mere udviklede lande, men ikke har status som vestlige lande<sup>5</sup>.

Med opdelingen af befolkningen på alder,  $x$ , køn,  $g$  og i de ti oprindelsesgrupper,  $o$ , er det differensligningssystem, der danner udgangspunkt for befolkningsfremskrivningen givet ved:

$$\begin{aligned} N_{t+1}^{x+1,g,o} &= N_t^{x,g,o} + \bar{I}_t^{x+1,g,o} + I_t^{x+1,g,o} - U_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} - S_t^{x+1,g,o} - D_t^{x+1,g,o} \\ N_{t+1}^{0,g,o} &= F_t^{g,o} + \bar{I}_t^{0,g,o} + I_t^{0,g,o} - U_t^{0,g,o} + \bar{S}_t^{0,g,o} - S_t^{0,g,o} - D_t^{0,g,o} \end{aligned} \quad (2)$$

hvor

- $N_t^{x,g,o}$  er befolkningen med alder  $x$ , køn  $g$  og oprindelse  $o$  primo år  $t$ ,
- $\bar{I}_t^{x,g,o}$  er det eksogent givne antal indvandrede med alder  $x$ , køn  $g$  og oprindelse  $o$  i år  $t$
- $I_t^{x,g,o}$  er antal indvandrede med alder  $x$ , køn  $g$  og oprindelse  $o$  i år  $t$
- $U_t^{x,g,o}$  er antal udvandrede med alder  $x$ , køn  $g$  og oprindelse  $o$  i år  $t$
- $\bar{S}_t^{x,g,o}$  er det eksogene antal med alder  $x$ , køn  $g$  og oprindelse  $o$  i år  $t$ , der skifter til dansk statsborgerskab (og dermed oprindelsesgruppe)
- $S_t^{x,g,o}$  er det endogent bestemte antal med alder  $x$ , køn  $g$  og oprindelse  $o$  i år  $t$ , der skifter til dansk statsborgerskab (og dermed oprindelsesgruppe)
- $D_t^{x,g,o}$  er antal døde med alder  $x$ , køn  $g$  og oprindelse  $o$  i år  $t$
- $F_t^{g,o}$  er antal fødte af køn  $g$  og oprindelse  $o$  i år  $t$ . Bemærk, at  $F_t^{g,o} = 0$  for  $o \in \{ixn, ixd, iwn, iwd\}$ , da indvandrere pr. definition ikke kan være født i Danmark.

Det er centralt at bemærke, at aldersdateringen af beholdninger og strømme er forskellig. Befolkningen  $N_t^{x,g,o}$  er en beholdning og angiver antallet af personer i

---

<sup>5</sup> En detaljeret oversigt over forskellen mellem oprindelsesfordelingerne mere/mindre udviklede lande og vestlige/ikke-vestlige lande kan fås ved henvendelse til DREAM.

oprindelsesgruppe  $o$ , der primo året, dvs. pr. 1. januar, har alderen  $x$ . De resterende variable er strømme og er ultimo daterede.  $I_t^{x,g,o}$  angiver f.eks. antallet af personer, der er indvandret til oprindelsesgruppe  $o$  i løbet af år  $t$ , og som ved dette års afslutning har alderen  $x$ . Denne lidt specielle dateringsteknik bruges ofte i demografiske sammenhænge, og har til formål at skaffe plads til ligningen, der beskriver udviklingen i antallet af 0-årige.

Hvor intet andet er anført gælder, at variablene i modellen er endogene. For befolkningsgrupper med dansk statsborgerskab, dvs.  $dad$ ,  $ixd$ ,  $iwd$ ,  $dxd$  og  $dwd$ , er den eksogene variabel  $\bar{S}_t^{x,g,o}$  defineret, idet den angiver antallet af personer, der tilgår grupperne pga. skift til dansk statsborgerskab. Omfanget af skiftet fastlægges af adfærden i de befolkningsgrupper personerne afgår fra og antages derfor for værende eksogent fra modtagergruppens synspunkt. Således eksisterer den endogene variabel  $S_t^{x,g,o}$  ikke for grupperne med dansk statsborgerskab, men er defineret for personer uden dansk statsborgerskab, dvs. for  $dan$ ,  $ixn$ ,  $iwn$ ,  $dxn$  og  $dwn$ . Der gælder, at

$$\begin{aligned}\bar{S}_t^{x,g,ixd} &= S_t^{x,g,ixn} \\ \bar{S}_t^{x,g,iwd} &= S_t^{x,g,iwn} \\ \bar{S}_t^{x,g,dxd} &= S_t^{x,g,dxn} \\ \bar{S}_t^{x,g,dwd} &= S_t^{x,g,dwn} \\ \bar{S}_t^{x,g,dad} &= S_t^{x,g,dan}\end{aligned}$$

så

$$\begin{aligned}\sum_o N_{t+1}^{x+1,g,o} &= \sum_o N_t^{x,g,o} + \bar{I}_t^{x+1,g,o} + I_t^{x+1,g,o} - U_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} - S_t^{x+1,g,o} - D_t^{x+1,g,o} \\ &= N_t^{x,g} + \bar{I}_t^{x+1,g} + I_t^{x+1,g} - U_t^{x+1,g} - D_t^{x+1,g}\end{aligned}\quad (3)$$

Den eksogene indvandring  $\bar{I}_t^{x,g,o}$  er relevant for indvandrere uden dansk statsborgerskab, dvs. for de to oprindelsesgrupper  $ixn$  og  $iwn$ , mens indvandring til de øvrige befolkningsgrupper har karakter af genindvandring og bestemmes ud fra historiske vandringstilbøjeligheder.

## 2.2 Bestemmelse af risikogrupper og vitale begivenheder

Generelt kan vitale begivenheder modelleres enten ved et antal personer, der oplever en vital begivenhed, eller ved en frekvens, der med udgangspunkt i en risikogruppe, beskriver andelen, der oplever den pågældende begivenhed. Fastlæggelsen af det såkaldte risikogruppebegreb er dermed centralt i

modelleringen af demografiske begivenheder. Nedenfor behandles 0-årige og den resterende befolkning særskilt, jf. de separate udtryk i (2).

### 2.2.1 Det generelle tilfælde

De endogene variable i ligningssystemet (2) fastlægges som

$$\begin{aligned} I_t^{x+1,g,o} &= i_t^{x+1,g,o} E_t^{x,g,o} & (4) \\ U_t^{x+1,g,o} &= u_t^{x+1,g,o} E_t^{x,g,o} \\ D_t^{x+1,g,o} &= d_t^{x+1,g,o} E_t^{x,g,o} \\ S_t^{x+1,g,o} &= s_t^{x+1,g,o} E_t^{x,g,o} \end{aligned}$$

hvor  $E_t^{x,g,o}$  for givet køn, alder og oprindelse, angiver det antal personer, der er eksponeret for risiko for hhv. at indvandre, udvandre, dø eller skifte statsborgerskab i løbet af år  $t$ .  $i_t^{x+1,g,o}$ ,  $u_t^{x+1,g,o}$ ,  $d_t^{x+1,g,o}$  og  $s_t^{x+1,g,o}$  angiver sandsynlighederne for hver af de vitale begivenheder. Idet det sædvanligvis antages, at befolkningens bevægelser eller de vitale begivenheder finder sted jævnt fordelt over året, antages risikogruppen at udgøres af mediobefolkningen, dvs.

$$\begin{aligned} E_t^{x,g,o} &= \frac{N_{t+1}^{x+1,g,o} + N_t^{x,g,o}}{2} & (5) \\ &= \frac{1}{2} \left( N_t^{x,g,o} + \bar{I}_t^{x+1,g,o} + I_t^{x+1,g,o} - U_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} - S_t^{x+1,g,o} - D_t^{x+1,g,o} + N_t^{x,g,o} \right) \\ &= N_t^{x,g,o} + \frac{1}{2} \left( \bar{I}_t^{x+1,g,o} + I_t^{x+1,g,o} - U_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} - S_t^{x+1,g,o} - D_t^{x+1,g,o} \right) \end{aligned}$$

Sandsynlighederne anvendt i (4) benævnes derfor mediasandsynligheder. Disse fastlægges ud historiske data for demografiske strømme og risikogrupper og antages herefter at være eksogene i fremskrivningen. Fastlæggelse af de demografiske begivenheder i fremskrivningen sker principielt ved, at der hver periode løses et ligningssystem bestående af ligningerne i (4) og (5), hvorefter næste periodes primobefolkning kan beregnes ved (2). Dette er en noget omstændig og ressourcerelevende fremgangsmåde, hvorfor det nedenfor vises, at der alternativt kan anvendes en ækvivalent tilgang, hvor der udelukkende indgår udtryk i risikogruppebegrebet, der er kendte primo fremskrivningsåret.

Indsættes (5) i (4) fås

$$\begin{aligned}
I_t^{x+1,g,o} &= i_t^{x+1,g,o} \left( N_t^{x,g,o} + \frac{1}{2} \left( \bar{I}_t^{x+1,g,o} + I_t^{x+1,g,o} - U_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} - S_t^{x+1,g,o} - D_t^{x+1,g,o} \right) \right) \quad (6) \\
U_t^{x+1,g,o} &= u_t^{x+1,g,o} \left( N_t^{x,g,o} + \frac{1}{2} \left( \bar{I}_t^{x+1,g,o} + I_t^{x+1,g,o} - U_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} - S_t^{x+1,g,o} - D_t^{x+1,g,o} \right) \right) \\
D_t^{x+1,g,o} &= d_t^{x+1,g,o} \left( N_t^{x,g,o} + \frac{1}{2} \left( \bar{I}_t^{x+1,g,o} + I_t^{x+1,g,o} - U_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} - S_t^{x+1,g,o} - D_t^{x+1,g,o} \right) \right) \\
S_t^{x+1,g,o} &= s_t^{x+1,g,o} \left( N_t^{x,g,o} + \frac{1}{2} \left( \bar{I}_t^{x+1,g,o} + I_t^{x+1,g,o} - U_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} - S_t^{x+1,g,o} - D_t^{x+1,g,o} \right) \right)
\end{aligned}$$

Ved løsning af (6) med hensyn til de fire endogene variable fås

$$\begin{aligned}
I_t^{x+1,g,o} &= \hat{i}_t^{x+1,g,o} \left( N_t^{x,g,o} + \frac{1}{2} \left( \bar{I}_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} \right) \right) \quad (7) \\
U_t^{x+1,g,o} &= \hat{u}_t^{x+1,g,o} \left( N_t^{x,g,o} + \frac{1}{2} \left( \bar{I}_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} \right) \right) \\
D_t^{x+1,g,o} &= \hat{d}_t^{x+1,g,o} \left( N_t^{x,g,o} + \frac{1}{2} \left( \bar{I}_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} \right) \right) \\
S_t^{x+1,g,o} &= \hat{s}_t^{x+1,g,o} \left( N_t^{x,g,o} + \frac{1}{2} \left( \bar{I}_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o} \right) \right)
\end{aligned}$$

hvor

$$\begin{aligned}
\hat{i}_t^{x+1,g,o} &= \frac{i_t^{x+1,g,o}}{1 - \frac{1}{2} \left( i_t^{x+1,g,o} - s_t^{x+1,g,o} - u_t^{x+1,g,o} - d_t^{x+1,g,o} \right)} \quad (8) \\
\hat{u}_t^{x+1,g,o} &= \frac{u_t^{x+1,g,o}}{1 - \frac{1}{2} \left( i_t^{x+1,g,o} - s_t^{x+1,g,o} - u_t^{x+1,g,o} - d_t^{x+1,g,o} \right)} \\
\hat{d}_t^{x+1,g,o} &= \frac{d_t^{x+1,g,o}}{1 - \frac{1}{2} \left( i_t^{x+1,g,o} - s_t^{x+1,g,o} - u_t^{x+1,g,o} - d_t^{x+1,g,o} \right)} \\
\hat{s}_t^{x+1,g,o} &= \frac{s_t^{x+1,g,o}}{1 - \frac{1}{2} \left( i_t^{x+1,g,o} - s_t^{x+1,g,o} - u_t^{x+1,g,o} - d_t^{x+1,g,o} \right)}
\end{aligned}$$

De korrigerede sandsynligheder benævnes primosandsynligheder, idet de udelukkende baserer sig på primostørrelser og eksogene variable. Disse kan historisk beregnes som

$$\begin{aligned}\hat{i}_t^{x+1,g,o} &= \frac{I_t^{x+1,g,o}}{\hat{E}_t^{x,g,o}} & (9) \\ \hat{u}_t^{x+1,g,o} &= \frac{U_t^{x+1,g,o}}{\hat{E}_t^{x,g,o}} \\ \hat{d}_t^{x+1,g,o} &= \frac{D_t^{x+1,g,o}}{\hat{E}_t^{x,g,o}} \\ \hat{s}_t^{x+1,g,o} &= \frac{S_t^{x+1,g,o}}{\hat{E}_t^{x,g,o}}\end{aligned}$$

hvor det primobaserede risikogruppebegreb er givet ved

$$\hat{E}_t^{x,g,o} = N_t^{x,g,o} + \frac{1}{2}(\bar{I}_t^{x+1,g,o} + \bar{S}_t^{x+1,g,o}) \quad (10)$$

Ved at benytte de historiske sandsynligheder beregnet som i (9)<sup>6</sup> kombineret med det primobaserede risikogruppebegreb, kan strømmene i (7) og efterfølgende næste års primobefolkning fastlægges i fremskrivningen ved (2).

Indsættes (10) og (7) i (5) fås nedenstående relation mellem den primo- og mediobaserede risikogruppe

$$E_t^{x,g,o} = \left(1 + \frac{1}{2}(\hat{i}_t^{x+1,g,o} - \hat{s}_t^{x+1,g,o} - \hat{u}_t^{x+1,g,o} - \hat{d}_t^{x+1,g,o})\right) \hat{E}_t^{x,g,o}$$

Såfremt den mediobaserede risikogruppe anvendes i fremskrivningen, er det for konsistensens skyld centralt, at de i fremskrivningen anvendte sandsynligheder også historisk er baseret på samme risikogruppebegreb og således er beregnet som i (9).

Definitionen af (10) varierer mellem oprindelsesgrupper, idet variablene ikke er aktuelle for samtlige grupper. Der gælder således, jf. afsnit 2.1, at

$$\begin{aligned}\hat{E}_t^{x,g,o} &= N_t^{x,g,o} + \frac{1}{2}\bar{I}_t^{x+1,g,o} \quad \text{for } o \in \{ixn, iwn\} \\ \hat{E}_t^{x,g,o} &= N_t^{x,g,o} + \frac{1}{2}\bar{S}_t^{x+1,g,o} \quad \text{for } o \in \{dad, idx, iwd, dxd, dwd\} \\ \hat{E}_t^{x,g,o} &= N_t^{x,g,o} \quad \text{for } o \in \{dan, dxn, dwn\}\end{aligned}$$

### 2.2.2 Nyfødte

Nedenstående vedrører både individer netop født i Danmark og individer, der er indvandret før det først fyldte år. Det antages, at børn fødes, dør, vandrer og

---

<sup>6</sup> Vedrørende dødsandsynligheder vil disse pga. tyndt data ikke blive beregnet for hver oprindelsesgruppe, men blot for befolkningen som helhed. Historisk beregning af disse dødsandsynligheder omtales i afsnit 2.2.3.

skifter statsborgerskab jævnt fordelt over året. Risikogruppen bestående af den 0-årige mediobefolkning kan med afsæt i (2) dermed udtrykkes som

$$E_t^{0,g,o} = \frac{1}{2} \left( F_t^{g,o} + \bar{I}_t^{0,g,o} + I_t^{0,g,o} - U_t^{0,g,o} + \bar{S}_t^{0,g,o} - S_t^{0,g,o} - D_t^{0,g,o} \right) \quad (11)$$

Antallet af fødte med givet køn og oprindelse antages i det følgende for værende eksogent. I afsnit 2.2.4 og 2.2.5 gennemgås henholdsvis hvorledes antallet af fødte fastlægges ud fra fertilitetsrater og hvorledes dette antal fordeles på oprindelse ud fra en sandsynlighedsfordeling for barnets oprindelse givet moderens.

Indvandring, udvandring og statsborgerskabsskifte kan ved anvendelse af ovenstående risikogruppe umiddelbart udtrykkes som

$$\begin{aligned} I_t^{0,g,o} &= i_t^{0,g,o} \frac{1}{2} \left( F_t^{g,o} + \bar{I}_t^{0,g,o} + I_t^{0,g,o} - U_t^{0,g,o} + \bar{S}_t^{0,g,o} - S_t^{0,g,o} - D_t^{0,g,o} \right) \quad (12) \\ U_t^{0,g,o} &= u_t^{0,g,o} \frac{1}{2} \left( F_t^{g,o} + \bar{I}_t^{0,g,o} + I_t^{0,g,o} - U_t^{0,g,o} + \bar{S}_t^{0,g,o} - S_t^{0,g,o} - D_t^{0,g,o} \right) \\ S_t^{0,g,o} &= s_t^{0,g,o} \frac{1}{2} \left( F_t^{g,o} + \bar{I}_t^{0,g,o} + I_t^{0,g,o} - U_t^{0,g,o} + \bar{S}_t^{0,g,o} - S_t^{0,g,o} - D_t^{0,g,o} \right) \end{aligned}$$

Som nævnt i afsnit 1.2.1 opdeles dødsfald for 0-årige i de, der indtræder inden for de første 10 dages levetid og de, der indtræder den resterende del af første leveår. Idet dødssandsynligheden er størst umiddelbart efter fødslen og spædbørn primært må formodes at indvandre efter denne periode, mindsker anvendelsen af to dødssandsynligheder risikoen for at overvurdere dødeligheden blandt indvandrende spædbørn. Historisk data for spædbørnsdødelighed afslører desuden forskellige tendenser over tid i de to typer af dødsfald blandt 0-årige, hvorfor en separat fremskrivning af disse tendenser også bidrager til at nuancere udviklingen i dødelighed relativt til en tilgang, der blot anvender en type af dødelighed for 0-årige. Risikogruppen for dødsfald indtrædende umiddelbart efter fødslen antages at være antallet af nyfødte, hvorfor antallet, der dør inden for de første 10 levedage er givet ved

$$D_{t,0}^{0,g,o} = d_{t,0}^{0,g,o} F_t^{g,o}$$

De resterende dødsfald blandt 0-årige kan beregnes ud fra samme risikogruppe, som er anvendt i (12), dvs.

$$D_{t,1}^{0,g,o} = d_{t,1}^{0,g,o} \frac{1}{2} \left( F_t^{g,o} + \bar{I}_t^{0,g,o} + I_t^{0,g,o} - U_t^{0,g,o} + \bar{S}_t^{0,g,o} - S_t^{0,g,o} - D_t^{0,g,o} \right) \quad (13)$$

hvor

$$D_t^{0,g,o} = D_{t,0}^{0,g,o} + D_{t,1}^{0,g,o}$$

Løsning af systemet (12) og (13) giver anledning til de demografiske bevægelser defineret ved

$$\begin{aligned} I_t^{0,g,o} &= \hat{i}_t^{0,g,o} \frac{1}{2} (F_t^{g,o} + \bar{I}_t^{0,g,o} + \bar{S}_t^{0,g,o} - D_{t,0}^{0,g,o}) & (14) \\ U_t^{0,g,o} &= \hat{u}_t^{0,g,o} \frac{1}{2} (F_t^{g,o} + \bar{I}_t^{0,g,o} + \bar{S}_t^{0,g,o} - D_{t,0}^{0,g,o}) \\ D_{t,1}^{0,g,o} &= \hat{d}_{t,1}^{0,g,o} \frac{1}{2} (F_t^{g,o} + \bar{I}_t^{0,g,o} + \bar{S}_t^{0,g,o} - D_{t,0}^{0,g,o}) \\ S_t^{0,g,o} &= \hat{s}_t^{0,g,o} \frac{1}{2} (F_t^{g,o} + \bar{I}_t^{0,g,o} + \bar{S}_t^{0,g,o} - D_{t,0}^{0,g,o}) \end{aligned}$$

hvor primosandsynlighederne er givet ved

$$\begin{aligned} \hat{i}_t^{0,g,o} &= \frac{i_t^{0,g,o}}{1 - \frac{1}{2} (i_t^{0,g,o} - s_t^{0,g,o} - u_t^{0,g,o} - d_{t,1}^{0,g,o})} & (15) \\ \hat{u}_t^{0,g,o} &= \frac{u_t^{0,g,o}}{1 - \frac{1}{2} (i_t^{0,g,o} - s_t^{0,g,o} - u_t^{0,g,o} - d_{t,1}^{0,g,o})} \\ \hat{d}_{t,1}^{0,g,o} &= \frac{d_{t,1}^{0,g,o}}{1 - \frac{1}{2} (i_t^{0,g,o} - s_t^{0,g,o} - u_t^{0,g,o} - d_{t,1}^{0,g,o})} \\ \hat{s}_t^{0,g,o} &= \frac{s_t^{0,g,o}}{1 - \frac{1}{2} (i_t^{0,g,o} - s_t^{0,g,o} - u_t^{0,g,o} - d_{t,1}^{0,g,o})} \end{aligned}$$

Disse beregnes historisk som

$$\begin{aligned} \hat{i}_t^{0,g,o} &= \frac{I_t^{0,g,o}}{\hat{E}_t^{0,g,o}} & (16) \\ \hat{u}_t^{0,g,o} &= \frac{U_t^{0,g,o}}{\hat{E}_t^{0,g,o}} \\ \hat{d}_{t,1}^{0,g,o} &= \frac{D_{t,1}^{0,g,o}}{\hat{E}_t^{0,g,o}} \\ \hat{s}_t^{0,g,o} &= \frac{S_t^{0,g,o}}{\hat{E}_t^{0,g,o}} \end{aligned}$$

hvor det primobaserede risikobegreb her er givet ved

$$\hat{E}_t^{0,g,o} = \frac{1}{2} (F_t^{g,o} + \bar{I}_t^{0,g,o} + \bar{S}_t^{0,g,o} - D_{t,0}^{0,g,o}) \quad (17)$$

Sandsynligheden for at dø inden for de første 10 dage efter fødslen baseres på en risikogruppe bestående af antallet af fødte og beregnes derfor som

$$\hat{d}_{t,0}^{0,g,o} = \frac{D_{t,0}^{0,g,o}}{F_t^{g,o}} \quad (18)$$

### 2.2.3 Identisk dødelighed på tværs af oprindelse

Specielt for dødssandsynlighederne antages som nævnt i afsnit 1.2.1, at disse ikke varierer mellem oprindelsesgrupperne, hvorfor de ikke historisk bestemmes som anført i (9), (16) og (18). Udgangspunktet for de historisk fastlagte dødshyppigheder er således i stedet en risikogruppe, der ikke afhænger af oprindelse. Den generelle sandsynlighed for at dø henholdsvis inden for de første 10 levedage, den resterende del af første leveår og i de øvrige aldersgrupper bestemmes historisk som

$$\hat{d}_{t,0}^{0,g} = \frac{D_{t,0}^{0,g}}{F_t^g} = \frac{\sum_o D_{t,0}^{0,g,o}}{\sum_o F_t^{g,o}} \quad (19)$$

$$\hat{d}_{t,1}^{0,g} = \frac{D_t^{0,g} - D_{t,0}^{0,g}}{F_t^g - D_{t,0}^{0,g} + \frac{1}{2}NI_t^{0,g}} = \frac{D_{t,1}^{0,g}}{F_t^g - D_{t,0}^{0,g} + \frac{1}{2}(N_{t+1}^{0,g} - F_t^g + D_t^{0,g})} = \frac{D_{t,1}^{0,g}}{\frac{1}{2}(N_{t+1}^{0,g} + F_t^g + D_{t,1}^{0,g} - D_{t,0}^{0,g})}$$

$$\hat{d}_t^{x+1,g} = \frac{D_t^{x+1,g}}{N_t^{x,g} + \frac{1}{2}NI_t^{x+1,g}} = \frac{D_t^{x+1,g}}{N_t^{x,g} + \frac{1}{2}(N_{t+1}^{x+1,g} - N_t^{x,g} - D_t^{x+1,g})} = \frac{D_t^{x+1,g}}{\frac{1}{2}(N_{t+1}^{x+1,g} + N_t^{x,g} + D_t^{x+1,g})}$$

hvor nettoindvandringen  $NI_t^{0,g}$  og  $NI_t^{x+1,g}$  er givet som (jf. summation af (2) over oprindelse)

$$NI_t^{0,g} \equiv \bar{I}_t^{0,g} + I_t^{0,g} - U_t^{0,g} = N_{t+1}^{0,g} - F_t^g - D_t^{0,g}$$

$$NI_t^{x+1,g} \equiv \bar{I}_t^{x+1,g} + I_t^{x+1,g} - U_t^{x+1,g} = N_{t+1}^{x+1,g} - N_t^{x,g} - D_t^{x+1,g}$$

Antallet af døde i hver oprindelsesgruppe bestemmes i fremskrivningen som anført i ligning (7), (14) og (18), men hvor  $\hat{d}_{t,0}^{0,g,o}$ ,  $\hat{d}_{t,1}^{0,g,o}$  og  $\hat{d}_t^{x+1,g,o}$  er erstattet af en fremskrivning baseret på de tilsvarende ikke oprindelsesfordelte størrelser i (19). De fremskrevne værdier for dødshyppighederne bestemmes for individer der dør inden for de første 10 levedage ved en simpel ekstrapolation af den historiske trend, mens Lee-Carters metode er anvendt for de resterende 0-årige og øvrige aldersgrupper.

### 2.2.4 Fødsler og fertilitet

Er primobefolkningen  $N_{t+1}^{x+1,g,o}$  i (2) fastlagt, vises det her hvorledes antallet af fødte  $F_t^{g,o}$  efterfølgende kan bestemmes ud fra fremskrevne fertilitetskvotienter og retningslinier for fordeling af nyfødte på oprindelsesstatus givet moderens oprindelse.

Antallet af fødsler blandt  $x+1$ -årige kvinder i oprindelsesgruppe  $\bar{o}$  kan bestemmes som

$$\Phi_t^{x+1,\bar{o}} = \phi_t^{x+1,\bar{o}} E_t^{x,\bar{o},g="kvinder"} = \phi_t^{x+1,\bar{o}} \frac{N_t^{x,\bar{o},g="kvinder"} + N_{t+1}^{x+1,\bar{o},g="kvinder"}}{2}$$

hvor  $E_t^{x,\bar{o},g="kvinder"}$  angiver mediobefolkningen af kvinder bestemt som i (2) og  $\phi_t^{x+1,\bar{o}}$  er den tilknyttede mediobaserede fertilitetskvotient. Fra afsnit 2.2.1 vides det, at ovenstående alternativt kan udtrykkes som

$$\begin{aligned} \Phi_t^{x+1,\bar{o}} &= \phi_t^{x+1,\bar{o}} \left( 1 + \frac{1}{2} \left( \hat{l}_t^{x+1,"kvinder",o} - \hat{s}_t^{x+1,"kvinder",o} - \hat{u}_t^{x+1,"kvinder",o} - \hat{d}_t^{x+1,"kvinder",o} \right) \right) \hat{E}_t^{x,"kvinder",o} \\ &\equiv \hat{\phi}_t^{x+1,\bar{o}} \hat{E}_t^{x,"kvinder",o} \end{aligned}$$

hvor  $\hat{\phi}_t^{x+1,\bar{o}}$  kan betegnes som den primobaserede fertilitetskvotient. Rapportering af samlet fertilitet tager dog generelt udgangspunkt i de mediobaserede kvotienter, hvorfor disse også er anvendt i fremskrivningen sammen med det mediobaserede risikobegreb, der enkelt kan beregnes fra (2), når blot strømmene i (7) er fastlagt, hvilket jo netop ikke fordrer kendskab til antallet af fødte. Det samlede antal fødte i oprindelsesgruppe  $\bar{o}$  bestemmes som i (20), idet man kun medregner fødsler, hvor moderens alder falder i intervallet 15-49 år.

$$\Phi_t^{\bar{o}} = \sum_{x=14}^{48} \Phi_t^{x+1,\bar{o}} \quad (20)$$

Metoden til fastlæggelse af nyfødtes køn og oprindelse gennemgås i nedenstående afsnit.

### 2.2.5 Nyfødtes fordeling på køn og oprindelse

Fødslerne opdeles simpelt på køn ved at antage, at en konstant andel af alle fødte er drenge. Det følger af definitionerne af opdelingen af befolkningen efter oprindelse, at barnets oprindelse afhænger af såvel faderens som moderens oprindelse, jf. afsnit 2.1. Derfor konstrueres på basis af historiske erfaringer en

sandsynlighedsfordeling for barnets oprindelse givet moderens. Denne sammenhæng anvendes i fremskrivningen til at fordele børnene på oprindelsesgrupper.

Til formålet defineres  $\bar{F}_t^g = (F_t^{g,ixn}, \dots, F_t^{g,dad})$  som en rækkevektor af fødte fordelt på barnets oprindelse og tilsvarende  $\bar{\Phi}_t^g = (\Phi_t^{g,ixn}, \dots, \Phi_t^{g,dad})$  som en rækkevektor af fødte fordelt på moderens oprindelse. Der gælder følgende sammenhæng

$$\bar{F}_t^g = \bar{\Phi}_t^g A_t$$

hvor  $A_t$  er en 9x9 matrix<sup>7</sup>, hvor hver række er en sandsynlighedsfordeling over barnets oprindelse givet moderens oprindelse.

Det er langt fra alle kombinationer af oprindelse, der er mulige, f.eks. kan ingen fødte være indvandrere. Matricen  $A$  angivende sandsynligheden for barnets oprindelse givet moderens indeholder derfor et betydeligt antal nuller, jf. Tabel 1.

**Tabel 1. Fordelingen af barnets oprindelse givet moderens oprindelse**

Barn										Række-
Moder	<i>ixn</i>	<i>ixd</i>	<i>lwn</i>	<i>iwd</i>	<i>dxn</i>	<i>dxd</i>	<i>dwn</i>	<i>dwd</i>	<i>da</i>	sum
<i>ixn</i>	0	0	0	0	$\lambda_t^{ixn,dxn}$	$\lambda_t^{ixn,dxd}$	0	0	$\lambda_t^{ixn,da}$	1
<i>ixd</i>	0	0	0	0	$\lambda_t^{ixd,dxn}$	$\lambda_t^{ixd,dxd}$	0	0	$\lambda_t^{ixd,da}$	1
<i>iwn</i>	0	0	0	0	0	0	$\lambda_t^{iwn,dwn}$	$\lambda_t^{iwn,dwd}$	$\lambda_t^{iwn,da}$	1
<i>iwd</i>	0	0	0	0	0	0	$\lambda_t^{iwd,dwn}$	$\lambda_t^{iwd,dwd}$	$\lambda_t^{iwd,da}$	1
<i>dxn</i>	0	0	0	0	$\lambda_t^{dxn,dxn}$	$\lambda_t^{dxn,dxd}$	0	0	$\lambda_t^{dxn,da}$	1
<i>dxd</i>	0	0	0	0	0	0	0	0	1	1
<i>dwn</i>	0	0	0	0	0	0	$\lambda_t^{dwn,dwn}$	$\lambda_t^{dwn,dwd}$	$\lambda_t^{dwn,da}$	1
<i>dwd</i>	0	0	0	0	0	0	0	0	1	1
<i>da</i>	0	0	0	0	0	0	0	0	1	1

Første række i tabellen angiver mulighederne for barnets oprindelse, hvis moderen er indvandrer fra et ikke-vestligt land uden dansk statsborgerskab. I dette tilfælde kan barnet blive efterkommer fra et ikke-vestligt land uden dansk

<sup>7</sup> Det antages, at der ikke fødes personer til restgruppen uden dansk statsborgerskab, *dan*.

statsborgerskab, hvis faderen er indvandrer uden dansk statsborgerskab. Andelen af børn, der får denne oprindelse kaldes  $\lambda_t^{ixn,dxn}$ .

Hvis faderen er indvandrer med dansk statsborgerskab, er der mulighed for, at barnet bliver karakteriseret som efterkommer fra et ikke-vestligt land med dansk statsborgerskab.<sup>8</sup> Andelen af børn med en moder, der er indvandrer fra et ikke-vestligt land, som får denne oprindelse betegnes  $\lambda_t^{ixn,dxd}$ .

Den tredje og sidste mulighed er, at barnet bliver af dansk oprindelse. Det er tilfældet, hvis faderen er født i Danmark og har dansk statsborgerskab (dvs. hvis faderen enten er efterkommer med dansk statsborgerskab eller af dansk oprindelse). Den andel af børnene, der tildeles dansk oprindelse og som har en moder, der er indvandret fra et ikke-vestligt land, benævnes  $\lambda_t^{ixn,da}$ .

Givet Danmarks Statistiks definitioner af befolkningens oprindelse vil alle børn, hvis moder er indvandrer fra et ikke-vestligt land blive karakteriseret ved en af disse tre oprindelser. Der gælder således

$$\lambda_t^{ixn,dxn} + \lambda_t^{ixn,dxd} + \lambda_t^{ixn,da} = 1$$

Det ved anvendelse af matricen fundne  $\bar{F}_t^g$  kan således efterfølgende indsættes i (11), så de øvrige flows for 0-årige kan beregnes.

### 2.3 Eksogene parametre i fremskrivningsmodellen

Som anført i gennemgangen af modellen er det nødvendigt at skønne over den fremtidige udvikling i en række eksogene parametre for at kunne gennemføre en fremskrivning. Fremskrivning af ind-, ud- og statsborgerskabsskiftssandsynlighederne i (9) og (16), omfanget af indvandring uden dansk statsborgerskab, dødsandsynlighederne i (19) samt fertilitetskvoteinterne i afsnit 2.2.4 er derfor centrale. Dødelighed og fertilitet estimeres ved ekstrapolation af den historiske udvikling. Frekvenser for genindvandring, udvandring og statsborgerskabsskifte bestemmes som et gennemsnit over de seneste tre års observationer kombineret med en antagelse om, at denne udvikling fastholdes i fremtiden. Indvandring af indvandrere uden dansk statsborgerskab fastholdes på et givet niveau i fremtiden, der er fastlagt ud fra en særskilt vurdering af hvorvidt de seneste års udvikling indikerer et permanent trendskifte og derfor i et vist omfang bør influere på det fremtidige niveau.

---

<sup>8</sup> Dansk statsborgerskab kan fravælges.

Fremskrivning af de enkelte eksogene variable gennemgås særskilt i denne udgives efterfølgende kapitler. Nedenfor er dog givet en oversigt over, hvad der skal fastlægges:

- $\hat{d}_t^{x,g}$ : Dødelighed fordelt på alder og køn.
- $\phi_t^{x,o}$ : Fertilitet fordelt på alder og moderens oprindelse.
- $\hat{u}_t^{x,g,o}$ : Udvandrings sandsynlighed fordelt på alder, køn og oprindelse.
- $\hat{i}_t^{x,g,o}$ : Genindvandringssandsynlighed fordelt på alder, køn og oprindelse.
- $\hat{s}_t^{x,g,o}$ : Sandsynlighed for at skifte til dansk statsborgerskab fordelt på alder, køn og oprindelsesgrupperne *dan*, *ixn*, *iwn*, *dxn* og *dwn*.
- $\bar{I}_t^{x,g,o}$  for  $o \in \{ixn, iwn\}$ : Indvandring fordelt på alder, køn og oprindelse, der ikke har karakter af genindvandring.

Datagrundlaget anvendt til fremskrivning af ovenstående parametre beskrives i det følgende kapitel.

### 3 Datagrundlaget for fremskrivningsmodellen

Med udarbejdelsen af samordnede fremskrivninger med Danmarks Statistik er datagrundlaget, som nævnt indledningsvist, søgt forbedret gennem udtræk fra PersonStatistikDatabasen (PSD). Databasen opdateres løbende såfremt, der tilkommer nye oplysninger eller korrektioner vedrørende historisk data. Typisk vil dette udmønte sig i, at der kan konstateres små ændringer relativt til det datagrundlag, der anvendtes til 2010-fremskrivningen. Data for demografiske strømme omfatter alt afhængigt af deltaljeringsniveau årene 1990/1992-2010, mens der for beholdninger også haves data for 2011.

For hver af de 10 befolkningsgrupper omfatter data følgende variable:

- Antal mænd og kvinder (fordelt på alder).
- Antal døde mænd og kvinder (fordelt på alder).
- Antal døde piger og drenge fordelt på dage inden for det første leveår.
- Antal fødte drenge og piger (fordelt på moders alder).
- Sandsynlighedsfordeling til fordeling af fødte på oprindelse givet moderens.
- Antal indvandrede mænd og kvinder (fordelt på alder).
- Antal udvandrede mænd og kvinder (fordelt på alder).
- Antal mænd og kvinder, der skifter til dansk statsborgerskab (fordelt på alder).

#### 3.1 Dødelighedsdata

Med overgangen til anvendelse af PSD som datagrundlag, beregnes dødssandsynligheder med udgangspunkt i data herfra, hvorfor dødeligheder udtrukket fra Human Mortality Database (HMD) ikke længere anvendes som estimationsgrundlag i Lee-Carter metoden. Befolkningsfremskrivningen fra 2010 anvendte heller ikke data fra HMD. Som udgangspunkt er datakilden til denne også Danmarks Statistik, men anvendelsen af HMD som kilde var problematisk, idet det nyeste data først var tilgængeligt i slutningen af kalenderåret. Således ville data for 2010 først kunne rekvireres i efteråret 2011. Med det formål at inddrage det nyeste data på en måde, der er konsistent med øvrigt historisk data, blev det seneste års data derfor tidligere konstrueret ud fra rådata fra DST pålagt samme imputeringsprincipper som anvendt af HMD.

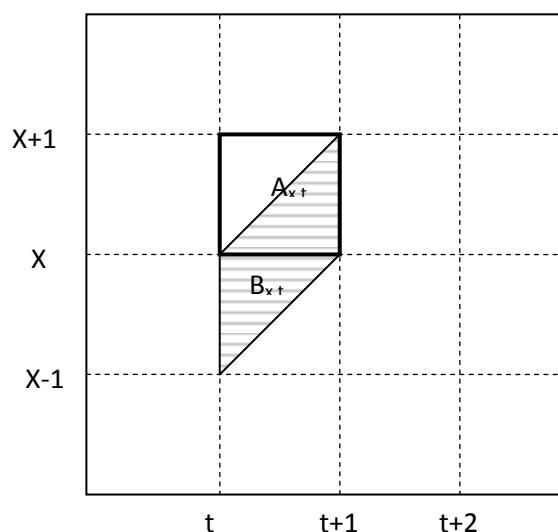
For at forenkle arbejdsgangen og sikre større gennemskuelighed er alt dødelighedsdata siden samordningen af befolkningsfremskrivningerne i 2010 derfor baseret på data leveret direkte fra DST.

Data fra HMD anvendes dog fortsat i nærværende publikation til illustration af udviklingen i dødelighed i perioden forud for år 1990<sup>9</sup>. Det er centralt at bemærke, at dødeligheder fra HMD er opgjort på såkaldte A-grupper, mens dødeligheder for perioden efter år 1990 er opgjort iht. B-gruppe formatet:

*Definition: Dødelighed for alder  $x$  opgjort iht. A-gruppe formatet vedrører antallet af personer, der i løbet af et år er døde med alderen  $x$ . Dødelighed for alder  $x$  opgjort på B-grupper, refererer til de personer, der er døde i løbet af året og som havde alderen  $x - 1$  ved årets begyndelse.*

Det i fremskrivningen anvendte format for data beskrivende demografiske strømme er B-grupper, hvor de demografiske begivenheder er kendetegnet ved at vedrøre kohorter. Figur 2 illustrerer definitionen af data opgjort efter henholdsvis A- og B-gruppe formatet.

**Figur 2. Datakonventioner i HMD-Databasen (A-grupper) og i DREAMS befolkningsfremskrivningsmodel (B-grupper)**



Den i diagrammet optrukne firkant indikerer, hvorledes døde  $x$ -årige på tidspunkt  $t$  opgøres iht. A-gruppeformatet anvendt i HMD. Denne gruppe benævnes  $A_{x,t}$ . Parallelogrammet benævnt  $B_{x,t}$  angiver den tilsvarende B-gruppedødelighed. Bemærk, at det er alderen ultimo år  $t$ , der her angiver, hvilken aldersklasse B-gruppen tilhører.

<sup>9</sup> Fordelen ved HMD er, at databasen bl.a. indeholder årlige dødelighedsdata for Danmark fordelt på alder og køn i en meget lang historisk periode (1835-2009), hvilket i vidt omfang muliggør eksperimenter med undersøgelse af fremskrivningspræcision af de anvendte metoder.

### 3.2 Rapportering af restlevetid

Til rapportering af restlevetid udglattes de estimerede dødelighedsdata over tid og alder, hvorved de bringes på en form, der for langt størstedelen af de historiske år, er sammenlignelig med data i dødelighedstavlerne i Statistikbanken. Denne tilpasning er hensigtsmæssig, idet det er indholdet i de officielle dødelighedstavler, der benyttes til fastlæggelse af indekseringstakten i tilbagerækningsalderen i Velfærdsaftalen 2006 og senest i Reformpakken 2020.

Til og med dødelighedstavlen 2007:2008 er restlevetiderne hos Danmarks Statistik baseret på et gennemsnit af fire B-grupper. De i Figur 3 indtegnede B-grupper danner således udgangspunkt for data for dødelighed og restlevetid for alder  $x$  i dødelighedstavlen dateret  $t$  til  $t + 1$ <sup>10</sup>. Fra og med dødelighedstavlen for 2008:2009 har Danmarks Statistik ændret metode og beregningsgrundlag for beregning af middellevetider. Disse baseres på nu oplysninger fra CPR-registeret og beregninger på C-grupper, jf. Danmark Statistik (2010b)<sup>11</sup>. I modsætning til de tidligere dødelighedstavler er det nye format vanskeligt at approksimere ved medioficering. Således vil de i nærværende publikation beregnede middellevetider for de seneste to historiske år 2008:2009 og 2009:2010 afvige fra DSTs opgørelse. DSTs metode er funderet i individdata og kan ikke anvendes på fremskrevet data.

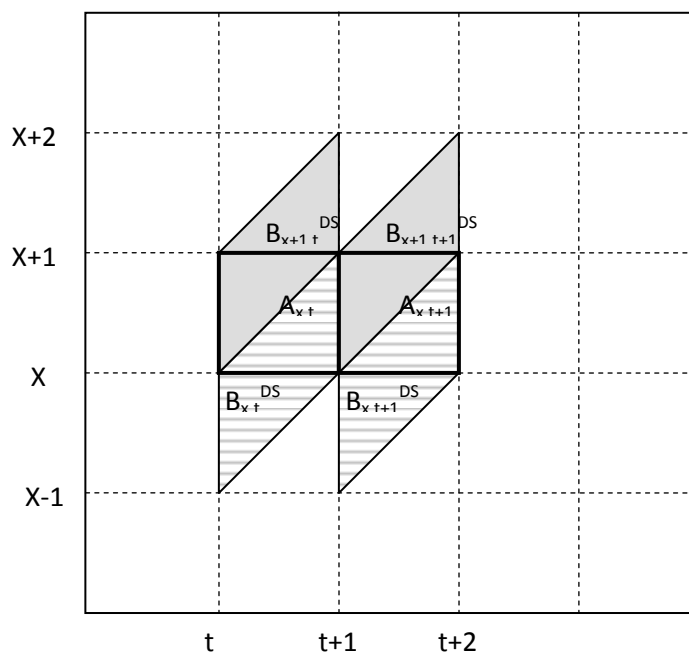
Til og med tavlen 1989:1990 approksimeres DSTs dødelighedstavler gennem medioficering i tid over to A-gruppe dødeligheder, da dette er det tilgængelige datagrundlag for den pågældende periode. I den efterfølgende historiske periode og gennem hele fremskrivningsperioden rapporteres restlevetider baseret på medioficering i tid og alder over fire B-gruppe dødeligheder, netop svarende til tilgangen anvendt af Danmarks Statistik forud for tavlen 2008:2009. Af Figur 3 fremgår det også hvorfor gennemsnittet af A-grupperne  $A_{x,t}$  og  $A_{x,t+1}$  er en rimelig approksimation til dødshyppigheden for en  $x$ -årig som denne fremgår af Danmarks Statistiks dødelighedstavle dateret  $t$  til  $t + 1$ .

---

<sup>10</sup> Danmarks Statistik (2003): s. 254. For  $x$ -årige beregnes dødeligheden i tavlen  $t$  til  $t + 1$  som følger. Først beregnes en dødshyppighed for  $B_{x,t}^{DS}$  og  $B_{x,t+1}^{DS}$  under ét som summen af de døde i de to grupper divideret med det samlede antal  $x-1$ -årige primo år  $t$  og primo år  $t + 1$ . På tilsvarende vis beregnes dødshyppigheden samlet for  $B_{x+1,t}^{DS}$  og  $B_{x+1,t+1}^{DS}$ , hvor der divideres med antallet af  $x$ -årige primo år  $t$  og primo år  $t + 1$ . Dødshyppigheden for  $x$ -årige konstrueres slutteligt som gennemsnittet over de to beregnede hyppigheder. Herefter kan restlevetiden beregnes.

<sup>11</sup> Se også <http://www.dst.dk/Vejviser/Portal/Befolkning/Middellevetid.aspx>.

Figur 3. Korrespondance mellem medioficerede A- og B-gruppe dødeligheder



Det er centralt at bemærke, at Danmarks Statistiks offentliggørelse af dødeligheden for 0-årige baserer sig på C-grupper i samtlige dødelighedstavler. Man kan derfor ikke opnå fuldstændig approksimation til disse tavler hvad angår dødeligheden og dermed restlevetiden for 0-årige.

*Definition: En dødelighed for alder  $x$  opgjort på C-grupper måler hvor mange personer, som opnår alderen  $x$  i løbet af år  $t$ , der er døde inden de bliver  $x+1$  år.*

## 4 Dødelighed og forventet levetid

Estimation af dødelighed til befolkningsmodellen bygger på den af Lee & Carter (1992) foreslåede metode til fremskrivning af aldersbetingede dødeligheder. Metoden er tidligere blevet brugt af Velfærdskommissionen (2006) og i DREAMs befolkningsfremskrivning siden 2006<sup>12</sup>. I de senere år har den opnået betydelig udbredelse, og anvendes af internationale institutioner som f.eks. FN's Population Division, der foretager sammenhængende langsigtede befolkningsfremskrivninger for alle verdens lande, jf. United Nations (2004). Herudover anvendes metoden til befolkningsfremskrivninger i en række lande, f.eks. Sverige, jf. Statistiska centralbyrån (2003, 2005)<sup>13</sup>, og endelig anvendes den også af en række forskergrupper. Den stigende udbredelse har betydet, at der er foretaget en ganske omfattende afprøvning af metoden på forskellige lande og generelt findes, at metoden har gode fremskrivningsegenskaber.<sup>14</sup>

I forbindelse med udarbejdelsen af DREAMs befolkningsfremskrivning i 2006 blev gennemført en analyse, der fokuserede på betydningen for estimation og fremskrivning af at opdele den samlede historiske dataserie i delperioder. På den baggrund fandtes bl.a., at man kunne opnå den bedste fremskrivningspræcision ved at vælge en relativt kort dataserie fra 1990 og frem. Grundet den relative markante stigning i væksten i restlevetiden fra og med 1995 medfører anvendelse af en dataserie startende før dette tidspunkt, at den fremskrevne kortsigtede udvikling i den aldersbetingede restlevetid vil ligge i underkanten af den seneste historiske udvikling. Anvendelse af Lee-Carter metoden indebærer, at væksten i restlevetiden over tid vil aftage, således at den langsigtede vækst svarer til trenden i en længere dataserie. Fordelen ved at anvende en dataserie omfattende relativt få år er, at udviklingstendensen for perioden fortsætter på kort sigt. Ved anvendelse af længere dataserier omfattende betydelige perioder med lav vækst, vil man i modsætning hertil opnå en noget lavere fremtidig vækst end de seneste år tilsiger. Ulempen ved at anvende en kort dataserie er, at estimationen af udviklingen i den aldersbetingede dødelighed i uhensigtsmæssig grad kan påvirkes af udsving gennem perioden.

I de to første afsnit præsenteres hovedtrækkene i den klassiske Lee-Carter tilgang, mens udvidelser hertil beskrives i det efterfølgende afsnit. I afsnit 4.4 omtales betydningen af valg af dataperiode og estimationsresultater præsenteres i afsnit 4.5. Dernæst redegøres for estimation og fremskrivning af dødsfald blandt de ældste ældre og nyfødte. I afsnit 4.8 og 4.9 præsenteres den af fremskrivningen foranledigede udvikling i aldersfordelt dødelighed. Den

---

<sup>12</sup> Se Danmarks fremtidige befolkning – Befolkningsfremskrivning 2006, DREAM for en evaluering af Lee-Carter metoden, samt dens udbredelse og i litteraturen foreslåede udvidelser.

<sup>13</sup> Norge anvender andre statistiske estimationsmetoder, som har en mere kompleks struktur end Lee-Carter metoden, jf. Keilman, Pham & Hetland (2001).

<sup>14</sup> En oversigt over forskellige anvendte metoder til befolkningsfremskrivning og en fortegnelse over hvilke principper der anvendes i en lang række institutioner findes i Andreev & Vaupel (2006)

resterende del af kapitlet er dedikeret til definition af restlevetid og fremskrivningens indvirkning på udviklingen i denne.

Hovedresultaterne i fremskrivningen er, at middellevetiden frem til år 2100 stiger med 12,8 år for mænd og 9,9 år for kvinder<sup>15</sup>. Der er således stort set tale om en halvering af den absolutte vækst i forhold til erfaringerne fra det 20. århundrede. I forhold til argumentationen i Oeppen & Vaupel (2002), som peger på en konstant årlig absolut vækst i middellevetiden, er der derfor tale om et meget forsigtigt skøn. En væsentlig del af forklaringen på den lavere vækst i middellevetiden er, at betydningen af fortsatte reduktioner i spæd- og småbørnsdødeligheden er mere begrænsede end i den historiske periode, fordi dødeligheden for disse grupper allerede er reduceret betydeligt. Betragtes i stedet udviklingen i restlevetiden for 60-årige, har forskellen mellem væksten i det 20. århundrede og den fremskrevne udvikling for det 21. århundrede det modsatte fortegn. I det 20. århundrede voksede restlevetiden for 60-årige mænd med 4,5 år, mens væksten var på 6,5 år for kvinder. I fremskrivningen fås, at den tilsvarende vækst frem til 2100 er 10,1 år for mænd og 7,9 år for kvinder. Dette afspejler at fremtidige reduktioner i dødelighederne forventes at få relativ større effekt på restlevetiden for ældre end for yngre.

Relativt til sidste års fremskrivning er middellevetiden i år 2100 steget med 0,2 år for mænds vedkommende, mens middellevetiden for kvinder i 2011-fremskrivningen på sigt er 0,4 år længere end i sidste fremskrivning. Ændringen i middellevetiden genfindes i restlevetiden for 60-årige mænd, mens restlevetiden for 60-årige kvinder er øget med 0,3 år.

#### 4.1 Lee-Carter metoden

Den grundlæggende idé i Lee-Carters metode er, at den fremtidige udvikling i en given aldersbetinget dødelighed kan beskrives ud fra den historiske udvikling i de aldersbetingede dødeligheder. Analysen gennemføres på tværs af oprindelse, men for hvert køn.

For at reducere dimensionen i problemet antages, at der er en betydelig regularitet i udviklingen af de aldersspecifikke dødeligheder, således at de kan beskrives ved hjælp af tre elementer: Et aldersspecifikt element, der er uafhængig af tidspunktet, et tidsafhængigt element, benævnt mortalitetsindekset, beskrivende udviklingen i den samlede dødelighed og et

---

<sup>15</sup> I publikationen Befolkningsfremskrivning 2010 var restlevetiderne ikke beregnet med afsæt i medioficerede dødeligheder beregnet iht. beskrivelsen i afsnit 3.2. Således kan niveauerne og ændringerne anført i Befolkningsfremskrivning 2011 afvige fra de, der fremgår af sidste års publikation. I forbindelse med sammenligninger mellem Befolkningsfremskrivning 2010 og 2011 er niveauerne i Befolkningsfremskrivning 2010 dog reberegnet. Samtlige niveauer og ændringer i restlevetiden i nærværende kapitel vedrører sidste kalenderår anvendt til medioficeringen. Således benyttes restlevetid for årene  $x_1:x_2$  og  $y_1:y_2$  til at angive ændringen fra  $x_2$  til  $y_2$ .

tidsafhængigt gennemslag af udviklingen i mortalitetsindekset på dødeligheden i den enkelte aldersgruppe. Konkret antages følgende sammenhæng:

$$\ln(m_t^{x,g}) = a^{x,g} + b^{x,g} k_t^g + \varepsilon_t^{x,g}$$

hvor definitionen af dimensionerne er identisk med de i kapitel 0 anvendte. I øvrigt gælder, at

- $m_t^{x,g}$  er de køns-, alders- og tidsafhængige dødssandsynligheder.
- $a^{x,g}$  er den køns- og aldersafhængige del, som i praksis er et gennemsnit over de logaritmiske dødeligheder i estimationsperioden.
- $k_t^g$  er den køns- og tidsafhængige udvikling i dødeligheden benævnt mortalitetsindekset.
- $b^{x,g}$  vægter gennemslaget af  $k_t^g$  på dødelighed for aldersgruppe  $x$ .
- $\varepsilon_t^{x,g}$  er et køns-, alders- og tidsafhængigt fejlded.<sup>16</sup>

Parameteren  $a^{x,g}$  angiver den grundliggende form på aldersprofilen over tid og beregnes i praksis som gennemsnittet over tid af de logaritmiske dødeligheder. Den aldersafhængige parameter  $b^{x,g}$  angiver afvigelser fra den generelle profil over tid. Central for fremskrivningen er den aldersafhængige parameter  $k_t^g$ , der beskriver udviklingen over tid i den generelle dødelighed og benævnes derfor mortalitetsindekset. Når det fælles niveau ændres over tid, vægter parameteren  $b^{x,g}$  således effekten på de aldersafhængige logaritmiske dødeligheder og tillader således forskellig vækst i dødelighed på tværs af alder. I det mortalitetsindekset vil være aftagende over tid, indikerer en positiv indgang i vektoren  $b^{x,g}$ , at dødeligheden for den pågældende aldersgruppe er aftagende, mens en negativ indgang indikerer det modsatte. Aftager mortalitetsindekset lineært, dvs.  $dk_t^g / dt$  er konstant, vil de logaritmiske dødeligheder også aftage lineært over tid med proportionalitetsfaktoren  $b^{x,g}$ . Dette medfører, at dødelighedshyppighederne er eksponentielt aftagende.

Denne formulering indebærer blandt andet, at forholdet mellem vækstraterne i dødeligheden for to forskellige aldersgrupper,  $x$  og  $y$  er konstant og lig med  $b^{x,g} / b^{y,g}$ . Udviklingen i dødelighederne bestemmes af udviklingen i mortalitetsindekset,  $k_t^g$ . Denne estimeres i Lee-Carter (1992) som en *random walk med drift*. Det betyder, at der i fremskrivningen antages en konstant fremtidig vækstrate i dødeligheden for en given årgang.

---

<sup>16</sup> For den konkrete estimationsprocedure henvises til f.eks. Haldrup (2004).

## 4.2 Estimation af Lee-Carter modellen

I henhold til Lee & Carter (1992) pålægges modellen restriktioner med det formål at bestemme en entydig løsning af parametrene  $a^{x,g}$ ,  $b^{x,g}$  og  $k_t^g$ . Således anvendes følgende to restriktioner, hvor  $t_1$  og  $t_n$  er henholdsvis første og sidste dataår:

$$\sum_{t=t_1}^{t_n} k_t^g = 0$$

$$\sum_x b^{x,g} = 1$$

Indledningsvist estimeres parameteren  $a^{x,g}$  som gennemsnittet over de logaritmiske dødssandsynligheder i dataperioden, dvs.

$$\hat{a}^{x,g} = \frac{\sum_{t=t_1}^{t_n} \ln(m_t^{x,g})}{t_n - t_1 + 1}$$

Efterfølgende beregnes en matrix bestående af de centrerede logaritmiske sandsynligheder som

$$\ln(m_t^{x,g}) - \hat{a}^{x,g}$$

hvorefter denne påføres en Singular Value Decomposition fra hvilken estimater for henholdsvis  $b^{x,g}$  og  $k_t^g$  kan udtrages som egenvektorer.

Således er fundet et initialt bud på modellens parametre. Der er imidlertid ikke nogen garanti for, at der for et givet år gælder, at antallet af døde beregnet ud fra estimationen, svarer til det faktiske antal døde. Derfor indeholder den oprindelige Lee-Carter formulering en efterfølgende opdatering af mortalitetsindekset, der sikrer dette. Haldrup (2004) følger denne procedure. I nærværende fremskrivning er det i stedet valgt at opdatere mortalitetsindekset, så der for hvert år i dataperioden opnås sammenfald med middellevetiden. Denne fremgangsmåde er ligeledes anvendt i Lee & Miller (2001). Årsagen til, at levealderen anvendes i stedet for antallet af døde er, at førstnævnte vurderes som en mere relevant målvariabel. Således er det også fremadrettet udviklingen i restlevetiden, der er interessant frem for antallet af dødsfald.

Ved anvendelse af de initiale parameterestimater for  $a^{x,g}$  og  $b^{x,g}$  bestemmes mortalitetsindekset ved anvendelse af en Newton-Raphson algoritme således, at forventet levealder i hele dataperioden opnås. I tråd med fremgangsmåden i Haldrup (2004) foretages efterfølgende en yderligere modifikation. Idet det nye bud på mortalitetsindekset tages for givet, findes ved OLS regression af dette på

de logaritmiske dødelighedsrater opdaterede bud på de aldersafhængige parametre  $a^{x,g}$  og  $b^{x,g}$ . Dette kan opfattes som første trin i en iterativ proces, der leder til værdier af  $a^{x,g}$  og  $b^{x,g}$ , som er indbyrdes konsistente med  $k_t^g$ . I denne analyse suppleres dette med andet trin i den iterative proces, idet der med de opdaterede værdier for  $a^{x,g}$  og  $b^{x,g}$  beregnes et nyt mortalitetsindeks, så der igen opnås sammenfald med faktisk levealder. Et eksperiment omhandlende fremskrivningspræcisionen afslører, at tilføjelsen af dette ekstra trin forbedrer præcisionen marginalt. Samtidig opnås tilpasning til den faktiske levealder.

Den centrale variabel for fremskrivningen af udviklingen i levetiden er udviklingen i mortalitetsindekset,  $k_t^g$ . Det er derfor centralt, at processen beskrivende udviklingen i dette indeks fastlægges og er statistisk velspecificeret. Denne fastlæggelse baseres på statistiske test suppleret med grafisk inspektion af serien.

Ovenstående procedure udgør den standardiserede estimation af Lee-Carter modellen. Med henblik på yderligere forbedring af fremskrivningspræcisionen gennemføres dog to yderligere korrektioner, der anbefales i den efterhånden omfattende litteratur, der søger at forbedre den klassiske tilgang.

### 4.3 Korrektioner af fremskrivningsmetoden

Som udvidelse til ovenstående foretages i nærværende afsnit to yderligere korrektioner, der søger at afhjælpe nogle åbenbare ulemper ved den klassiske tilgang for estimation af parametrene i Lee-Carter modellen.

For det første sikres, at den aldersbetingede restlevetid i fremskrivningens første år ikke foretager spring i forhold til den aldersbetingede restlevetid i det seneste år i dataperioden. Da Lee-Carter metoden ekstrapolerer med udgangspunkt i estimerede værdier for aldersbetinget dødelighed, kan dette afstedkomme en ikke ubetydelig afvigelse mellem den seneste observation og estimatet for samme. Bell (1997) foreslår en metode til at forhindre sådanne uhensigtsmæssige afvigelser, så spring mellem den historiske og fremskrevne dataserie undgås. Der anvendes en modificeret udgave af denne metode i nærværende fremskrivning.

For det andet kan Lee-Carter metoden føre til uhensigtsmæssige forskelle i de årlige reduktioner af den aldersbetingede dødelighed i aldersgrupper, der ligger i nærheden af hinanden. Dette tilskrives, at effekten på den enkelte aldersspecifikke dødelighed fremkommer ved, at udviklingen i mortalitetsindekset multipliceres med den aldersspecifikke vektor  $b^{x,g}$ . Tilfældige udsving mellem to på hinanden følgende aldre i  $b^{x,g}$  forstærkes derfor uhensigtsmæssigt. Problemet er størst, hvis metoden baseres på en kortere dataserie og samtidig anvendes til længere fremskrivninger. De Jong & Tickle

(2005) foreslår, at der anvendes en udglatningsmekanisme til afhjælpning af problemet. Den neden for anvendte udglatningsmekanisme afviger fra den i den originale fremstilling foreslåede, men har kvalitativt samme effekt.

#### 4.3.1 Bell's metode samt modifikation heraf

Med henblik på at eliminere de spring i dødshyppigheder og dermed restlevetider, der kan forekomme mellem det sidste historiske år og det første fremskrevne, foreslår Bell (1997) anvendelse af en korrektionsmetode. I henhold til Bells fremstilling kan problemet med spring- eller knækpunkter elimineres ved at påføre hele fremskrivningen en biaskorrektion, hvis omfang bestemmes af forskellen mellem nyeste tilgængelige historiske dødeligheder og estimerede dødeligheder for det samme år. Hvis  $t$  således angiver sidste år i den historiske periode, da vil den aldersafhængige bias mellem de logaritmiske dødeligheder være givet som

$$\Delta^{x,g} = \ln(m_t^{x,g}) - \ln(\hat{m}_t^{x,g})$$

hvor  $t$  er det sidste historiske år, mens  $m_t^{x,g}$  og  $\hat{m}_t^{x,g}$  er henholdsvis de observerede og estimerede dødeligheder det pågældende år.

Bell's foreslåede biasjustering indebærer, at den beregnede bias tillægges samtlige logaritmiske dødeligheder i fremskrivningen. Den biasjusterede fremskrivning bliver for fremskrivningsperiode  $s$  dermed

$$\ln(\hat{m}_{t+s}^{x,g})_{\Delta} = \ln(\hat{m}_{t+s}^{x,g}) + \Delta^{x,g}, \quad s > 0$$

Ovenstående kan alternativt formuleres som

$$\begin{aligned} \ln(\hat{m}_{t+s}^{x,g})_{\Delta} &= \ln(\hat{m}_{t+s}^{x,g}) + \ln(m_t^{x,g}) - \ln(\hat{m}_t^{x,g}) \\ &= a^{x,g} + b^{x,g} k_{t+s}^g + \ln(m_t^{x,g}) - a^{x,g} - b^{x,g} k_t^g \\ &= b^{x,g} [k_{t+s}^g - k_t^g] + \ln(m_t^{x,g}) \end{aligned}$$

Det følger, at biaskorrekturen påvirker dødshyppighederne på følgende måde

$$\begin{aligned} \ln(\hat{m}_{t+s}^{x,g})_{\Delta} &= \ln(\hat{m}_{t+s}^{x,g}) + \Delta^{x,g} \\ &\Updownarrow \\ (\hat{m}_{t+s}^{x,g})_{\Delta} &= \exp(a^{x,g} + b^{x,g} k_{t+s}^g) \exp(\Delta^{x,g}) \end{aligned}$$

hvorfor det kan konstateres, at biaskorrekturen fremkommer ved multiplikation af de oprindelige fremskrevne dødeligheder med en aldersafhængig konstant. Som det ses af første linje i systemet ovenfor, er forskellen mellem de biaskorrigerede estimerede logaritmiske dødeligheder og de ikke-biaskorrigerede

estimerede logaritmiske dødeligheder per definition konstant over tid. Tages udgangspunkt i døds hyppighederne kan differencen mellem det biaskorrigerede og det ikke-biaskorrigerede estimat udtrykkes som

$$\left(\hat{m}_{t+s}^{x,g}\right)_{\Delta} - \hat{m}_{t+s}^{x,g} = \hat{m}_{t+s}^{x,g} \left(\exp(\Delta^{x,g}) - 1\right)$$

Eftersom de fremskrevne døds hyppigheder aftager eksponentielt over tid, vil forskellen mellem de biaskorrigerede og ikke-biaskorrigerede døds hyppigheder dermed også følge en eksponentiel funktionsform over tid og konvergere mod et givet niveau. Idet restlevetiden er baseret på udviklingen i dødeligheden, vil forskellen over tid mellem den biaskorrigerede og den ikke-biaskorrigerede restlevetid også være afledt af den eksponentielle form.

Til trods for, at Bells metode eliminerer et eventuelt spring mellem restlevetider i det sidste dataår og første fremskrivningsår, kan metodens berettigelse diskuteres. Ved anvendelse af metoden forbedrer man fremskrivningens egenskaber første år, men idet korrektionen som specificeret ovenfor ikke er tidsafhængig, bør der gøres forsøg på at redegøre for vedvarende konsekvenser og vurdere disse i forhold til gevinsten første år.

Det er i forbindelse med udarbejdelsen af tidligere befolkningsfremskrivninger undersøgt, hvorvidt fremskrivningspræcisionen påvirkes af anvendelse af metoden<sup>17</sup>. Undersøgelserne resulterede i, at en modificeret version af Bell's metode anvendes. Modifikationen indebærer for det første, at korrektionen udelukkende anvendes på 50-99-årige. Idet døds hyppigheden er væsentlig højere i disse aldersgrupper end i de yngre, er det for disse grupper, at korrektionen alt andet lige vil have den største effekt. For det andet søges korrektionens negative effekter på længere sigt reduceret ved at lade effekten af korrektionen aftage over tid. Biaskorrektionen  $\Delta^{x,g}$  multipliceres således med potensfunktionen  $f_t = \gamma^t$ , der for  $\gamma < 1$  er aftagende over tid. Funktionen sikrer desuden, at biaskorrektionen i sidste dataår ( $t = 0$ ) er uændret.

Der er to yderpunkter for potensfunktionen. I tilfældet  $\gamma = 0$  forekommer ingen biaskorrektion af de estimerede og fremskrevne dødeligheder. I det andet yderpunkt er  $\gamma = 1$  og her er  $f_t = 1$  for alle  $t$ , således at Bells biaskorrektion implementeres i sin oprindelige form.

Det er her valgt at fastlægge parameteren  $\gamma$  således, at halveringstiden for korrektionen,  $T_{1/2}$ , bliver 10 år. Idet halveringstiden defineres som

$$T_{\frac{1}{2}} = \frac{\ln(\frac{1}{2})}{\ln(\gamma)}$$

---

<sup>17</sup> Se DREAMs befolkningsfremskrivning 2006.

da kan  $\gamma$  beregnes ved

$$\gamma = \exp\left(\frac{\ln(\frac{1}{2})}{T_{\frac{1}{2}}}\right)$$

som med en halveringstid på 10 år giver værdien 0,93303. For fuldstændighedens skyld bliver den modificerede biasjusterede fremskrivning for fremskrivningsperiode  $s$  dermed

$$\ln(\hat{m}_{t+s}^{x,g})_{\Delta} = \ln(\hat{m}_{t+s}^{x,g}) + \gamma^s \Delta^{x,g}, \quad s > 0, x \geq 50$$

#### 4.3.2 Udglatning af dødeligheder for forskellige aldre i fremskrivningen

For at forhindre, at der ved lange fremskrivningshorisonter baseret på kortere dataserier kan forekomme forholdsvis store og ikke-demografisk forklarlige fluktuationer i dødelighedsprofilerne, påføres fremskrivningen slutteligt en udglattende korrektion. Dermed modificeres de estimerede dødshyppigheder i fremskrivningen, hvilket også indebærer, at de ovenfor præsenterede biasjusterede rater korrigeres yderligere.

Lad  $\ln(\hat{m}_{t+s}^{x,g})_S$  angive de fremskrevne logaritmiske dødeligheder efter den udglattende modifikation. Denne kan for  $s > 0$  udtrykkes som

$$\begin{aligned} \ln(\hat{m}_{t+s}^{x,g})_S &= a^{x,g} + b^{x,g} k_t^g + \tilde{b}^{x,g} [k_{t+s}^g - k_t^g] + \gamma^s \Delta^{x,g} \\ &= a^{x,g} + b^{x,g} k_t^g + \tilde{b}^{x,g} [k_{t+s}^g - k_t^g] + \gamma^s \Delta^{x,g} + b^{x,g} k_{t+s}^g - b^{x,g} k_{t+s}^g \\ &= \ln(\hat{m}_{t+s}^{x,g}) + \tilde{b}^{x,g} [k_{t+s}^g - k_t^g] + \gamma^s \Delta^{x,g} - b^{x,g} [k_{t+s}^g - k_t^g] \\ &= \ln(\hat{m}_{t+s}^{x,g})_{\Delta} + (\tilde{b}^{x,g} - b^{x,g}) [k_{t+s}^g - k_t^g] \end{aligned}$$

hvor  $t$  her angiver det sidste år for hvilket data er tilgængeligt og  $\Delta^{x,g} = 0$  for  $x < 50$ . Parameteren  $\tilde{b}^{x,g}$  er en udglattet version af estimatet  $b^{x,g}$ . I praksis er  $\tilde{b}^{x,g}$  en såkaldt *cubic spline* af  $b^{x,g}$ , hvilket indebærer, at der mellem samtlige punkter i  $b^{x,g}$  er tilpasset et polynomium af tredje grad under antagelser om minimering af såvel kurvatur i  $\tilde{b}^{x,g}$  som af kvadratisk afstand til indgangene i  $b^{x,g}$ .<sup>18</sup>

Udglatningen indebærer, at afvigelser i logaritmen af dødeligheden fra en alder til den næste reduceres eller elimineres. Derimod påvirker udglatningsproceduren ikke det fænomen, at estimationen kan indebære, at

<sup>18</sup> En formel definition af *cubic spline* begrebet gives i afsnit 5.3 omhandlende estimation af fertilitet.

reduktionen i den aldersbetingede dødelighed er voksende med alderen for visse aldersgrupper.

#### 4.4 Valg af dataperiode

I forbindelse med DREAMs befolkningsfremskrivning fra 2006 blev der lavet en analyse af forskellige dataperioders betydning for fremskrivningen af dødelighed. Den historiske udvikling viser, at både det gennemsnitlige fald i (logaritmen til) dødelighederne og sammensætningen af faldet på aldersgrupper har tendens til at være konstant over lange perioder. At dette er tilfældet for stort set alle lande, er netop årsagen til Lee-Carter metodens betydelige succes og udbredelse. Samtidig viser analysen af dødelighedsudviklingen i Danmark, at der er givne forholdsvis kortvarige perioder, hvor stabiliteten i dødelighedsudviklingen ophører for herefter at blive genetableret med en ændret vækstrate og en ændret aldersmæssig sammensætning.

Lee-Carter metoden fører til, at fremskrivningen vil have tendens til at forlænge det gennemsnitlige observerede årlige fald i (logaritmen til) dødelighederne igennem estimationsperioden til fremskrivningsperioden. Samtidig vil alderssammensætningen af faldet i dødelighederne også have tendens til at afspejle den gennemsnitlige aldersfordeling af dødelighedsfaldet. Analysen af den estimerede effekt på faldet i de aldersafhængige dødeligheder viser, at estimererne af  $b^{x,s}$  og  $k_t^s$  ikke er robuste over for valg af dataperiode. Således kan valg af dataperiode derfor have afgørende indflydelse på fremskrivningens resultater.

Baseres fremskrivningen på den seneste periode med ensartet vækst i restlevetiden, hvilket omfatter årene fra 1995 og frem, vil det indebære, at den vil have tendens til at forlænge udviklingen fra denne periode. Da data indikerer, at strukturen i dødelighedsudviklingen fastholdes i længere perioder, er der grund til at forvente, at en fremskrivning baseret på denne dataserie vil have de bedste fremskrivningsegenskaber på kortere sigt.

Årsagen, til at det kan være relevant at inddrage længere dataperioder, er både den betydelige usikkerhed med hensyn til estimererne og hensynet til det længere sigt. Den seneste periode er karakteriseret ved at have de største fald i (logaritmen til) dødelighederne målt over alle perioder i det 20. århundrede. Der er derfor en risiko for, at en kort dataserie vil have en tendens til at overvurdere den årlige vækst i middellevetiden på langt sigt.

Det endelige valg af længden af dataserien er derfor et kompromis mellem den undervurdering af middellevetidsvæksten på kortere sigt, som en lang dataserie medfører og den mulige fremtidige overvurdering af middellevetidsvæksten, som anvendelsen af en kort dataserie kan føre til.

I lighed med de generelle anbefalinger i litteraturen er det derfor valgt at basere fremskrivningens hovedforløb på en kort dataperiode, som hovedsageligt dækker perioden efter 1995. Af hensyn til usikkerheden på de aldersbetingede estimater er det nødvendigt at udvide antallet af observationer. Den anvendte dataserie omfatter perioden 1990-2010. En fremskrivning baseret på denne tidsserie har for både mænd og kvinder den egenskab, at den langsigtede vækstrate i middellevetiden konvergerer mod vækstraten for en fremskrivning baseret på en længere dataperiode. På det korte sigt indebærer denne fremskrivning en vis opbremsning i væksten i middellevetiden i forhold til den historiske trend.

## 4.5 Estimationsresultater

I det følgende gennemgås estimationsresultaterne baseret på dataserien omfattende årene 1990-2010. De aldersfordelte døds hyppigheder beregnes indledningsvist som anført i ligning (19) afsnit 2.2.3. Det er udelukkende døds sandsynlighederne for individer, der ikke dør inden for de første 10 levedage, der inddrages i Lee-Carter estimationen, dvs.  $\hat{d}_{t,1}^{0,g}$  og  $\hat{d}_t^{x+1,g}$ . Tilpasningen af modellens parametre til middellevetiden, jf. afsnit 4.2, er med udeladelsen af dødsfald inden for de første 10 levedage i realiteten en tilpasning til restlevetiden for et individ, der har overlevet de første 10 dage.

Med udeladelsen af dødsfald inden for de første 10 levedage i Lee-Carter estimationen tillades som tidligere nævnt estimation og fremskrivning af en særskilt trend for dødelighed indtrædende umiddelbart efter fødslen, hvilket beskrives i afsnit 4.7. Det er udelukkende dødeligheder for 0-99-årige, der inddrages i Lee-Carter estimationen. Dette skal tilskrives, at datatætheden for ældre aldersgrupper er relativ begrænset. Alternativt er der derfor valgt en særskilt ekstrapolativ tilgang til modellering af dødelighedsudviklingen for de ældste ældre, som baseres på resultatet af Lee-Carter fremskrivningen. Den anvendte fremgangsmåde er beskrevet i afsnit 4.6.

Nedenstående modeller for mortalitetsindekset,  $k_t^g$ , findes efter test at være velspecificerede for fremskrivningerne for hhv. mænd og kvinder. Standardafvigelsen på de estimerede koefficienter er angivet i parentes

$$\begin{aligned} \Delta k_t^{mænd} &= -2,7225 & \Delta k_t^{kvinder} &= -2,5287 \\ & (0,4276) & & (0,6695) \\ \sigma_{\varepsilon}^{mænd} &= 1,9121 & \sigma_{\varepsilon}^{kvinder} &= 2,9941 \end{aligned}$$

Fremskrivningen af mortalitetsindekset er for både mænd og kvinder beskrevet ved en *random walk med drift*. Den valgte proces for kvinder adskiller sig fra den, der er anvendt til DREAMs seneste fremskrivninger. Tidligere var ovenstående proces for kvinder præget af autokorrelation i residualerne, hvorfor en 1. ordens

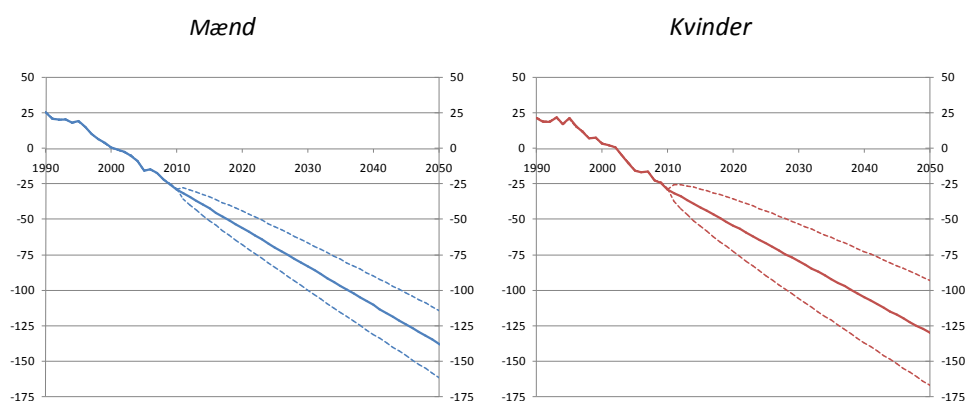
*autoregressiv proces med trend* alternativt anvendtes til at beskrive udviklingen i mortalitetsindekset.

De estimerede og fremskrevne værdier for  $k_t^g$  er sammen med det tilhørende 95 % konfidensinterval illustreret i Figur 4. Mortalitetsindekset er den eneste kilde til angivelse af statistisk usikkerhed i fremskrivningen, og det kan således synes utilfredsstillende, at der ikke er knyttet usikkerhed til de aldersafhængige parametre  $a^{x,g}$  og  $b^{x,g}$ . Dette er dog en generel egenskab ved ikke-stokastiske modeller, hvor de vitale begivenheder ikke tilknyttes en sandsynlighed. At begivenhederne ikke tillægges en bestemt sandsynlighed for at indfinde sig i fremtiden er desuden det, der kendetegner en fremskrivning frem for en prognose.

Sammen med de estimerede værdier af gennemslaget på de aldersbetingede dødeligheder,  $b^{x,g}$ , af et givet fald i dødelighedsindekset,  $k_t^g$ , kan dette anvendes til at finde de aldersbetingede procentvise reduktioner i dødelighederne i de to fremskrivninger. Bemærk, at det fremgår af Figur 4,

hvorledes restriktionen  $\sum_{t=I_1}^{I_n} k_t^g = 0$  er opfyldt i estimationsperioden.

**Figur 4. Estimeret og fremskrevet mortalitetsindeks samt 95 % konfidensinterval**



Kilde: Egne beregninger

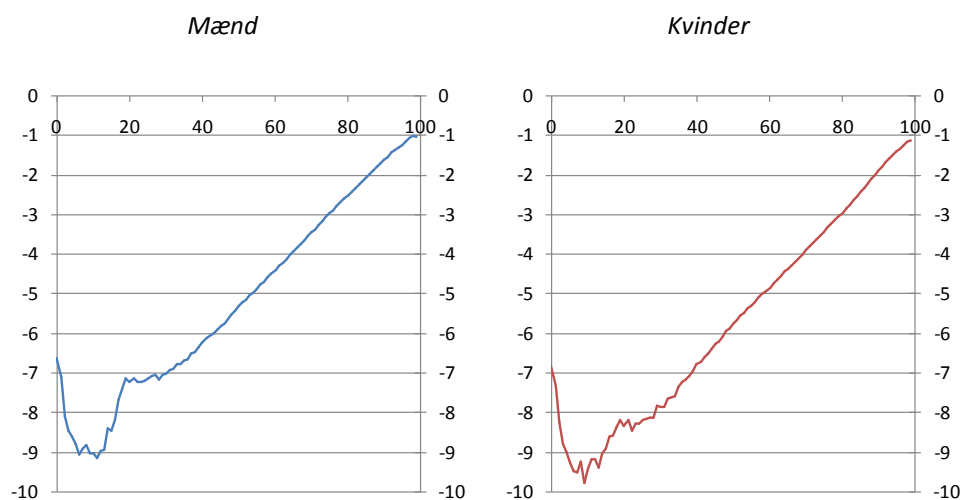
I Figur 5 og Figur 6 er afbilledet estimerterne for henholdsvis den gennemsnitlige aldersbetingede dødelighed,  $a^{x,g}$ , samt den aldersbetingede vægtning,  $b^{x,g}$ , af ændringen i mortalitetsindekset,  $k_t^g$ . Endvidere er i Figur 6 angivet udviklingen i  $\tilde{b}^{x,g}$ , der udglatte  $b^{x,g}$  over alder, jf. afsnit 4.3.2.

Det ses, at den gennemsnitlige aldersbetingede dødelighed for både mænd og kvinder er forholdsvis glat, og at fluktuationer primært indtræffer i aldersklasser

med meget få dødsfald. Endvidere kan det bemærkes, at der for 30-99-årige er en tendens til at dødeligheden vokser eksponentielt med alderen. Dette svarer til, at risikoen for at dø vokser med en fast procent pr. alderstrin. Fænomenet benævnes Gompertz' lov, jf. Gompertz (1825).

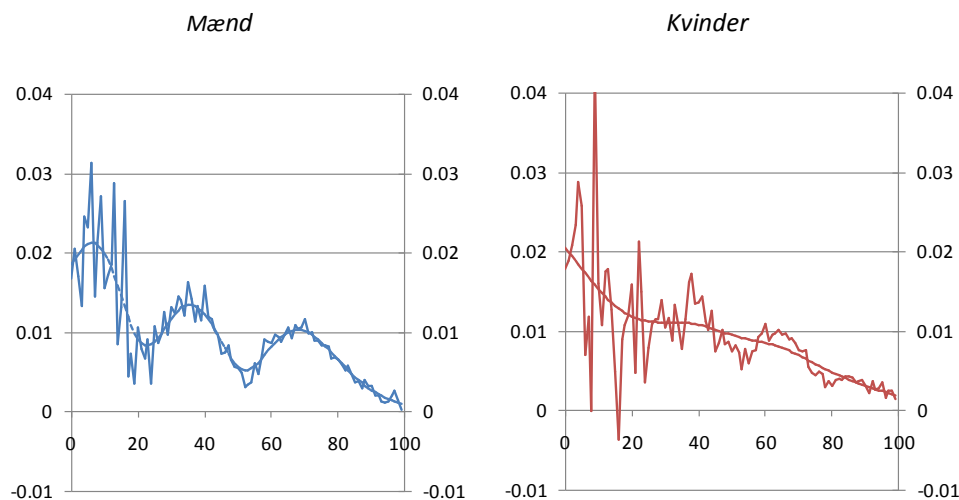
Niveauet for den gennemsnitlige dødelighed er for alle aldersklasser lavere for kvinder, og specielt er ungdomsdødeligheden i aldersklasserne 15-30 år markant højere for mænd, jf. Figur 5.

**Figur 5. Estimerede værdier for  $a(x)$**



Kilde: Egne beregninger

Den aldersbetingede vægtning,  $b^{x \cdot g}$ , af ændringen i mortalitetsindekset er langt mere fluktuerende end den gennemsnitlige aldersbetingede dødelighed, hvilket skyldes den korte dataperiode, jf. Figur 6.

Figur 6. Estimerede og udglattede værdier for  $b(x)$ 

Kilde: Egne beregninger.

For mænd betyder specielt den markant lavere vægtning af aldersklasserne 45-55 år, at den loglineære aldersstruktur på sigt vil bryde samme, hvilket også vil afspejle sig i analysen af fremskrivningsresultaterne. For kvinder er tendensen til, at de procentvise årlige fald er størst for de laveste aldre og gradvist faldende med alderen mere klar. Dette er dog ikke tilfældet for de yngre aldre op til 25 år, hvor der er stor fluktuation og tilmed aldre, hvor ændringen er negativ, hvilket betyder at dødeligheden for de pågældende aldersgrupper vil stige over tid. Bruddet med Gompertz' lov er mindre udtalt for kvinderne end for mændenes vedkommende. Som nævnt i afsnit 4.3.2 er udglatning af  $b^{x,g}$  ingen garanti for, at bruddet med Gompertz' lov undgås, men tendensen mindskes. Det ses også, at fluktuationen i parameteren angivende den aldersbetingede vægtning af gennemslaget af udviklingen i mortalitetsindekset for yngre kvinder er så stor, at den udglatningsform, der er mest optimal, udgøres af en ret linie, jf. Figur 6.

#### 4.6 Konstruktion af dødshyppigheder for de ældste ældre

Grundet relativt tyndt data for personer i alderen 100 år og ældre er det valgt ikke at inkludere disse aldre i Lee-Carter estimationen ovenfor. Alternativt modelleres udviklingen i dødshyppigheder for 100-120-årige ved at ekstrapolere den fremskrevne alderstrend i dødeligheden for de 80-99-årige.

Aldersfordelte dødshyppigheder for 80-99-årige fremskrevet ved Lee-Carter metoden, her benævnt  $\hat{m}_t^{x,g}$ , transformeres indledningsvist som

$$\hat{M}_t^{x,g} = \ln \left( \frac{\hat{m}_t^{x,g}}{1 - \hat{m}_t^{x,g}} \right)$$

Ved en simpel OLS estimation findes efterfølgende for hvert tidspunkt relationen mellem de transformerede dødeligheder og alder, således at den estimerede transformerede dødelighed for de 80-99-årige kan beskrives som

$$\hat{M}_t^{x,g} = \hat{\alpha}_t^g + \hat{\beta}_t^g x$$

Koefficientestimerterne anvendes nu enkelt til at konstruere estimerede transformerede dødeligheder for de ældste aldersgrupper, som efterfølgende kan retransformeres som

$$\hat{m}_t^{x,g} = \frac{\exp\left(\hat{M}_t^{x,g}\right)}{1 + \exp\left(\hat{M}_t^{x,g}\right)}$$

Ekstrapolation af transformerede sandsynligheder sikrer dermed, at de ekstrapolerede dødssandsynligheder,  $\hat{m}_t^{x,g}$ , aldrig vil kunne blive større end 1. Problemet med tilgangen er, at der ingen relation er mellem de ekstrapolerede sandsynligheder på tværs af fremskrivningsperioder. Således er det som ved en fremskrivning anvendende Lee-Carter baserede estimerter ikke nødvendigvis tilfældet, at dødeligheden for en given aldersgruppe udvikler sig monotont (hyppigst aftagende) over tid. Idet risikogruppen for dødsfald blandt de ældste ældre såvel nu som i fremtiden er relativt lille, er inkonsistensen vurderet som værende af mindre betydning.

#### 4.7 Estimation og fremskrivning af dødsfald inden for de første 10 levedage

Som anført skelnes der mellem spædbørnsdødelighed indtrædende hhv. de første 10 dage efter fødslen og dødsfald indtrædende den resterende del af første leveår. Opdelingen er dels funderet i ønsket om at kunne fremskrive dødsfald for 0-årige ved brug af mere end en trend og dels at mindske sandsynligheden for at overvurdere dødsfald blandt 0-årige indvandrere.

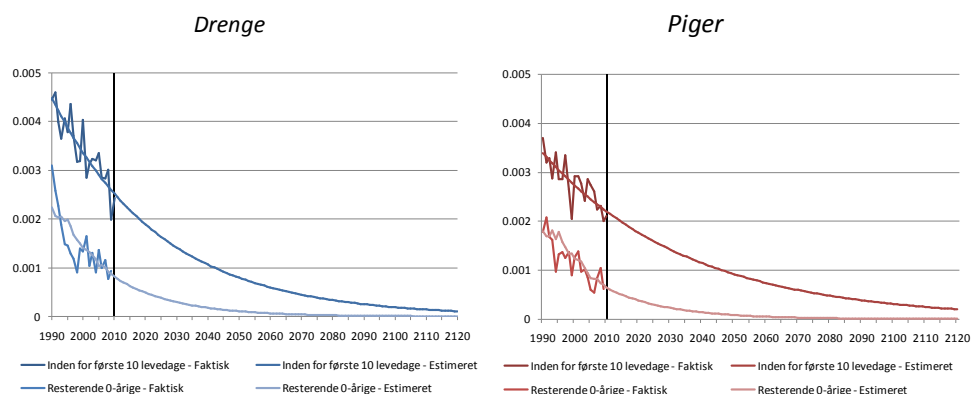
En historisk analyse af de to typer af dødsfald blandt 0-årige viser, at udviklingen over tid ikke er ensartet. Ofte skal dødsfald umiddelbart efter fødslen tilskrives specielle forhold, hvilket i højere grad er funderet i det enkelte individ end i udviklingen i dødeligheden generelt. På denne baggrund kan det retfærdiggøres, at netop denne type dødsfald ikke indregnes i Lee-Carter estimationen.

Idet dødssandsynligheden blandt 0-årige er specielt høj kort tid efter fødslen og 0-årige indvandrere primært kommer til landet efter denne periode, opnås ved differentieringen en mere realistisk vurdering af dødsfald blandt indvandrede spædbørn. Ved kun at betragte en enkeltstående dødelighed blandt 0-årige vil der, grundet det høje bidrag fra helt nyfødte, alt andet lige være tendens til at overvurdere dødsfald blandt indvandrede spædbørn.

De historiske døds sandsynligheder for dødsfald inden for de første 10 levedage beregnes ud fra formelen i (19) for årene 1990-2010<sup>19</sup>. Der udtrages ved anvendelse af OLS en trend for udviklingen over tid i de logaritmisk transformerede sandsynligheder. Denne ekstrapoleres lineært ved brug af de estimerede koefficienter. Efterfølgende påføres fremskrivningen en eksponentiel transformation, hvilket tilvejebringer de fremskrevne spædbørnsdødeligheder<sup>20</sup>.

Såvel faktiske som estimerede sandsynligheder for dødsfald blandt 0-årige er i Figur 7 illustreret for hhv. dødsfald indtrædende inden for de første 10 levedage og resterende dødsfald. For begge køn kan det bemærkes, at sandsynligheden for at dø umiddelbart efter fødslen er ca. dobbelt så stor som sandsynligheden for resterende dødsfald blandt 0-årige. Det kan også bemærkes, at trenden i de seneste år ikke varierer betydeligt mellem de to typer af dødsfald, men at estimationen baseret på hele dataserien dog vil give anledning til en ikke-parallel udvikling af de fremskrevne hyppigheder. Den største gevinst ved opdelingen er tilsyneladende en mindsket risiko for at overvurdere dødsfald blandt indvandrere.

**Figur 7. Hyppigheder for dødsfald indtrædende inden for de første 10 levedage og for resterende 0-årige, faktisk og estimeret**



Kilde: DST og Befolkningsfremskrivning 2011.

## 4.8 Resultater af fremskrivningen

De aldersbetingede dødeligheder er i Figur 8 illustreret for år 1900, år 2010, som er det seneste år i dataserien samt for fremskrivningsårene 2050 og 2100. Der er en udpræget tendens til, at dødeligheden på et givet tidspunkt vokser eksponentielt med alderen og således udvikler sig iht. Gompertz' lov.

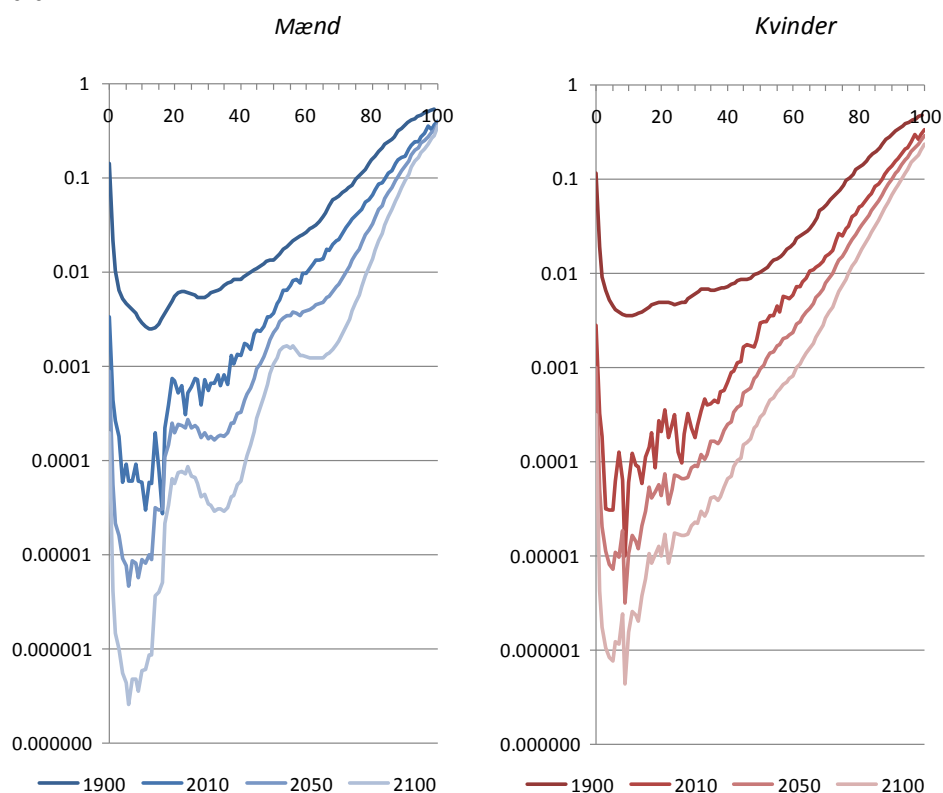
<sup>19</sup> Data for spædbørnsdødelighed fordelt på dage inden for det første leveår vedrører flere kalenderår, men det er her forsimpelt antaget, at alle dødsfald indtrædende inden for de første 10 levedage vedrører samme kalenderår som de øvrige dødsfald blandt 0-årige. Data for dødsfald fordelt på dage efter fødslen vedr. kun personer født i Danmark.

<sup>20</sup> Idet sandsynligheden for spædbørnsdødelighed er meget tæt på nul, vil OLS estimation baseret på en logittransformation (som anvendt i afsnit 4.6) resultere i stort set den samme udvikling som den anvendte logtransformation giver anledning til.

Tendensen til at dødeligheden vokser med en fast procent for hvert alderstrin, gælder i 2010 afhængigt af køn fra omkring 20- til 30-årsalderen og indtil omkring 90-95 års alderen. Det ses, at dødeligheden blandt personer, der er yngre end 30 år, historisk har haft en tendens til at være højere end den lineære (logaritmiske) udvikling tilsiger. Særligt spædbørns- og småbørnsdødelighed har haft tendens til at trække dødsrisikoen op.

Den historiske udvikling i de aldersbetingede dødeligheder gennem det 20. århundrede har indebåret, at dødelighederne er faldet mest for de yngre aldersgrupper (spæd- og småbørnsdødelighed samt ungdomsdødelighed, herunder død under barsel for kvinder). Faldet i dødeligheden har været mindre for de ældre aldersgrupper. Udviklingen har betydet, at Gompertz' lov gradvist har haft tendens til at gælde for et stigende antal aldersgrupper.

**Figur 8. Aldersbetingede dødshyppigheder for mænd og kvinder, logaritmisk skala**



Anm.: Data fra år 1900 stammer fra HMD og således er opgjort på A-gruppe format i modsætning til øvrigt data, der er opgjort på B-gruppe format.

Kilde: Befolkningsfremskrivning 2011, DST og Human Mortality Database (HMD).

De estimerede årlige reduktioner i den aldersbetingede dødelighed er generelt positive for alle aldersgrupper, jf. Figur 6. Det betyder, at fremskrivningen over tid vil føre til, at dødelighederne reduceres for alle aldersgrupper. Det ses, at for både mænd og kvinder reduceres den aldersbetingede dødelighed frem mod år 2050 og år 2100.

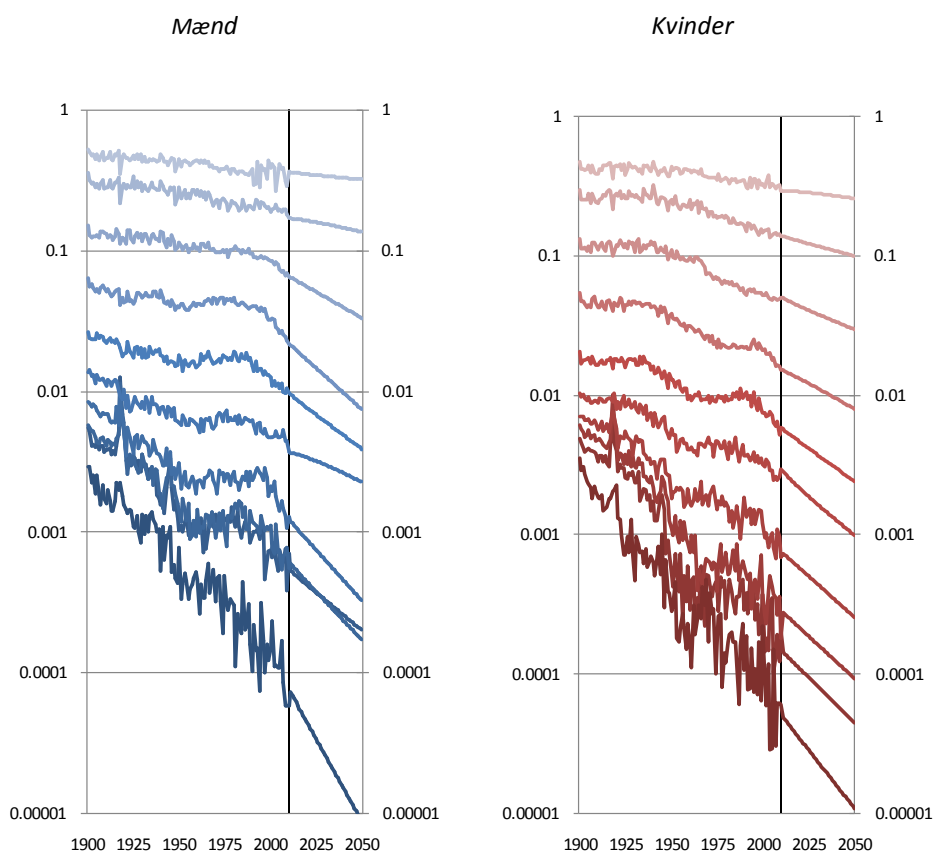
For mænd i aldersgrupperne ca. 23-35 år og 50-60 år observeres i år 2050 stagneret eller faldende dødelighed ved stigende alder, jf. Figur 8. Fænomenet forværres blot frem mod år 2100. Dette medfører et ganske markant brud på Gompertz' lov om, at dødeligheden vokser eksponentielt med alderen. For kvinder ses omvendt, at fremskrivningen leder til en stort set perfekt lineær vækst i logaritmen til dødeligheden med alderen i år 2050 og år 2100.

Det kan derfor sluttet, at valget af den korte dataserie som udgangspunkt for fremskrivningen har en negativ konsekvens for aldersstrukturen i dødeligheden for mænd sammenlignet med serien fra år 2010, mens dette ikke er tilfældet for kvinder.

#### **4.9 Udviklingen i dødeligheden for de enkelte aldersgrupper**

Betragtes den historiske udvikling gennem det 20. århundrede i den aldersbetingede dødelighed for de individuelle aldre, ses det, at der er betydelig forskel på de relative størrelser af faldet i forskellige tidsperioder. Lee-Carter metoden indebærer, at den historiske tendens forlænges således, at den fremskrevne fordeling af faldet i dødelighederne svarer til den historiske udvikling. Betydningen af den valgte estimationsperiode illustreres klarest ved at tage udgangspunkt i logaritmen til dødeligheden for en given alder. Lee-Carter metoden implicerer, at den fremskrevne udvikling i dødeligheden for en given aldersgruppe kan vurderes ud fra trenden i dødeligheden gennem den valgte estimationsperiode.

**Figur 9. Udviklingen i dødshyppigheder 1900-2050 for udvalgte aldre (10, 20, ... , 90, 99 år), mænd og kvinder, logaritmisk skala**



*Anm.: Data fra år 1900-1989 stammer fra HMD og således er opgjort på A-gruppe format i modsætning til øvrigt data, der er opgjort på B-gruppe format. Da dødeligheden er voksende med alderen er udviklingen for 99-årige øverst, efterfulgt af 90-årige osv. Bemærk dog at for mænd bliver dødeligheden for 30-årige lavere end dødeligheden for 20-årige i fremskrivningen.*

*Kilde: Befolkningsfremskrivning 2011, DST og Human Mortality Database (HMD).*

I Figur 9 ses udviklingen i de aldersbetingede dødeligheder for henholdsvis mænd og kvinder gennem det 20. århundrede for 10-årige, 20-årige, ..., 90-årige og 99-årige. Dødelighederne for 99-årige ligger øverst, herefter følger de øvrige aldre i faldende orden.

Figureerne for begge køn viser, som en overordnet tendens, at (logaritmen til) dødeligheden er faldende for alle aldersgrupper, og at der på langt sigt er en tydelig tendens til, at faldet er størst for de yngre aldersgrupper. Der er således langt større spredning i dødelighederne ved indgangen til det 21. århundrede end ved indgangen til det 20. århundrede. Herudover er det for begge køn en udbredt tendens, at faldet i dødelighederne flader ud omkring 1950 og først for alvor begynder at falde igen omkring 1995, jf. Figur 9.

For kvinder er der frem til omkring midten af 1930'erne en tendens til, at det kun er de unge aldersgrupper, hvor der er fald i dødeligheden. Herefter falder dødelighederne generelt i en periode frem til 1960'erne, hvor dødelighedsfaldet aftager og gradvist ophører. Først omkring 1995 er der – ligesom hos mændene – tendens til, at dødeligheden igen falder for alle aldersgrupper.

Af fremskrivningerne fremgår det, at Lee-Carter metoden fortsætter de enkelte aldres udvikling siden 1990. Udtrykt logaritmisk fortsætter den procentvise ændring siden 1990. For mænd medfører en videreførelse af den procentvise reduktion, at strukturen i dødelighedsmønstret efter bare få år begynder at bryde sammen, idet enkelte aldres dødeligheder forholdsvis hurtigt begynder at krydse hinanden. Dette reflekteres også i den nedbrudte aldersprofil for mænd i år 2050, jf. Figur 8. Hos kvinder bevares den initiale struktur i fremskrivningen.

Den historiske tendens til at de senere års reduktion i dødeligheden er vokset relativt, fremgår ligeledes af Figur 9, hvor det såvel for mænd som kvinder gælder, at dødeligheden for 60-årige (og i mindre grad 80-årige) reduceres betydeligt.

## 4.10 Definition af restlevetid

Udviklingen i befolkningens dødelighedsmønster kan alternativt beskrives ved udviklingen i restlevetiden for de enkelte aldersgrupper. Restlevetiden for en given aldersgruppe afhænger af dødeligheden i alle ældre aldersgrupper. Middellevetiden, hvilket er den normale betegnelse for restlevetiden for en 0-årig, indeholder dermed information på en sammenvejet kompakt form om dødeligheden i alle aldersgrupper. Udviklingen i middellevetiden er derfor en måde, hvor på udviklingen i den samlede dødelighed kan udtrykkes.

I nærværende afsnit gennemgås fremgangsmåden for beregning af restlevetid. Det vises i denne forbindelse, at den generelle tilgang, som er funderet i traditionelle demografiske begreber er ækvivalent med en simplere og mindre regnetung fremstilling. Slutteligt vises det, at begge udtryk er konsistente med en intuitiv udlægning af restlevetidsbegrebet.

### 4.10.1 Sædvanlig fremgangsmåde for beregning af restlevetid i dødelighedstavler

I det følgende anvendes i tråd med den klassiske notation betegnelsen  $q_t^{x,g}$  for de køns- og aldersfordelte døds sandsynligheder. Sandsynligheden for at overleve fra alder  $x$  til  $x+1$  er givet ved

$$p_t^{x,g} = 1 - q_t^{x,g}$$

Indledningsvist beregnes en fælles døds sandsynlighed for 0-årige baseret på hyppighederne knyttet til død inden for de første 10 levedage,  $q_{t,0}^{0,g}$ , og de resterende dødsfald for 0-årige,  $q_{t,1}^{0,g}$ .

$$q_t^{0,g} = q_{t,0}^{0,g} + (1 - q_{t,0}^{0,g})q_{t,1}^{0,g}$$

Sandsynligheden for at dø som 0-årig er dermed givet som sandsynligheden for at dø inden for de første 10 levedage tillagt produktet af sandsynligheden for at overleve de første 10 levedage og sandsynligheden for at dø efter de 10 første dage.

Antag at parameteren  $\lambda^{x,g}$  udtrykker den gennemsnitlige andel af året, der gennemlevs af de, der dør. Det antages sædvanligvis, at dødsfald for både mænd og kvinder sker jævnt fordelt over året, hvorfor  $\lambda^{x,g} = 0,5$  for  $x > 0$ . For 0-årige er der dog tendens til, at den gennemlevede del af året er blot ca. 10 %, hvor piger desuden kan forventes at gennemleve en større del af første leveår end drenge. For drenge anvendes  $\lambda^{0,dreng} = 0,111$ , mens størrelsen for piger er  $\lambda^{0,pige} = 0,112$ .<sup>2122</sup>

En person, der fiktivt gennemlever hele sin livscyklus i periode  $t$ , vil have sandsynlighed  $l_t^{x,g}$  for stadig at være i live ved alder  $x$  defineret ved

$$l_t^{x,g} = l_t^{x-1,g} (1 - q_t^{x-1,g}) \quad (21)$$

(21) kan omskrives til

$$l_t^{x,g} = \prod_{s=0}^{x-1} (1 - q_t^{s,g}) \quad (22)$$

For en fiktiv initialbefolkning,  $l_t^{0,g} = 100000$ , kan (21) alternativt udtrykke det antal personer, der overlever til alder  $x$ . Ud af disse vil antallet  $d_t^{x,g}$  dø før de fylder  $x + 1$  år, dvs.

$$d_t^{x,g} = l_t^{x,g} - l_t^{x+1,g} = l_t^{x,g} - l_t^{x,g} (1 - q_t^{x,g}) = l_t^{x,g} q_t^{x,g} \quad (23)$$

<sup>21</sup> Jf. Danmarks Statistik, Befolkningens bevægelser 2001 s. 243. Størrelserne varierer årligt, men ændringerne har blot marginal betydning. Fastholdelse af disse konstanter relativt til sidste befolkningsfremskrivning muliggør en sammenligning mellem middellevetiderne, hvor forskelle ikke kan tilskrives ændringer i denne konstant.

<sup>22</sup> Idet tilpasning til middellevetiden under Lee-Carter estimationen, jf. afsnit 4.2 grundet udeladelsen af dødsfald inden for de første 10 levedage i realiteten er en tilpasning til den forventede restlevetid for en 0-årig, der har overlevet de første 10 levedage, anvendes der i stedet  $\lambda = 0,3$  for begge køn. Givet man med sikkerhed overlever de første 10 levedage, men dør inden den næste fødselsdag, kan man altså regne med at leve næsten en tredjedel af året.  $\lambda$  reduceres således markant ved udeladelsen af den høje dødelighed inden for de første 10 levedage.

Det forventede antal personår,  $L_t^{x,g}$ , der leves i aldersintervallet  $[x, x+1)$  er defineret som

$$L_t^{x,g} = l_t^{x+1,g} + \lambda^{x,g} d_t^{x,g} \quad (24)$$

hvilket netop er sammensat af de, der overlever til alder  $x + 1$  og produktet af den forventede del af året, der gennemlevs af de, der dør og antallet af dødsfald gennem aldersintervallet.

Det forventede antal tilbageværende personår for individer med alder  $x$ ,  $T_t^{x,g}$ , fås netop som summen af det forventede antal personår for hvert af de efterfølgende aldersintervaller, dvs.

$$T_t^{x,g} = \sum_{s=x}^{\infty} L_t^{s,g}$$

hvilket i praksis baseres på eksistensen af en maksimal levealder  $X$  og derfor kan formuleres som

$$T_t^{x,g} = \sum_{s=x}^X L_t^{s,g} \quad (25)$$

Med udgangspunkt i ovenstående kan en person med alder  $x$  forvente at kunne leve yderligere i  $e_t^{x,g}$  år defineret som forholdet mellem det samlede antal forventede tilbageværende personår delt med det antal personår, der er i live med alder  $x$ .

$$e_t^{x,g} = \frac{T_t^{x,g}}{L_t^{x,g}} \quad (26)$$

Man kan forstå dette som, at det samlede antal tilbageværende leveår skal fordeles blandt de medlemmer af populationen, der stadig er i live ved den givne alder. Udtrykket i (26) er netop det traditionelle udtryk for den på tidspunkt  $t$  forventede restlevetid for personer af køn  $g$  med alder  $x$ .

#### 4.10.2 Alternativ beregning af restlevetid

Det kan vises, at (26) under antagelse om eksistens af en maksimal levealder  $X$  er helt ækvivalent med formuleringen af restlevetiden for alder  $x$  på tid  $t$  udtrykt som

$$e_t^{x,g} = \lambda^{x,g} q_t^{x,g} + (1 - q_t^{x,g}) (1 + e_t^{x+1,g}) \quad (27)$$

Den alternative formulering er praktisk, idet den reducerer antallet af mellemregninger betragteligt. Indsættes (24) og (25) i (26) fås

$$e_t^{x,g} = \frac{\sum_{s=x}^X l_t^{s+1,g} + \lambda^{s,g} d_t^{s,g}}{l_t^{x,g}} = \frac{l_t^{x+1,g} + \lambda^{x,g} d_t^{x,g}}{l_t^{x,g}} + \frac{\sum_{s=x+1}^X l_t^{s+1,g} + \lambda^{s,g} d_t^{s,g}}{l_t^{x,g}} \quad (28)$$

(21) og (23) giver nu, at

$$\begin{aligned} e_t^{x,g} &= (1 - q_t^{x,g}) + \lambda^{x,g} q_t^{x,g} + \frac{\sum_{s=x+1}^X l_t^{s+1,g} + \lambda^{s,g} d_t^{s,g}}{l_t^{x,g}} \\ &= (1 - q_t^{x,g}) + \lambda^{x,g} q_t^{x,g} + \frac{l_t^{x+1,g}}{l_t^{x,g}} \frac{\sum_{s=x+1}^X l_t^{s+1,g} + \lambda^{s,g} d_t^{s,g}}{l_t^{x+1,g}} \end{aligned} \quad (29)$$

Benyt igen (24), (25) samt (21) og (26) og få (27)

$$\begin{aligned} e_t^{x,g} &= (1 - q_t^{x,g}) + \lambda^{x,g} q_t^{x,g} + \frac{l_t^{x+1,g}}{l_t^{x,g}} \frac{T_t^{x+1,g}}{l_t^{x+1,g}} \\ &= (1 - q_t^{x,g}) + \lambda^{x,g} q_t^{x,g} + (1 - q_t^{x,g}) e_t^{x+1,g} \\ &= \lambda^{x,g} q_t^{x,g} + (1 - q_t^{x,g}) (1 + e_t^{x+1,g}) \end{aligned}$$

For ethvert tidspunkt,  $t$ , kan restlevetiderne dermed beregnes iterativt efter ovenstående formel. Bemærk, at med antagelsen om eksistensen af en maksimal levealder  $X$ , vil det gælde, at  $q_t^{x,g} = 1$  for alle aldre  $x \geq X$ . I nærværende rapport er restlevetiden beregnet ved anvendelse af en maksimalalder på 110 år, dvs.  $e_t^{110,g} = \frac{1}{2}$ <sup>23</sup>.

Udtrykket ovenfor forstås givetvist bedst ved at se på de ekstreme tilfælde, hvor dødeligheden  $q_t^{x,g}$  måtte være henholdsvis 0 og 1. Såfremt individet dør på tidspunkt  $t$ , svarende til  $q_t^{x,g} = 1$ , da vil restlevetiden for den pågældende aldersgruppe være et halvt år, idet det antages, at man dør ligeligt fordelt over året. Hvis individet overlever til næste tidsperiode, svarende til  $q_t^{x,g} = 0$ , da kan det se frem til at leve yderligere et år tillagt restlevetiden for en et år ældre aldersgruppe.

---

<sup>23</sup> Ved tilpasning til middellevetiden i Lee-Carter estimationen, jf. afsnit 4.2, anvendes en maksimal levealder på 100 år, da estimationen som nævnt kun omfatter data for 0-99-årige.

### 4.10.3 Intuitiv formulering af restlevetid

Udtrykket i (27) er enkelt at beregne, men er måske ikke specielt intuitivt med undtagelse af ekstremtilfældene. Generelt kan man tænke på forventet levealder som det hele antal år man kan forvente at gennemleve, dvs. som

$$\hat{e}_x = \sum_{s=0}^{\infty} s p_{x,s} q_{x+s}$$

hvilket gælder for hvert køn og tidspunkt (pågældende indeks er her undertrykt for at lette notationen).  $p_{x,s}$  angiver ovenfor sandsynligheden for at overleve fra alder  $x$  til  $x + s$ , mens  $q_{x+s}$  angiver sandsynligheden for at dø mellem alder  $x + s$  og  $x + s + 1$ . Eftersom alderen rundes ned til nærmeste hele fødselsdag, vil et individ generelt leve en hvis andel af året efter den sidste fødselsdag. Dette er netop udtrykt ved variabelen  $\lambda^{x,g}$  ovenfor (i det følgende betegnet  $\lambda_x$ ) og den samlede restlevetid for en given alder  $x$ ,  $e_x$ , kan derfor udtrykkes som

$$e_x = \hat{e}_x + \lambda_x = \lambda_x + \sum_{s=0}^{\infty} s p_{x,s} q_{x+s} \quad (30)$$

Det kan vises, at (30) er ækvivalent med (26). Substitution af  $p_{x,s}$  ovenfor giver sammen med definitionen  $\gamma_x = 1 - q_x$ , at

$$\begin{aligned} e_x &= \lambda_x + \sum_{s=0}^{\infty} s p_{x,s} q_{x+s} \\ &= \lambda_x + \sum_{s=0}^{\infty} s q_{x+s} \prod_{i=0}^{s-1} (1 - q_{x+i}) \\ &= \lambda_x + \sum_{s=1}^{\infty} s q_{x+s} \prod_{i=0}^{s-1} (1 - q_{x+i}) \\ &= \lambda_x + \sum_{s=1}^{\infty} s (1 - \gamma_{x+s}) \prod_{i=0}^{s-1} \gamma_{x+i} \\ &= \lambda_x + \sum_{s=1}^{\infty} s \prod_{i=0}^{s-1} \gamma_{x+i} - \sum_{s=1}^{\infty} s \prod_{i=0}^s \gamma_{x+i} \\ &= \lambda_x + \sum_{s=1}^{\infty} s \prod_{i=0}^{s-1} \gamma_{x+i} - \sum_{s=1}^{\infty} (s+1) \prod_{i=0}^s \gamma_{x+i} + \sum_{s=1}^{\infty} \prod_{i=0}^s \gamma_{x+i} \\ &= \lambda_x + \sum_{s=1}^{\infty} s \prod_{i=0}^{s-1} \gamma_{x+i} - \sum_{s=2}^{\infty} s \prod_{i=0}^{s-1} \gamma_{x+i} + \sum_{s=1}^{\infty} \prod_{i=0}^s \gamma_{x+i} \\ &= \lambda_x + \gamma_x + \sum_{s=1}^{\infty} \prod_{i=0}^s \gamma_{x+i} \\ &= \lambda_x + \sum_{s=1}^{\infty} \prod_{i=0}^{s-1} \gamma_{x+i} \\ &= \lambda_x + \sum_{s=1}^{\infty} \prod_{i=0}^{s-1} (1 - q_{x+i}) \end{aligned}$$

Ved anvendelse af (22), (24) og (25) haves, at ovenstående kan skrives som

$$\begin{aligned}
 e_x &= \lambda_x + \sum_{s=1}^{\infty} \frac{\prod_{i=0}^{x+s-1} (1-q_i)}{\prod_{i=0}^{x-1} (1-q_i)} \\
 &= \lambda_x + \sum_{s=1}^{\infty} \frac{l_{x+s}}{l_x} \\
 &= \frac{\lambda_x l_x + \sum_{s=x+1}^{\infty} l_s}{l_x} \\
 &= \frac{\sum_{s=x}^{\infty} \lambda_s l_s - \sum_{s=x+1}^{\infty} \lambda_s l_s + \sum_{s=x+1}^{\infty} l_s}{l_x} \\
 &= \frac{\sum_{s=x}^{\infty} l_{s+1} + \lambda_s (l_s - l_{s+1})}{l_x} \\
 &= \frac{\sum_{s=x}^{\infty} l_{s+1} + \lambda_s d_s}{l_x} \\
 &= \frac{\sum_{s=x}^{\infty} L_s}{l_x} \\
 &= \frac{T_x}{l_x}
 \end{aligned}$$

hvilket jo netop er identisk med (26).

Definitionen af restlevetiden illustrerer tydeligt afhængigheden mellem restlevetid for aldersgruppe  $x$  og dødeligheden for aldersgrupperne  $x$  og ældre. Samtidig kan det bemærkes, at udviklingen i restlevetiden i kraft af værende en ikke-lineær transformation af dødelighederne, vil være præget af antagelsen om en eksponentiel udvikling i sidstnævnte.

Det skal samtidig præciseres, at forventet restlevetid er et demografisk begreb, der ikke relaterer til, hvor gamle eksisterende personer forventes at blive. Begrebet giver et stationært billede af, hvilken alder et fiktivt individ, der gennemlever hele sin livscyklus i tidsperiode  $t$ , kan forventes at blive.

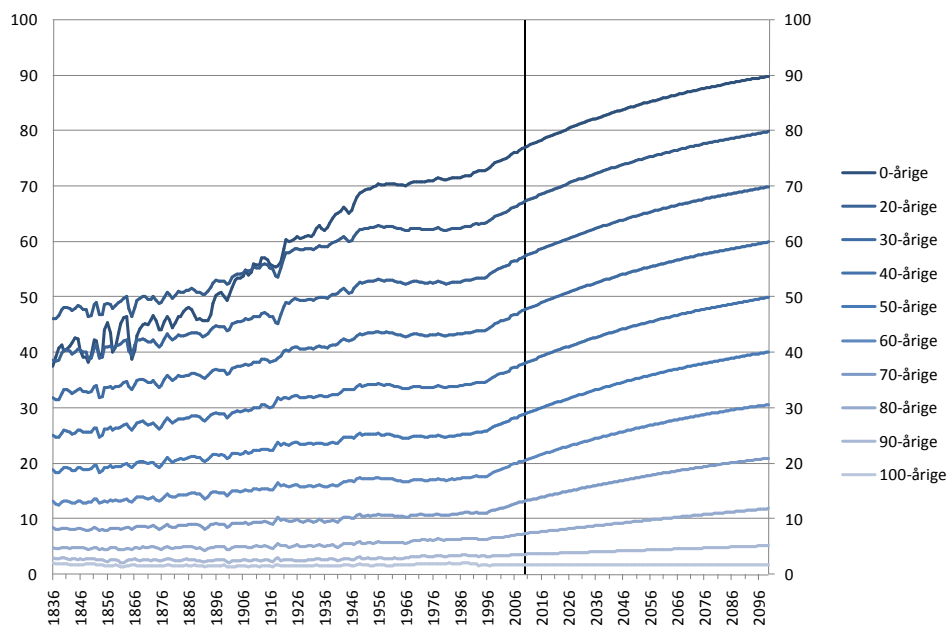
Det kan bemærkes, at en konstant årlig reduktion i logaritmen til de aldersbetingede dødeligheder, som er resultatet af en fremskrivning med Lee-Carter, vil have tendens til at føre til en faldende vækst i middellevetiden. Det

skyldes, at dødelighedernes niveau gradvist reduceres, hvilket betyder, at reduktionen i antallet af døde især koncentrerer sig i de højere aldre, hvilket ikke giver en tilsvarende stigning i ekstra leveår som en reduktion i dødeligheden blandt unge. Middellevetidsvæksten vil derfor have en tendens til at være faldende over tid, jf. Lee (2000). Dette betyder, at Lee-Carter metoden ikke reproducerer den lineære stigning i middellevetiden, som findes i Oeppen & Vaupel (2002).

#### 4.11 Restlevetidsudviklingen for mænd

Restlevetidsudviklingen for mænd i aldersgrupperne 0, 10, ..., 90, 100 år fra 1836<sup>24</sup> og fremefter fremgår af Figur 10. Da restlevetiden netop er en sammenvejning af dødelighederne på et givet tidspunkt, vil brud i trenden i dødelighederne indebære et tilsvarende brud i udviklingen i restlevetiden. Derfor er der i den historiske udvikling i mænds restlevetid to tydelige brud svarende til de brud, der blev identificeret i udviklingen i dødelighederne.

**Figur 10. Restlevetidsudviklingen for mænd, udvalgte aldersgrupper**



Anm.: Restlevetider for årene 1836-1990 er beregnet ud fra dødeligheder fra HMD opgjort på A-gruppe format og medioficeret over tid. Restlevetid for år  $t$  og alder  $a$  er beregnet ud fra dødelighederne for alder  $a$  på tidspunkt  $t-1$  og  $t$ . Restlevetider for 1991 og frem er beregnet med afsæt i dødelighedsdata fra DST opgjort på B-gruppe format og medioficeret over både alder og tid. Restlevetid for år  $t$  og alder  $a$  er beregnet ud fra dødelighederne for alder  $a$  og  $a + 1$  på tidspunkt  $t-1$  og  $t$ . Dateringen anvendt i figuren henviser til sidste kalenderår anvendt til medioficeringen.

Den lodrette streg angiver sidste rene dataår, dvs. 2010 (2009:2010).

Kilde: Befolkningsfremskrivning 2011, DST og Human Mortality Database (HMD).

<sup>24</sup> Dateringen dækker over restlevetid baseret på dødeligheder medioficeret over årene 1835 og 1836 iht. fremgangsmåde beskrevet i afsnit 3.2.

Den historiske udvikling i restlevetiden for mænd kan derfor opdeles i tre adskilte perioder karakteriseret ved forskellige vækstrater for alle aldersgrupper.

Den første periode indtil 1953 er karakteriseret ved en betydelig vækst i den mandlige restlevetid. Vækstraten er stort set den samme for alle aldersgrupper fra 10-årige og derover. For disse aldersgrupper ligger væksten på ca. 0,3 pct. pr. år. Målt absolut er væksten i restlevetiden derfor aftagende med alderen. Således aftager den gennemsnitlige årlige vækst i perioden 1836-1953 fra 3,3 måneder pr. år for 0-årige til 1,5 måneder for 20-årige, hvorefter den falder gradvist med stigende alder. Middellevetidsudviklingen skiller sig ud på grund af et betydeligt større fald i spædbørnsdødeligheden end i dødeligheden for de øvrige aldersgrupper. Stigningen i middellevetiden er derfor omkring 30 pct. i perioden fra 1904 til 1953, eller omtrent dobbelt så høj en vækst som de øvrige aldersgrupper. Udviklingen svarer til en gennemsnitlig årlig stigning i middellevetiden på små 4 måneder. Det bemærkes, at der for de yngre aldersgrupper er et betydeligt knæk i væksten omkring 1918, hvilket skal tilskrives den spanske syge. For 60-årige stiger restlevetiden med gennemsnitligt 0,6 måneder om året fra 1904-1953.

Den anden periode – lavvækstperioden – strækker sig fra omkring 1953 til 1995. I denne periode er der stort set ikke vækst i restlevetiden for aldersgrupperne fra 10 til 60 år. Vækstraten over hele den 40-årige periode ligger på omkring 1,2-2,0 pct. for disse aldersgrupper. Selvom der også er en meget betydelig opbremsning i væksten i middellevetiden, er vækstraten i denne små 5 pct. over den godt 40-årige periode, hvilket svarer til en gennemsnitlig årlig vækst på lidt under en måned. Restlevetiden for 60-årige stiger gennemsnitligt med 0,1 måneder årligt over perioden.

Den tredje og sidste periode er den nuværende og strækker sig foreløbigt fra 1995 til 2010. I denne periode er restlevetiden for alle aldersgrupper vokset mere end i den foregående periode, selvom den foregående periode strækker sig over en årrække, som er ca. 3 gange så lang. Det er endvidere bemærkelsesværdigt, at strukturen i væksten er ændret, således at der nu er en meget tydelig tendens til, at vækstraterne er størst for de ældre aldersgrupper. Middellevetiden for mænd er således vokset med 5,7 pct. siden 1995, restlevetiden for 20-årige mænd med 7,3 pct. og restlevetiden for 60-årige mænd med 16,9 pct. Forskellene i vækstraten betyder, at der kun er begrænsede forskelle i den absolutte vækst for de forskellige aldersgrupper. Således er restlevetiden for en 20-årig mand vokset med 3,9 år over perioden, mens restlevetiden for en 60-årig mand er vokset med 3 år. Gennemsnitligt er middellevetiden vokset med 3,3 måneder om året, mens restlevetiden for 60-årige er vokset med 2,4 måneder årligt.

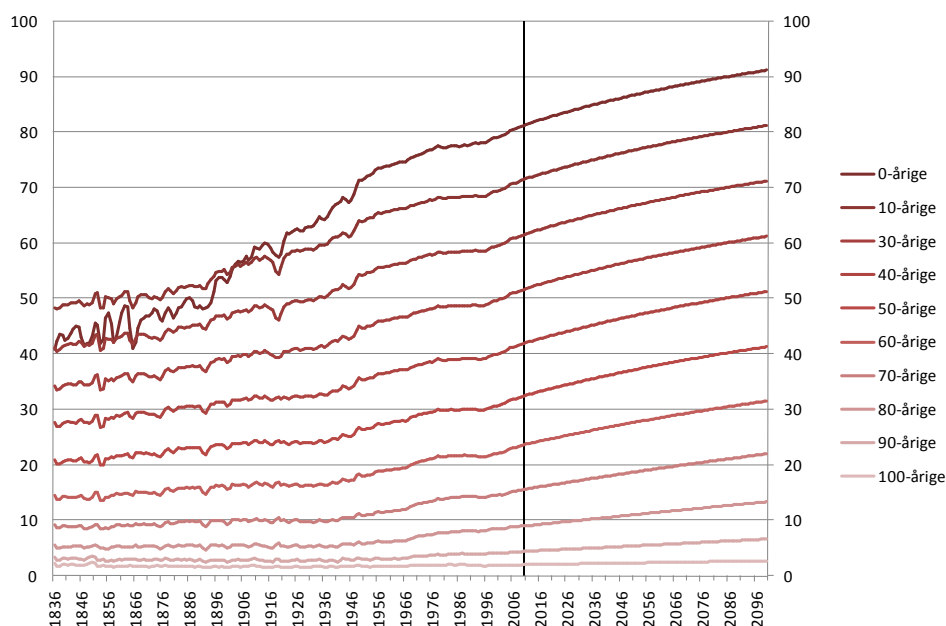
På basis af de internationale sammenligninger er der ligeledes belæg for, at restlevetidsudviklingen for mænd igennem de sidste 100 år kan opdeles i tre faser. Den første fase er karakteriseret ved høj og relativt ensartet vækstrate for

aldersgrupperne bortset fra et ekstraordinært stort fald i spædbørnsdødeligheden. Den næste fase er karakteriseret ved stort set 0-vækst for de fleste aldersgrupper, men dog lidt vækst i restlevetiden for ældre og et forsåt, men begrænset fald i spæd- og småbørnsdødeligheden. Den tredje fase er karakteriseret ved høj vækst for alle aldersgrupper og en tendens til, at vækstraten i restlevetiden er højest for de ældre aldersgrupper. Der henvises til DREAMs Befolkningsfremskrivning 2006 for en international sammenligning af restlevetid.

## 4.12 Restlevetidsudviklingen for kvinder

Svarende til beskrivelsen for mænd fremgår restlevetidsudviklingen for kvinder i aldersgrupperne 0, 10, ..., 90, 100 år fra 1836 og fremefter af Figur 11. Udviklingen i dødelighederne for kvinder kan opdeles i fem adskilte perioder, som ikke er helt så markante som i mændenes tilfælde. De fem perioder genfindes (på samme måde som hos mændene) i udviklingen i den aldersbetingede restlevetid for kvinder.

**Figur 11. Restlevetidsudviklingen for kvinder, udvalgte aldersgrupper**



Anm.: Restlevetider for årene 1836-1990 er beregnet ud fra dødeligheder fra HMD opgjort på A-gruppe format og medioficeret over tid. Restlevetid for år  $t$  og alder  $a$  er beregnet ud fra dødelighederne for alder  $a$  på tidspunkt  $t-1$  og  $t$ . Restlevetider for 1991 og frem er beregnet med afsæt i dødelighedsdata fra DST opgjort på B-gruppe format og medioficeret over både alder og tid. Restlevetid for år  $t$  og alder  $a$  er beregnet ud fra dødelighederne for alder  $a$  og  $a + 1$  på tidspunkt  $t-1$  og  $t$ . Dateringen anvendt i figuren henviser til sidste kalenderår anvendt til medioficeringen.

Den lodrette streg angiver sidste rene dataår, dvs. 2010 (2009:2010).

Kilde: Befolkningsfremskrivning 2011, DST og Human Mortality Database (HMD).

Den første periode indtil år 1900 er som for mænd karakteriseret ved en betydelig vækst i restlevetiden. Vækstraten er stort set den samme for alle aldersgrupper fra 10- til 40-årige, mens middellevetidens udviklingen igen skiller sig ud på grund af et betydeligt større fald i spædbørnsdødeligheden end i dødeligheden for de øvrige aldersgrupper. Således øges middellevetiden i gennemsnit med 2,3 måneder om året, mens restlevetiden for 60-årige øges med 0,2 måneder årligt.

Den anden periode varer fra 1900 - 1936 og er karakteriseret ved, at der er en meget lille vækst i restlevetiden for personer over 30 år. For de yngre aldersgrupper er der en positiv vækst, som er aftagende med alderen. Specielt spædbørnsdødeligheden er faldende og giver anledning til en gennemsnitlig vækst i middellevetiden på ca. 3,8 måneder årligt. Restlevetiden for 60-årige øges med gennemsnitligt 0,2 måneder om året gennem perioden.

Den tredje periode fra 1936 til 1978 er karakteriseret ved, at der er en betydelig vækst i restlevetiden for alle aldersgrupper, og at vækstraten er stigende med alderen indtil omkring 80 år. Undtaget er dog middellevetiden, hvor vækstraten er høj som følge af fortsat faldende spæd- og småbørnsdødelighed. I perioden fra 1936 til 1978 vokser middellevetiden for kvinder med godt 20 pct., mens restlevetiden for 10-årige piger vokser med 15 pct., restlevetiden for 30-årige kvinder med 18 pct., for 60-årige kvinder med 33 pct. og for 80-årige med 48 pct. Ændringen i middellevetiden svarer til en gennemsnitlig årlig stigning på 3,8 måneder om året, mens restlevetiden for 60-årige årligt stiger med 1,5 måneder i den tredje periode.

Fjerde periode løber fra 1978 til 1995, og som for mænd er denne en 0-vækstperiode. I løbet af denne periode er der således stort set ikke vækst i restlevetiden for kvinder i nogen aldersgrupper. Den gennemsnitlige årlige vækst i restlevetiden for de forskellige aldersgrupper varierer omkring et niveau på 0,2 måneder.

Femte og sidste periode er den nuværende, der foreløbig dækker perioden fra 1995 til 2010. I denne periode er der atter en betydelig vækst i restlevetiden for kvinder i alle aldersgrupper. Som for mænd er der en tendens til, at vækstraten i restlevetiden vokser med alderen i denne periode. Tendensen er dog mindre udtalt end hos mændene og gør sig ikke monotont gældende for aldersgruppen på 60 år og opefter, som har forholdsvis ensartede vækstrater. Middellevetiden for kvinder vokser 4,1 pct. eller gennemsnitligt 2,5 måneder årligt fra 1995 til 2010. Restlevetiden for 60-årige kvinder øges med 10,3 pct. i perioden, hvilket svarer til en gennemsnitlig årlig ændring på 1,8 måneder.

Internationale sammenligninger peger på, at opbremsningen i væksten i levetiden for kvinder i 1970'erne, 80'erne og begyndelsen af 1990'erne er et særskilt dansk fænomen, og at den seneste periodes udvikling tyder på, at middellevetidens udviklingen igen svarer til udviklingen i de omkringliggende lande.

## 4.13 Fremtidig udvikling i den aldersbetingede restlevetid

Som det fremgår af Figur 10 og Figur 11 giver den Lee-Carter fremskrevne reduktion i aldersbetingede dødeligheder over tid anledning til en stigning i den aldersfordelte restlevetid.

Stigningen i den aldersbetingede restlevetid er i fremskrivningen en forlængelse af de seneste års udvikling kombineret med en tendens til gradvist aftagende vækst. I det følgende fokuseres på udviklingen i restlevetiden for to udvalgte aldersgrupper: Middellevetiden, dvs. restlevetiden for en 0-årig, og restlevetiden for en 60-årig.

### 4.13.1 Fremtidig udvikling i middellevetiden<sup>25</sup>

I perioden fra 1995 til 2010 er middellevetiden for mænd vokset fra 72,9 år til 77,1 år, dvs. med 4,2 år. Det svarer til en gennemsnitlig årlig vækst på 3,3 måneder om året. Den samlede vækst over perioden 1995-2010 er af samme størrelsesorden som den samlede vækst fra 1950-1995, hvor den gennemsnitlige årlige vækst er knap 1,1 måned pr. år. Udgangspunktet for væksten i den mandlige middellevetid er derfor højt.

Fremskrivningen indebærer en gradvis opbremsning i denne vækst, hvor det forventes, at middellevetiden vokser med yderligere 2,2 år frem til 2020, hvor den vil være nået op på 79,3 år. I perioden fra 2020-2030 vil middellevetiden ifølge fremskrivningen vokse med yderligere næsten 2 år og således nå op på 81,2 år i 2030. Det svarer til en gennemsnitlig vækst i middellevetiden på 2,3 måneder om året. I den efterfølgende periode vokser den med små to måneder om året og når således 82,9 år i 2040 og 84,5 år i 2050.

Udviklingen i den årlige tilvækst i middellevetiden for mænd i den historiske periode fra 1950 til 1990 udviser årlig variation omkring en vækst på ca. 1 måned pr. år. Efter 1990 er væksten i middellevetiden steget og har ligget på ca. tre

---

<sup>25</sup> Angivelse af middellevetider i dette afsnit er for perioden forud for 1990 baseret på beregninger udført på medioficerede A-gruppe hyppigheder og efterfølgende på beregninger baseret på medioficerede B-gruppe dødeligheder. Danmarks Statistik har fra og med 2010 (dvs. dødelighedstavlens for 2008:2009) ændret metode og beregningsgrundlag for beregning af middellevetider. Disse baseres på nu oplysninger fra CPR-registeret og beregninger på C-grupper, jf. Danmark Statistik (2010b) og <http://www.dst.dk/Vejviser/Portal/Befolkning/Middellevetid.aspx>. I modsætning til DSTs tidligere dødelighedstavler er det nye format vanskeligt at approksimere ved medioficering. Den i afsnittet angivne middellevetid vil derfor for de seneste to historiske år 2008:2009 og 2009:2010 afvige fra DSTs opgørelse. Der er i nærværende publikation konsistens mellem opgørelse og metode knyttet til beregning af hhv. historisk (fra 1990) og fremskrevet middellevetid. DSTs metode kan i øvrigt ikke anvendes på fremskrevet data.

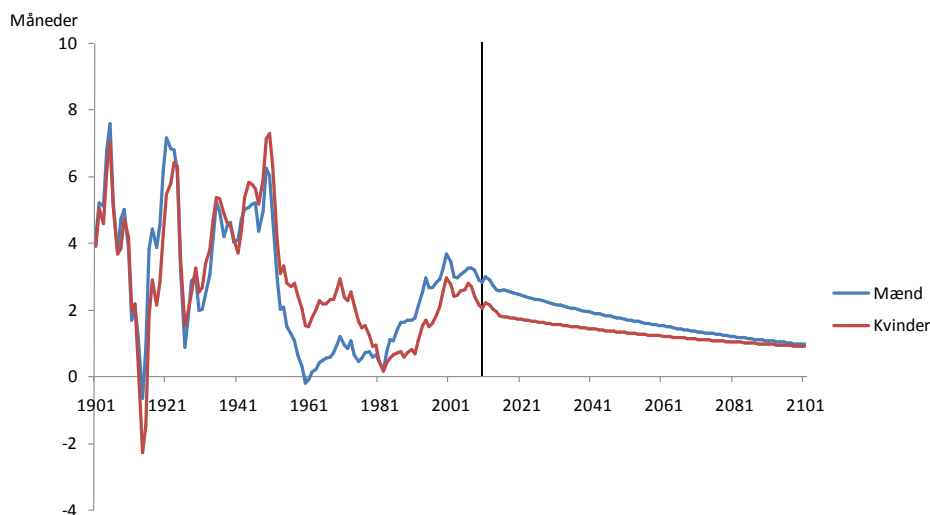
måneder pr. år i gennemsnit. I fremskrivningen antages, at den årlige vækst i middellevetiden gradvist reduceres gennem det 21. århundrede fra et niveau på omkring tre måneder i starten til et niveau på ca. 1,0 måned imod slutningen af århundredet, jf. Figur 12.

I perioden fra 1995 til 2010 er middellevetiden for kvinder steget med 3,2 år svarende til 2,6 måneder om året. Selvom det er væsentligt lavere end væksten i mændenes middellevetid i samme periode (3,3 måneder pr. år), er der også for kvinderne tale om en betydelig acceleration i væksten. Til sammenligning steg middellevetiden for kvinder med 3,4 år over perioden 1966 til 1995. Den gennemsnitlige årlige vækst er kun på 1,4 måneder i denne periode.

Fremskrivningen indebærer, at væksten i middellevetiden i de første år er lavere end i den seneste historiske periode og ligger på ca. 1,9 måneder pr. år indtil 2020. Væksten reduceres gradvist gennem fremskrivningsperioden og ligger på ca. 1 måned om året ved udgangen af århundredet, jf. Figur 12.

Middellevetiden vokser således fra 81,3 år i 2010 til 82,9 år i 2020. I 2050 er middellevetiden for kvinder vokset til 86,7 år. Det betyder dels, at væksten i middellevetiden for kvinder frem til 2050 er mindre end for mænd og dels, at væksten ligger betydeligt under den gennemsnitlige vækst i middellevetiden for kvinder i det 20. århundrede.

**Figur 12. Vækst i middellevetid. 10 års glidende gennemsnit**



Anm.: Væksten frem år 1985 er beregnet alene på dødelighedsdata fra HMD, mens vækstrater derefter er helt eller delvist baseret på data fra DST. Den lodrette streg angiver sidste dataår 2010.

Kilde: Befolkningsfremskrivning 2011, DST, Human Mortality Database (HMD) og egne beregninger.

#### 4.13.2 Fremtidig udvikling i restlevetiden for 60-årige

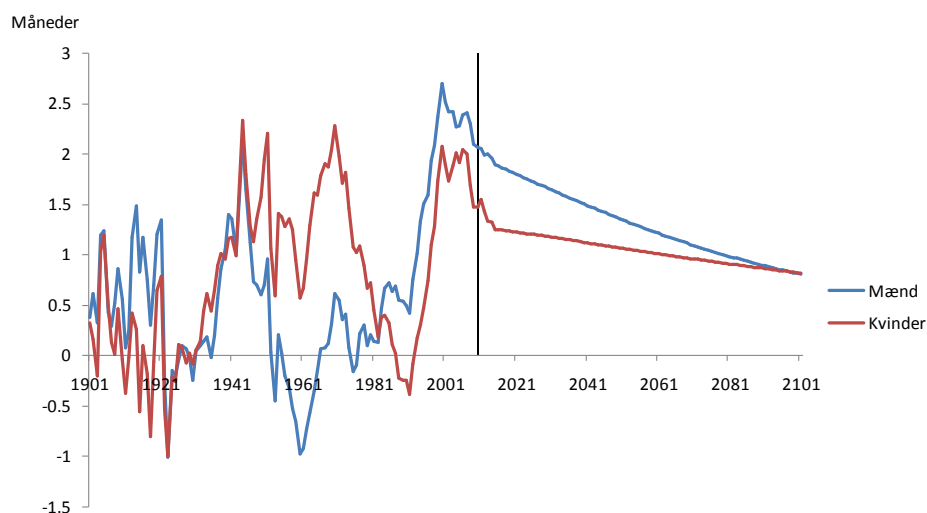
Restlevetiden for 60-årige mænd er vokset fra 17,7 år i 1995 til 20,6 år i 2010, hvilket svarer til en gennemsnitlig årlig vækst i perioden på 2,4 måneder pr. år. Relativt til den tidligere historiske udvikling er der tale om en bemærkelsesværdig høj vækst i denne periode. Til sammenligning kan det nævnes, at restlevetiden for en 60-årig mand også steg absolut med 2,8 år i perioden fra 1905 til 1995, svarende til gennemsnitligt 0,4 måneder årligt.

Resultatet af fremskrivningen er, at væksten i restlevetiden for 60-årige mænd falder tilbage mod niveauet før 1995, men at væksten også på langt sigt forbliver højere end i perioden før 1995. Fremskrivningen indebærer, at væksten i de første år er på omkring 1,9 måneder om året – hvilket er noget lavere end væksten siden 1995. Fra dette niveau falder den årlige vækst gradvist mod et langsigtet niveau på godt 0,8 måneder om året, jf. Figur 13.

Som for mænd er der sket en mærkbar stigning i væksten i restlevetiden for 60-årige kvinder i de seneste år. I denne periode er restlevetiden vokset med 2,2 år, hvilket svarer til en gennemsnitlig stigning på 1,8 måneder om året. Til sammenligning steg restlevetiden for 60-årige kvinder også med godt 2 år i perioden fra 1966 til 1995, hvilket svarer til en årlig vækst på lidt under 1 måned.

Fremskrivningen indebærer, at væksten på kort sigt bliver noget lavere end i den seneste periode og vil ligge på omkring 1,3 måneder om året, hvilket dog er markant højere end i perioden umiddelbart forud for 1995. Væksten er svagt faldende over tid og er på langt sigt godt 0,8 måneder om året, jf. Figur 13.

**Figur 13. Vækst i restlevetid for 60-årige. 10 års glidende gennemsnit**



Anm.: Væksten frem år 1985 er beregnet alene på dødelighedsdata fra HMD, mens vækstrater derefter er helt eller delvist baseret på data fra DST. Den lodrette streg angiver sidste dataår 2010.

Kilde: Befolkningsfremskrivning 2011, DST, Human Mortality Database (HMD) og egne beregninger.

#### 4.14 Sammenfatning af resultater om restlevetidstigning

Den hastigere reduktion i de aldersbetingede dødeligheder fra 1995 har betydet, at middellevetiden og specielt restlevetiden for 60-årige er vokset markant hurtigere gennem de seneste år end i de foregående 40 til 50 år. Udviklingen er særlig kraftig for mænd.

Tilsvarende stigninger i vækstraten for levetiden er indtrådt omkring 1980 i de øvrige nordiske lande og en række andre vestlige lande. I disse lande har de højere vækstrater været fastholdt i den mellemliggende periode. Vækstraten i Danmark er fra 1995 på niveau med de øvrige vesteuropæiske lande, jf. DREAMs Befolkningsfremskrivning 2006.

På denne baggrund er det valgt at opfatte ændringen i 1995 som et strukturelt brud og derfor basere estimationen af de aldersbetingede dødeligheder på en kort dataserie fra 1990. Det betyder, at fremskrivningen på det korte sigt får en vækstrate i restlevetiden, som ligger i underkanten af væksten i perioden 1995-2009, og som gradvist reduceres til et niveau, der svarer til den langsommere vækst i perioden fra 1950-1995.

2011-fremskrivningen afstedkommer en mindre stigning i den forventede vækst i den aldersbetingede restlevetid for mænd i forhold til sidste års fremskrivning, som ligeledes var baseret på en dataserie fra 1990. På langt sigt, vil mænds middellevetid være ca. 0,2 år højere end tidligere forventet, mens der for kvinders vedkommende er tale om en stigning på 0,4 år. Ændringen i middellevetiden genfindes kvantitativt i restlevetiden for 60-årige mænd, mens restlevetiden for 60-årige kvinder er øget med 0,3 år relativt til det langsigtede skøn i Befolkningsfremskrivning 2010.

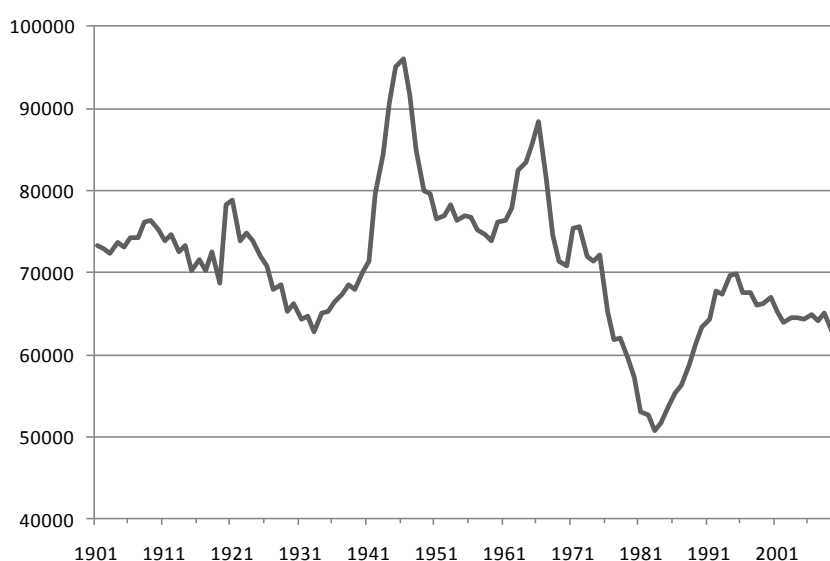
For yderligere bemærkninger til forskellen i middellevetidsudviklingen i de to fremskrivninger henvises til kapitel 8.

## 5 Fødsler og Fertilitet

### 5.1 Historisk udvikling i fødsler og fertilitet

Størrelsen af fødselsårgangene har varieret betydeligt gennem det 20. århundrede. Største årgang (1946) er med godt 96.000 fødsler omtrent dobbelt så stor som mindste årgang (1983) med knap 51.000 fødsler. De største årgange – på mere end 80.000 fødsler – findes i perioden fra 1943-50 og blandt disse generationers børn i 1963-67. De mindste årgange på under 60.000 findes i perioden 1979-88, jf. Figur 14.

**Figur 14. Antal fødte pr. år, 1901-2010**



Kilde: Statistikbanken, tabel HISB3 (1901-1989) og udtræk fra PSD (1990-2010)

Årsagen til variationen i størrelsen af fødselsårgangene er dels variation i antallet af kvinder i den fødedygtige alder og dels variation i den samlede fertilitet.

Igennem det 20. århundrede er den samlede fertilitet – dvs. det samlede antal børn pr. kvinde i den fødedygtige alder i et givet år – faldet fra godt 4 børn pr. kvinde til et niveau i år 2000 på 1,77 barn pr. kvinde. I år 2010 er den samlede fertilitet 1,88 barn pr. kvinde.

I perioden fra 1900 til 1930 skete der en meget markant reduktion i den samlede fertilitet fra godt 4 børn til godt 2 børn pr. kvinde. Fra starten af 1940'erne til 1946 steg fertiliteten til 3,0 barn pr. kvinde for herefter at falde til omkring 2,5 barn pr. kvinde i starten af 1950'erne. Dette niveau blev fastholdt frem til midten af 60'erne, hvor en ny periode med markante reduktioner i den samlede fertilitet indtrådte. Faldet fortsatte frem til 1983, hvor fertiliteten nåede ned på 1,38 barn pr. kvinde. Det er det hidtil laveste niveau for den samlede fertilitet i Danmark.

Herefter steg fertiliteten frem til 1995, hvor den nåede et niveau på 1,81 barn pr. kvinde. Herefter har fertiliteten svinget mellem 1,7 og 1,9 barn pr. kvinde, jf. Figur 15.

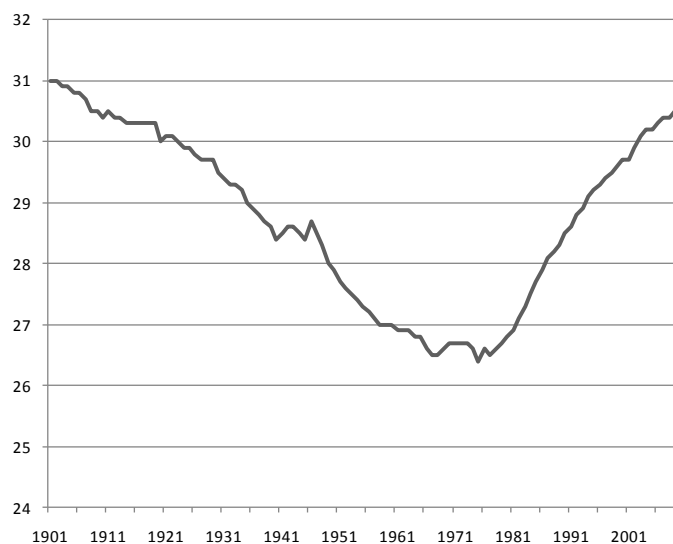
**Figur 15. Udviklingen i den samlede fertilitet i perioden 1901 – 2010**



Kilde: Statistikbanken, tabel FOD3.

De meget store årgange i 1940'erne hænger således i meget høj grad sammen med, at fertiliteten i denne periode var ekstraordinært høj.

**Figur 16. Gennemsnitsalder for fødende kvinder, 1901-2010**



Kilde: Statistikbanken, tabel FOD11.

De store årgange fra starten af 1960'erne er derimod alene en ekkovirkning af de store generationer fra 1940'erne, idet fertiliteten i denne periode var forholdsvis konstant. De meget små generationer i 1980'erne er en kombination af, at

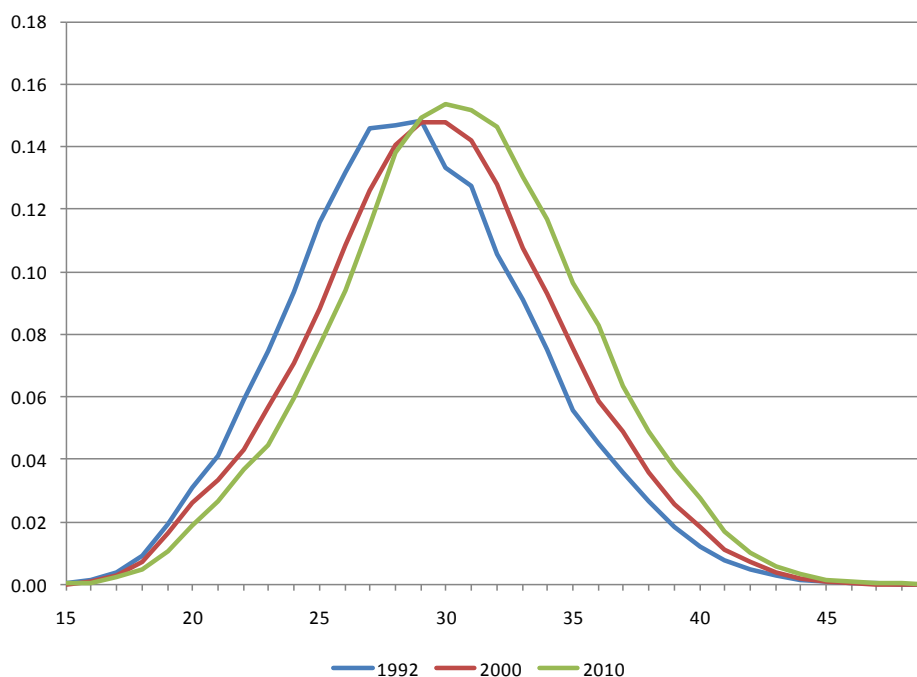
årgangene i 1950'erne ikke var store som i 1940'erne og at disse årgange også havde en lavere fertilitet.

Den relativt beskedne ændring i den samlede fertilitet i de seneste år dækker over en betydelig forskydning i den aldersfordelte fertilitet. Der er en klar tendens til, at moderens alder ved fødslen er voksende, jf. Figur 16.

Frem til 1975 var der tale om en gradvist faldende gennemsnitsalder for fødende kvinder. I de første 75 år af det 20. århundrede faldt gennemsnitsalderen med 4,6 år til 26,4 år. I den sidste fjerdedel af sidste århundrede har gennemsnitsalderen derimod været markant stigende og er i 2010 30,6 år, hvilket er næsten lige så højt som ved starten af det 20. århundrede.

Stigningen i gennemsnitsalderen for fødende kvinder igennem de seneste årtier er karakteriseret ved, at fordelingen for aldersbetinget fertilitet forskydes. Således har der været et betydeligt fald i den aldersbetingede fertilitet op til det 28. år, mens der har været en endnu større stigning i den aldersbetingede fertilitet for kvinder over 28 år. Bevægelsen er sket gradvist gennem perioden, jf. Figur 17.

**Figur 17. Aldersbetinget fertilitet i 1992, 2000 og 2010**



Kilde: Egne beregninger på PSD.

## 5.2 Fremskrivning af aldersbetingede fertilitetskvoienter

Fremskrivningen af fertilitet foretages ved at fremskrive den aldersbetingede fertilitet for hver oprindelsesgruppe. I modsætning til estimationen af dødeligheder estimeres den enkelte aldersgruppes fertilitet uafhængigt af den samlede fertilitet og vil derfor også være uafhængig af de øvrige aldersgruppes fertilitet. Udviklingen i den samlede fertilitet i fremskrivningen fremkommer som en sum af udviklingen i de enkelte aldersgruppes fertilitet. Der er historisk set større udsving i fertilitet end i dødelighed, hvorfor udviklingen på kort sigt bestemmes på baggrund af den nuværende trend i fertiliteten, mens fertiliteten på langt sigt konvergerer mod et eksogent ligevægtsniveau.

Det er udelukkende for de tre største oprindelsesgrupper, at der er et tilstrækkeligt antal observationer til at kunne estimere den aldersbetingede fertilitet med et-års alderstrin. De tre grupper er: Personer med dansk oprindelse, indvandrere fra ikke-vestlige lande uden dansk statsborgerskab og indvandrere fra vestlige lande uden dansk statsborgerskab. For disse grupper uddrages for hver alder mellem 15 og 49 år trenden af den historiske udvikling i fertiliteten og den aldersbetingede fertilitet fremskrives derefter ved hjælp af disse trends. Trenden i den aldersbetingede fertilitet udledes ved anvendelse af en metode benævnt *Cubic Spline Smoothing*. Fremgangsmåden beskrives i afsnit 5.3.

De resterende befolkningsgrupper er små og det er problematisk at fremskrive fertiliteten for disse grupper af to grunde. For det første medfører gruppernes størrelse en betydelig statistisk usikkerhed. For det andet er de kvindelige medlemmer af efterkommergrupperne hovedsageligt under 25 år, hvilket betyder, at det er særdeles vanskelig at skønne over den aldersbetingede fertilitet for den ældste halvdel af den fødedygtige alder på grund af manglende data. Fremskrivningen konstrueres derfor ved at sammenligne den aldersbetingede fertilitetsprofil for hver enkelt af de små befolkningsgrupper med den tilsvarende aldersbetingede fertilitetsprofil for de 3 store grupper. Ved hjælp af lineær regressionsanalyse på observationer for de seneste 5 år estimeres de små befolkningsgruppes aldersbetingede fertilitet som et vejet gennemsnit af de tre store gruppes aldersbetingede fertilitet. Herefter fremskrives den aldersbetingede fertilitet for de små befolkningsgrupper på grundlag af fremskrivningerne af de 3 store gruppes aldersbetingede fertilitet. I afsnit 5.4 uddybes metoden.

De fundne estimater for den aldersbetingede fertilitet er meget afhængige af de seneste observerede dataår og må således betragtes som det bedste bud på den kortsigtede udvikling i fertilitet. På kort sigt kan der dermed være stor forskel i udviklingen i fertilitet for befolkningsfremskrivninger med forskelligt basisår. Idet tilføjelsen af et enkelt dataår imidlertid ikke indeholder information, der gør os mere vidende om det mellemlange og lange sigt, er det ikke hensigtsmæssigt at denne variation tillades at afspejle sig i den langsigtede udvikling. For således at sikre, at langsigtet udviklingen ikke ændres nævneværdigt ved indførelsen af et nyt

dataår, bestemmes ved anvendelse af en *Richardskurve* konvergenen mod et eksogent fastlagt niveau. Dette behandles i afsnit 5.5.

Skift i oprindelsesdimensionen har nødvendiggjort en revision af det hidtil anvendte langsigtsniveau. De hertil knyttede overvejelser præsenteres i afsnit 5.6.

### 5.3 Estimation af den aldersbetingede fertilitet for de store grupper

Trenden i den aldersbetingede fertilitet udledes for hver af de tre store befolkningsgrupper ved anvendelse af *Cubic Spline Smoothing* (CSS), jf. Hyndman, King & Billah (2002). Metoden indebærer tilpasning af en speciel funktionsform til data, hvilket resulterer i en udglattet kontinuert form, der i sin simpleste version har konstant hældning i endepunkterne, hvor de andenafledte således er nul. Metoden indebærer, at et tredjegradspolynomium tilpasses mellem hvert par af datapunkter under antagelse om, at det respektive interval opdeles i delintervaller af samme længde. Mellem  $n$  datapunkter tilpasses således  $n - 1$  tredjegradspolynomier med individuelle koefficienter. Haves således datapunkterne  $(x_i, y_i)$  hvor  $i = 1, \dots, n$ , estimeres for hvert interval en funktion af formen

$$S_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \text{ for } x \in [x_i, x_{i+1}]$$

Samlet omtales disse individuelle funktionssektioner som *the spline*,  $S(x)$ . Tilpasningen til den polynomiale funktionsform sker under antagelser om kontinuitet mellem de individuelle funktioner og deres første- og andenafledte

$$\begin{aligned} S_{i-1}(x_i) &= S_i(x_i) \\ S'_{i-1}(x_i) &= S'_i(x_i) \\ S''_{i-1}(x_i) &= S''_i(x_i) \end{aligned}$$

Anvendes *Cubic Spline Interpolation* stilles yderligere det krav, at  $S(x)$  skal gennemløbe samtlige datapunkter. Således opnås ikke en udglattet form, men derimod en funktionsform, der fluktuerer kraftigt. I nærværende situation er man netop interesseret i en udglattet form, hvorfor der i tilpasningen slækkes på kravet om gennemløb af samtlige datapunkter. Kravet erstattes af en kombineret afvejning af kriterier indehavende, at funktionens afstand til data minimeres og at den samtidig udviser begrænset kurvatur. Afstanden til data,  $D$ , måles som summen af de kvadratiske afvigelse mellem  $y_i$  og  $S(x_i)$ , dvs.

$$D = \sum_{i=1}^n (S(x_i) - y_i)^2$$

$D$  er oplagt mindst, når  $S(x_i)$  gennemløber samtlige datapunkter. Dette indbefatter dog som nævnt voldsom kurvatur, hvilken måles som den integrerede værdi af den andenaflædte spline

$$\int |S''(x)|^2 dx$$

Dette udtryk minimeres, når den andenaflædte er nul, dvs. når  $S(x)$  er en ret linie. Dette ville formodentligt resultere i en høj værdi af afstandsmålet  $D$ , hvis minimum jo er forbundet med en høj grad af kurvatur. Kravet til tilpasningen af  $S(x)$  er derfor givet som en kombination af de to mål. Restriktionen af kravene om kontinuitet søges derfor følgende udtryk minimeret

$$W = D + \alpha \int |S''(x)|^2 dx$$

hvor parameteren  $\alpha$  angiver den vægt, der pålægges afstandskravet. Er  $\alpha$  lille vil minimering af  $W$  medføre, at  $S(x)$  ligger tæt på de faktiske data. Parameteren bestemmes her automatisk ved minimering af det såkaldte *Generalized Cross Validation* (GCV) kriterium<sup>26</sup>.

Efter estimation af  $S(x)$  noteres hældningen mellem de to sidste punkter i splinen samt værdien i endepunktet, dvs.  $S(x_n)$ , hvilket angiver henholdsvis retningen og udgangspunktet for fremskrivningen. Fremskrivningen fremkommer ved anvendelse af funktionen  $f(t) = \exp\left(\frac{at+b}{t+c}\right)$ .

Parameteren  $c$  angiver den såkaldte tilpasningshastighed for fremskrivningen, som her er fastsat til 5. Denne parameter er et udtryk for, hvor hurtigt udviklingen i fertiliteten stabiliseres. En værdi på 5 indikerer, at havde udviklingen i fertiliteten fortsat lineært fra splinens endepunkt med den her beregnede hældningskoefficient, ville fertiliteten stabiliseres efter fem perioder.

Generelt indebærer en lav værdi af  $c$ , at udviklingen stabiliseres relativt hurtigt, idet afstanden til det for dette  $c$  gældende stabile niveau er rimelig kort. Et større  $c$  betyder, at det stabile niveau afviger mere fra niveauet angivende udgangspunktet for fremskrivningen. Samtidig vil et større  $c$  afstedkomme en langsommere tilpasning til det stabile niveau.

Således indikerer størrelsen af  $c$  altså ikke hastigheden for tilpasning til et for alle værdier af  $c$  gældende langsigtetsniveau, men derimod hastigheden af tilpasning til

---

<sup>26</sup> Se eksempelvis Burrage, Williams, Erhel & Pohl (1994) for en teoretisk specificering af denne almindeligt anvendte fremgangsmåde til fastlæggelse af parameteren  $\alpha$ . Papiret tager udgangspunkt i Wahbas originalfremstilling fra 1979.

en individuel stabil tilstand, i fald udviklingen havde været lineær. At funktionen  $f(t)$  er sammensat af en eksponential- og potensfunktion betyder, at udviklingen i fertiliteten netop ikke vil være lineær, hvilket resulterer i en over tid langsommere tilpasning til det stabile niveau.

Koefficienterne  $a$  og  $b$  bestemmes således, at såvel hældningen som funktionsværdien i det sidste datapunkt, svarende til  $t = 0$ , er sammenfaldende med henholdsvis hældningen mellem splinens sidste to punkter og værdien i dennes endepunkt. Med andre ord bestemmes  $a$  og  $b$  således at følgende er opfyldt

$$\begin{aligned} f(0) &= S(0) \\ f'(0) &= S'(0) \end{aligned}$$

hvor  $t = x_n = 0$  svarer til det sidste historiske år i den anvendte dataserie. Dette medfører, at

$$\begin{aligned} a &= c \frac{S'(0)}{S(0)} + \ln(S(0)) \\ b &= c \ln(S(0)) \end{aligned}$$

Med fastlæggelse af disse koefficienter kan fremskrivningen af de aldersafhængige fertilitetskvotienter gennemføres ved simpel indsættelse i fremskrivningsfunktionen  $f(t)$ .

Udgangspunktet for estimationen er de mediobaserede alders- og oprindelsesfordelte fertilitetskvotienter for årene 1992-2010 beregnet iht. beskrivelsen i afsnit 2.2.4. For personer af dansk oprindelse samt indvandrere fra hhv. vestlige og ikke-vestlige lande uden dansk statsborgerskab fremskrives fertilitetskvotienterne for de 15-49-årige ved først at anvende *cubic spline smoothing* på hver enkel aldersgruppe og dernæst indsætte resultatet i funktionen  $f(t)$ . Med dette udgangspunkt kan udviklingen i aldersfordelt fertilitet bestemmes for de resterende befolkningsgrupper.

#### 5.4 Bestemmelse af fertilitetskvotienter for de øvrige befolkningsgrupper

Fremskrivning af fertilitetskvotienter for de øvrige befolkningsgrupper gennemføres grundet det tynde datamateriale ved en alternativ fremgangsmåde. Ved at sammenligne fertilitetsprofilen over alder på et givet tidspunkt for hver enkelt af de små befolkningsgrupper med den tilsvarende profil i de tre store

befolkningsgrupper, kan det konstateres hvorvidt der er sammenfald i udviklingen de seneste fem år.

Fertilitetskvoteinterne for hver af de små grupper regresseres på kvotienterne i de af de store befolkningsgrupper, hvor der er konstateret en ensartet udvikling. Til brug for denne regressionsanalyse pooles data for op til de seneste fem år sammen, således der for hver af de små grupper maksimalt fremkommer tre signifikante regressionskoefficienter, svarende til at samtlige af de store grupper har forklaringssevne. Regressionskoefficienterne anvendes sammen med fremskrivningen af fertilitetskvoteinterne for de relevante hovedgrupper til at konstruere fremskrivningen for de mindre befolkningsgrupper.

På grund af manglende observationer er det ikke samtlige aldersgrupper, der indgår i regressionen for visse af de små befolkningsgrupper. Således er estimationen for efterkommere fra vestlige lande uden dansk statsborgerskab blot baseret på de 15-35-årige, mens der i estimationen for grupperne efterkommere fra ikke-vestlige lande uden dansk statsborgerskab og efterkommere fra ikke-vestlige lande med dansk statsborgerskab blot indgår data for 15-33-årige.

Det findes ved analysen, at udviklingen i fertiliteten for befolkningsgrupperne indvandrere fra vestlige lande med dansk statsborgerskab, efterkommere fra vestlige lande uden dansk statsborgerskab samt efterkommere fra vestlige lande med dansk statsborgerskab forklares af udviklingen i fertiliteten for de to store befolkningsgrupper personer af dansk oprindelse og indvandrere fra vestlige lande uden dansk statsborgerskab.

Den seneste udvikling i fertiliteten for befolkningsgruppen indvandrere fra ikke-vestlige lande med dansk statsborgerskab er bedst forklaret af en kombination af udviklingen i fertiliteten for befolkningsgrupperne personer med dansk oprindelse og indvandrere fra ikke-vestlige lande uden dansk statsborgerskab.

Til fremskrivningen af fertiliteten for henholdsvis efterkommere fra ikke-vestlige lande uden dansk statsborgerskab og efterkommere fra ikke-vestlige lande med dansk statsborgerskab anvendes udviklingen i fertiliteten for indvandrere fra ikke-vestlige lande uden dansk statsborgerskab og personer med dansk oprindelse.

## 5.5 Konvergens af fertilitet på langt sigt

På kort sigt beregnes udviklingen i fertiliteten som nævnt ved at ekstrapolere de eksisterende trends fundet ved *Cubic Spline Smoothing*. Der er historisk set meget store udsving i fertiliteten, hvorfor tilføjelsen af et nyt dataår til estimationen kan betyde, at den ekstrapolerede fertilitet på lang sigt ændres markant relativt til sidste års fremskrivning. Disse ændringer repræsenterer ikke, at vi er blevet meget klogere med hensyn til udviklingen i fertiliteten på lang sigt

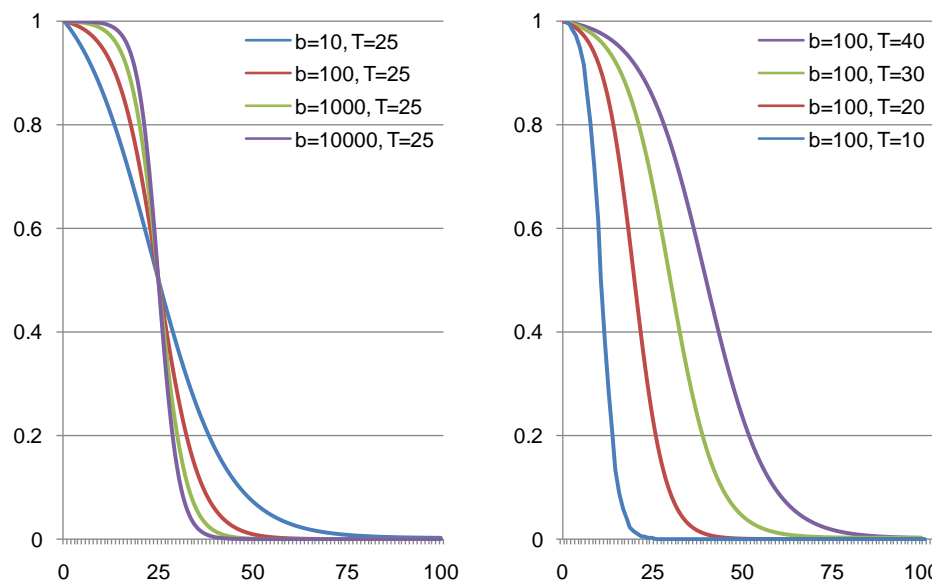
og udgør herudover en kilde til uforholdsmæssig stor variation i skønnet for den fremtidige befolkningsudvikling. For at undgå den markante variation i fertiliteten på lang sigt indføres derfor en funktion, der vægter fertiliteten mellem den kortsigtede ekstrapolerede udvikling og en eksogen fastlagt langsigtet udvikling, der alt andet lige ikke ændres mellem to fremskrivninger.

Funktionen, der benyttes til at vægte mellem det korte og lange sigt benævnes en *Richardskurve*, og er defineret ved

$$r_t = \left( 1 - \frac{1}{1 + b \left( \frac{1}{2+b} \right)^{\frac{t}{T}}} \right) \frac{1+b}{b}$$

Funktionen har den egenskab, at den antager værdien 1 for  $t = 0$ , konvergerer mod 0 for  $t$  gående mod  $\infty$  og antager værdien  $\frac{1}{2}$  for  $t = T$ . Således er halveringstiden givet ved parameteren  $T$ , mens parameteren  $b$  er udtryk for, om konvergensen sker gradvist. På figuren ses funktionen for forskellige værdier af  $b$  og  $T$ .

**Figur 18. – Richardskurve ved forskellige værdier af  $b$  og  $T$ .**



Den endelige udvikling i fertiliteten,  $\phi_t^{x,o}$ , bestemmes således ved en vægtning af kort- og langsigtudviklingen,  $\phi_t^{SR,x,o}$  og  $\phi_t^{LR,x,o}$ ,

$$\phi_t^{x,o} = r_t \phi_t^{SR,x,o} + (1 - r_t) \phi_t^{LR,x,o}$$

hvor koefficienterne i Richardskurven er sat til  $b=100$  og  $T=10$ .

## 5.6 Fastlæggelse af det langsigtede fertilitetsniveau

Når der bortses fra statsborgerskabsstatus anvendes til 2011-fremskrivningen samme langsigtede fertilitetsskøn som til 2010-fremskrivningen. For personer med dansk oprindelse samt efterkommere fra ikke-vestlige lande anvendes uafhængigt af statsborgerskab et langsigtsniveau på 1,9. For efterkommere fra vestlige lande er det langsigtede niveau uagtet statsborgerskab 1,75. Indvandrere fra ikke-vestlige lande havde i 2010-fremskrivningen en langsigtet fertilitet på 1,9, mens indvandrere fra vestlige lande har en fertilitet i 2040 på 1,75. Historisk noteres for indvandrere en betragtelig variation i fertiliteten på tværs af statsborgerskab, hvilken ønskes bevaret i fremskrivningen. For indvandrergrupperne bestemmes derfor statsborgerskabsspecifikke langsigtede fertilitetsrater.

Langsigtede fertilitetsniveauer for indvandrergrupperne fordelt på statsborgerskabsstatus er bestemt ud fra den på nuværende tidspunkt gældende relation mellem den samlede fertilitet for grupperne som helhed og den samlede fertilitet for den enkelte gruppe fordelt på statsborgerskab. Hvis det eksempelvis findes, at fertiliteten for indvandrere fra ikke-vestlige lande uden dansk statsborgerskab historisk er 1,5 pct. højere end fertiliteten for indvandrere fra ikke-vestlige lande som helhed, er det antaget, at denne relation også er bevaret på sigt<sup>27</sup>. Hvis det langsigtede skøn for gruppen som helhed således er fastsat til 1,9, vil langsigtsniveauet for gruppen uden dansk statsborgerskab være 1,5 pct. højere, dvs. ca. 1,93. Eftersom indvandrerbefolkningens sammensætning på statsborgerskab ikke ændres væsentligt i fremskrivningen, vil den samlede fertilitet for gruppen som helhed på sigt være ca. 1,9 selv om der er anvendt særskilte niveauer for de enkelte statsborgerskabskategorier<sup>28</sup>. Det skal bemærkes, at de langsigtede niveauer for efterkommere ikke varierer på tværs af statsborgerskab. Dette skyldes, at efterkommeres sammensætning på statsborgerskab ændres markant over tid, hvorfor de ønskede gennemsnitsniveauer ikke ville kunne opnås ved varierende langsigtsniveauer.

Fastlæggelsen af det langsigtede fertilitetsniveau kunne tænkes forbedret på flere punkter. Flere studier viser, at beslutningen om at få børn afhænger af en lang række faktorer. Fleksibilitet af arbejdsmarkedet, den enkeltes højst opnåede uddannelsesniveau samt indkomst er blandt de faktorer, der kan influere på fertiliteten, jf. D'Addio & d'Ercole (2005). Fertilitetsfremskrivning indbefatter på nuværende tidspunkt ikke sammenligning af nuværende kohorter med kohorter, der har gennemlevet de fertile aldre. Suppleret med paritetsfordelt fertilitet kunne dette muligvis bidrage til en forbedring af det langsigtede skøn, men man

<sup>27</sup> Det relative forhold beregnes som et gennemsnit over de sidste tre år, senest 2008-2010.

<sup>28</sup> Variation i fertilitetsniveauerne på tværs af statsborgerskabsstatus er stort set kvantitativt ækvivalent med hvad benyttedes i 2010-fremskrivningen. Det langsigtede fertilitetsniveau for indvandrere fra ikke-vestlige lande var 1,93 for personer uden dansk statsborgerskab og 1,87 for personer med dansk statsborgerskab i både 2010- og 2011-fremskrivningen. For vestlige indvandrere uden dansk statsborgerskab var det langsigtede niveau 1,77 i begge fremskrivninger, mens niveauet for vestlige indvandrere med dansk statsborgerskab i 2011-fremskrivningen er 1,63 mod 1,61 i sidste års fremskrivning.

kan naturligvis ikke ad denne vej tage højde for modefænomener og konjunkturtendenser.

## 5.7 Fremskrivningsresultater

Hovedresultatet af fremskrivningen er, at den samlede fertilitet for befolkningen er voksende de første 10-15 år af fremskrivningsperioden for derefter at stabilisere sig på et niveau omkring 1,88, jf. Tabel 2. For indvandrere fra ikke-vestlige lande uden dansk statsborgerskab er der tale om en initialt faldende tendens med efterfølgende stabilisering på langt sigt. Fertiliteten for indvandrere fra vestlige lande er stigende gennem hele fremskrivningen. For efterkommere fra ikke-vestlige lande med dansk statsborgerskab noteres en svagt faldende fertilitet, mens en markant stigning kan bemærkes for samme befolkningsgruppe uden dansk statsborgerskab. Udviklingen i sidstnævnte afspejler pr. konstruktion udviklingen i fertiliteten for indvandrere fra ikke-vestlige lande uden dansk statsborgerskab. Trods et lavere udgangspunkt end det sidste dataår giver anledning til, stiger fertiliteten for efterkommere fra vestlige lande med dansk statsborgerskab gennem hele fremskrivningen, således at niveauet på sigt svarer til, hvad var gældende i år 2010. For tilsvarende gruppe uden dansk statsborgerskab observeres ligeledes en jævn stigende tendens. Nedenstående tabel tydeliggør desuden hvorledes den langsigtede fertilitet tillades at variere på tværs af statsborgerskabsstatus iht. fremgangsmåden beskrevet i afsnit 5.6.

Generelt indebærer tilgangen, at fertiliteten er stort set uændret fra 2050 og fremefter.

**Tabel 2. Samlet fertilitet for befolkningen som helhed og fordelt på oprindelse**

	2010	2050	2100
Personer med dansk oprindelse	1.90	1.90	1.90
Indvandrere fra ikke-vestlige lande uden dansk statsborgerskab	1.99	1.93	1.93
Indvandrere fra ikke-vestlige lande med dansk statsborgerskab	1.86	1.87	1.87
Indvandrere fra vestlige lande uden dansk statsborgerskab	1.55	1.77	1.77
Indvandrere fra vestlige lande med dansk statsborgerskab	1.52	1.63	1.63
Efterkommere fra ikke-vestlige lande uden dansk statsborgerskab	1.68	1.90	1.90
Efterkommere fra ikke-vestlige lande med dansk statsborgerskab	1.89	1.90	1.90
Efterkommere fra vestlige lande uden dansk statsborgerskab	1.88	1.75	1.75
Efterkommere fra vestlige lande med dansk statsborgerskab	1.75	1.75	1.75
I alt	1.87	1.88	1.88

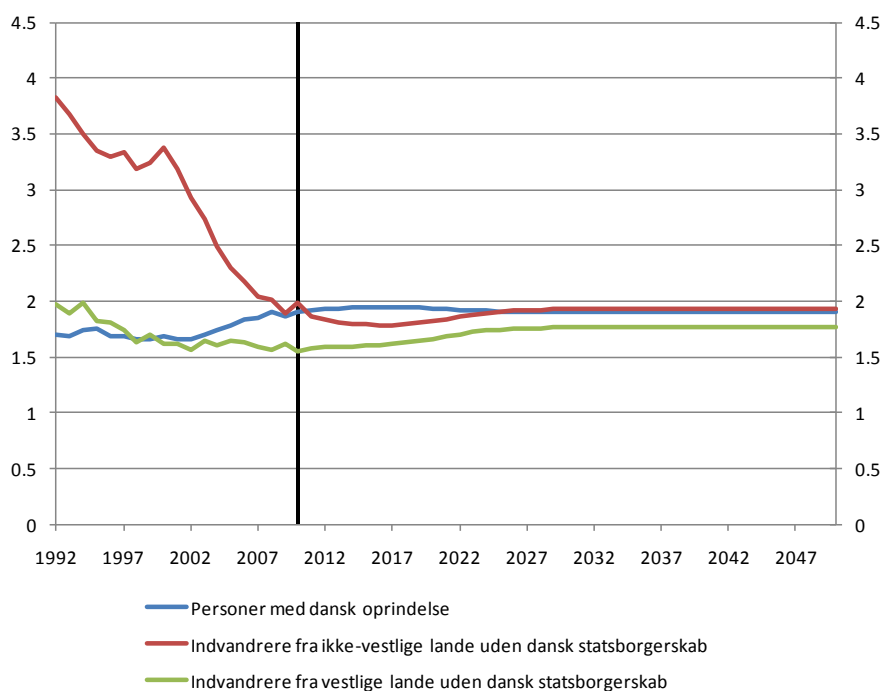
Kilde: Befolkningsfremskrivning 2011

Indvandrere fra ikke-vestlige lande uden dansk statsborgerskab havde i 1992 en samlet fertilitet på omkring 3,8 (børn pr. kvinde). Frem til omkring år 2000 lå den omkring 3,3 barn pr. kvinde, men i de seneste år har der været en dramatisk reduktion i fertiliteten for denne gruppe, således at den i 2010 er 1,99 barn pr. kvinde, hvilket dog er marginalt højere end niveauet for 2009. Der er næppe tvivl om, at såvel de ændrede regler for familiesammenføring som den ændrede sammensætning af indvandringen efter lovændringen i år 2002, har betydning

for denne udvikling. Det er imidlertid vanskeligt at skønne over, hvor stor den langsigtede effekt vil være. Estimation baseret på den historiske udvikling indebærer, at der forventes en fortsat svagt faldende tendens i den indledende del af fremskrivningen, mens tilpasning til langsigtsniveauet sikrer en stabilisering på et højere niveau, jf. Figur 19.

For personer med dansk oprindelse har der i de seneste år været en tendens til en mindre stigning i den samlede fertilitet. Det forventes i fremskrivningen, at dette fortsætter således at den samlede fertilitet på kort sigt vil stige til 1,94 barn pr. kvinde, for derefter at konvergere mod et langsigtet niveau på omkring 1,9 barn pr. kvinde. Omvendt har den samlede fertilitet for indvandrergruppen fra vestlige lande uden dansk statsborgerskab vist en tendens til reduktion fra et niveau på omkring 2,0 børn pr. kvinde til ca. 1,55 barn pr. kvinde. I fremskrivningen forventes initialt en stabilisering på det nuværende niveau, men tilpasningen til langsigtsudviklingen afstedkommer en gradvis stigning på sigt til en fertilitet på 1,77 i 2050, jf. Figur 19 og Tabel 2.

**Figur 19. Udviklingen i den samlede fertilitet for de tre største befolkningsgrupper 1992-2050**



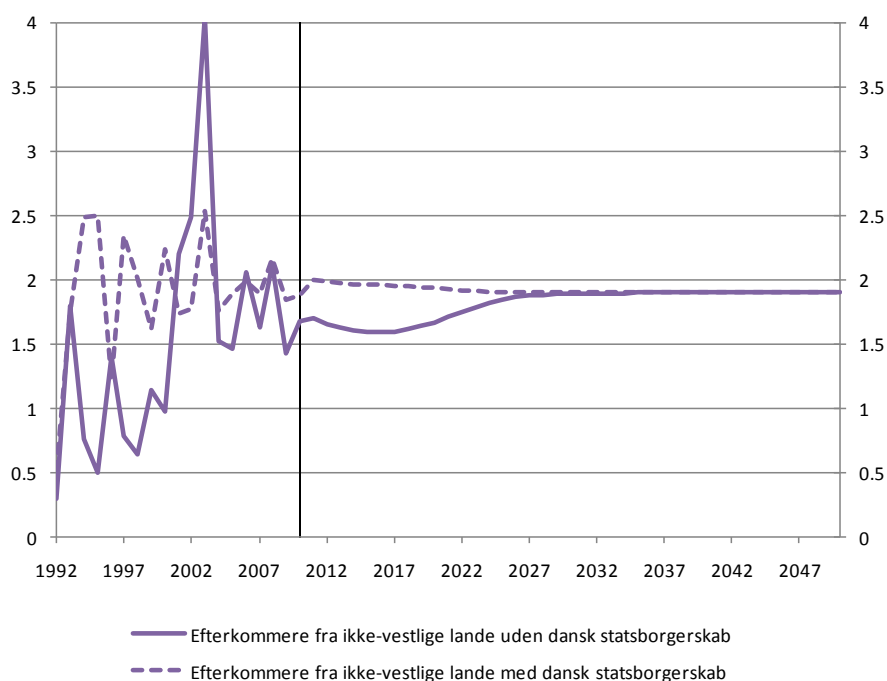
Kilde: Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste dataår 2010.

Alle fire efterkommergrupper har en lav og meget fluktuerende fertilitet. Det sidste skyldes befolkningsgruppernes begrænsede størrelse, som betyder, at tilfældige udsving påvirker det samlede niveau. Det forudsættes i fremskrivningen, at aldersprofilen for fertiliteten kan beskrives ud fra to af de tre store befolkningsgrupper, dvs. personer af dansk oprindelse, indvandrere fra ikke-vestlige lande uden dansk statsborgerskab og indvandrere fra vestlige lande uden dansk statsborgerskab. Dette medfører, at udviklingen i de store grupper på

kort sigt afspejles i fertiliteten for efterkommergrupperne. Generelt opnås dog relativt hurtigt tilpasning til de eksogent valgte langsigtssniveauer, jf. Figur 20 og Figur 21.

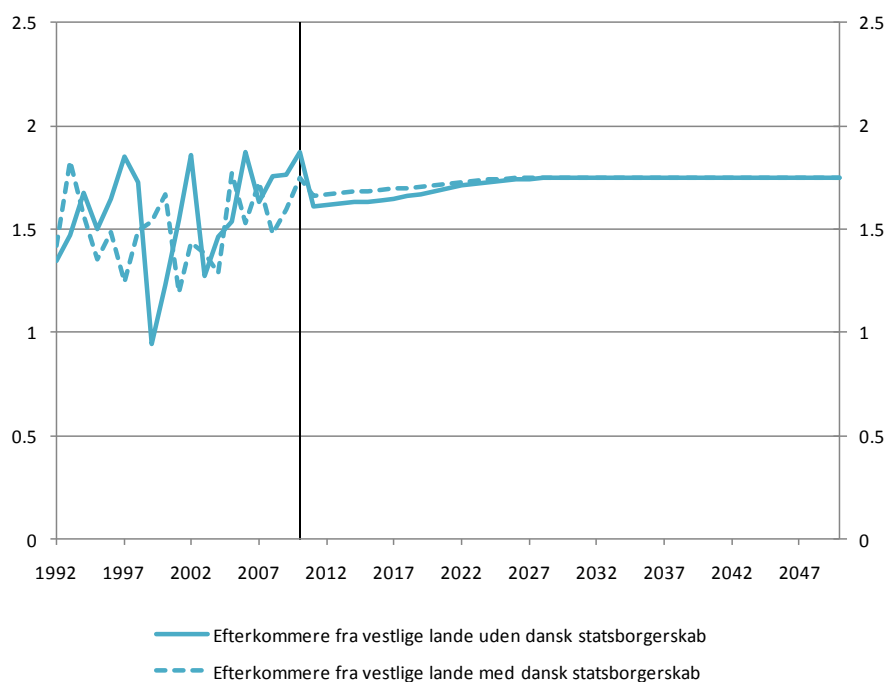
Det bemærkes, at der er en meget betydelig usikkerhed knyttet til fertilitetsudviklingen for efterkommergrupperne. Udover størrelsen af befolkningsgrupperne skyldes dette, at de kvindelige efterkommere er unge og således ikke har gennemlevet de fertile aldre, hvilket giver anledning et tyndt datamateriale. Der er således en risiko for, at den anvendte procedure indebærer, at aldersfordelingen af fertiliteten får for lidt vægt i de ældre aldersgrupper. Endelig er fertiliteten for gruppen af efterkommere fra mindre udviklede lande i betydeligt omfang påvirket af både 24-årsreglen og mest-tilknytningsprincippet i indvandringslovgivningen. Som tidligere nævnt er det endnu for tidligt at skønne over disse bestemmelser langsigtede konsekvenser.

**Figur 20. Udviklingen i den samlede fertilitet for efterkommere fra ikke-vestlige lande 1992-2050**



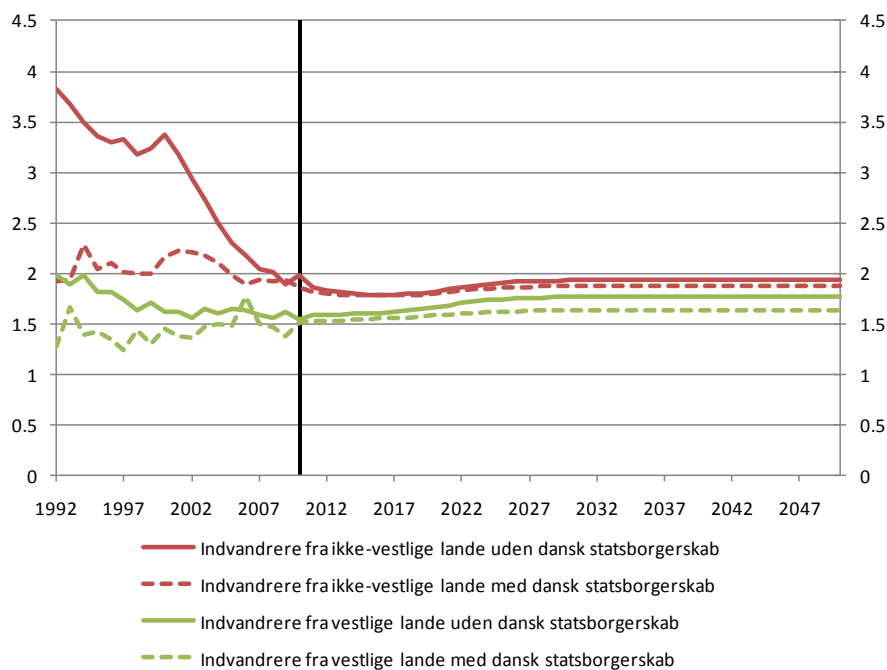
Kilde: Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste dataår 2010.

**Figur 21. Udviklingen i den samlede fertilitet for efterkommere fra vestlige lande 1992-2050**



Kilde: Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste dataår 2010.

**Figur 22. Udviklingen i den samlede fertilitet for indvandrergupperne 1992-2050**



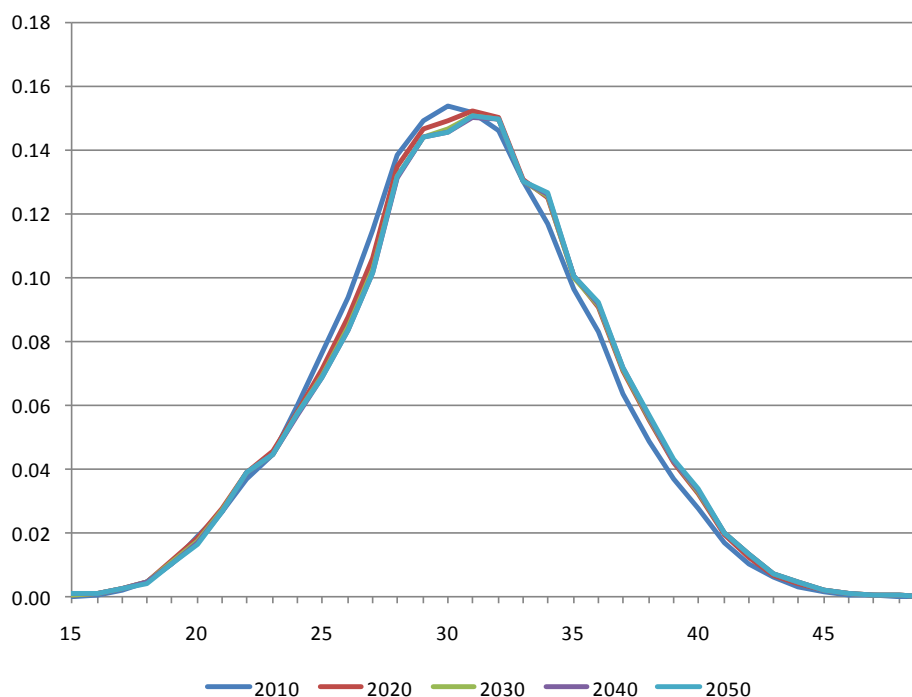
Kilde: Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste dataår 2010.

Som nævnt har der historisk set været et meget markant fald i fertiliteten for indvandrere fra ikke-vestlige lande uden dansk statsborgerskab. En tilsvarende, men mindre markant udvikling ses for indvandrere fra ikke-vestlige lande med dansk statsborgerskab, jf. Figur 22. For denne gruppe er fertiliteten faldet med knapt 0,4 barn pr. kvinde siden 2002. Forløbet har dog ikke været monotont, idet der har været tale om en stabilisering siden 2006. Det forventes, at faldet udfases yderligere i fremskrivningen således, at fertiliteten for denne befolkningsgruppe inden for en kortere årrække stabiliseres på et niveau omkring 1,87 barn pr. kvinde. For indvandrere fra vestlige lande antages i fremskrivningen, at fertiliteten for personer uden dansk statsborgerskab over tid stabiliseres på et niveau, der er omkring 0,2 højere end det nuværende. For personer med dansk statsborgerskab er niveauændringen ca. 0,1 over tid, jf. Figur 22.

## 5.8 Udviklingen i den aldersbetingede fertilitet

Den aldersbetingede fertilitetsprofil er blevet forskudt i retning mod højere alder for fødende kvinder igennem de seneste år, jf. Figur 17. I fremskrivningen forudsættes denne bevægelse at fortsætte med væsentlig reduceret styrke og det fås, at aldersprofilen konvergerer mod en form, som med få undtagelser ikke afviger markant fra profilen i 2010<sup>29</sup>, jf. Figur 23.

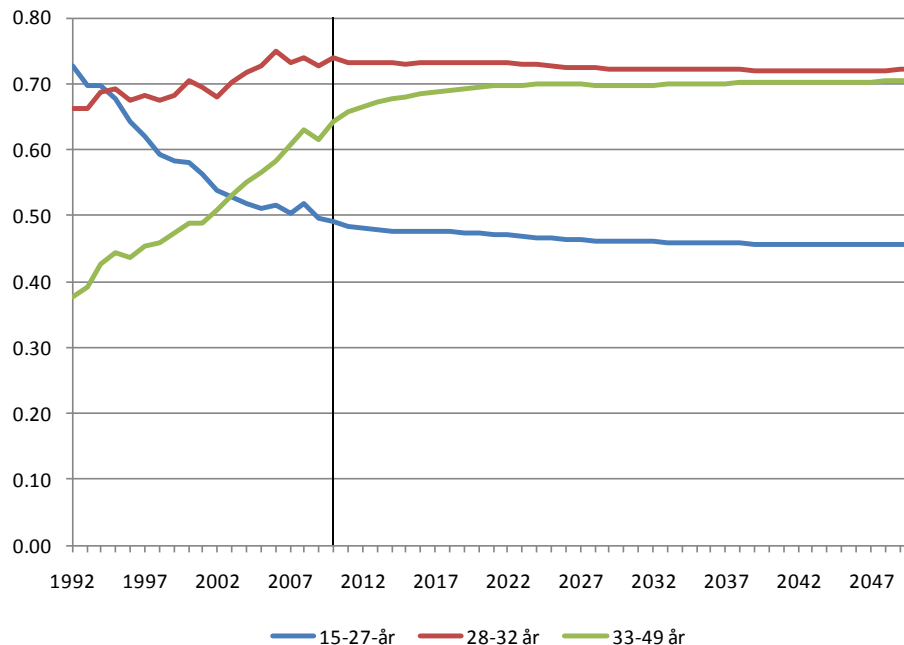
**Figur 23. Aldersbetinget fertilitet for hele befolkningen, udvalgte år**



<sup>29</sup> Mindre afvigelser fra formen i 2010 specielt omkring 30-35 års alderen, skal tilskrives, at den anvendte metode til fremskrivning som nævnt indebærer, at den historiske tendens i hver aldersgruppe ekstrapoleres uafhængigt af tendenserne i de øvrige. At trenden for de ca. 30-årige er relativt beskeden kan ses i Figur 17.

Kilde: Befolkningsfremskrivning 2010

**Figur 24. Udviklingen i fertiliteten i aldersgrupperne 15-27, 28-32 og 33-49 år**



Kilde: Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste dataår 2010.

Den varierende udvikling på tværs af alder er ligeledes illustreret i Figur 24, hvor den samlede fertilitet er opdelt i aldersgrupperne 15-27 år, 28-32 år og 33-49 år. Opdelingen af den samlede fertilitet på aldersintervaller afslører, at den samlede fertilitet for 15-27-årige er faldet stødt gennem hele den historiske periode, bortset fra de sidste år, hvor den synes at stabiliseres. Samtidig er den samlede fertilitet for både de 28-32-årige og de 33-49-årige vokset i den historiske periode. For den yngste aldersgruppe er der dog tale om en mindre markant stigning relativt til den ældste gruppe.

På lang sigt ses det, at den samlede fertilitet for de 28-32-årige er stort set uændret relativt til det nuværende niveau, mens fertiliteten for de 33-49-årige og de 15-27-årige er hhv. højere og lavere end det nuværende niveau. Den fremtidige stigning i den samlede fertilitet, jf. Tabel 2, dækker således over en aldersforskydning i fertilitetsmønstret.

## 5.9 Nyfødtes fordeling på oprindelsesgrupper

De nyfødte har ikke nødvendigvis samme oprindelse som deres moder, jf. Tabel 1. Derfor fordeles de nyfødte på oprindelsesgrupper betinget af moderens

oprindelsesgruppe. Dette gøres med udgangspunkt i data for de tre seneste år, dvs. for årene 2008-2010 i nærværende fremskrivning. Denne forudsætning indebærer at mønstret for, hvordan de forskellige befolkningsgrupper blandes i forbindelse med fødsler fastholdes i fremtiden. Der er således ikke øget integration – forstået som øget tendens til at befolkningsgrupperne blandes i forbindelse med fødsler – i fremskrivningsperioden.

Der er i dag kun en beskedent tilbøjelighed til, at indvandrere fra ikke-vestlige lande får børn med personer af dansk oprindelse. Det er kun 18 til 28 pct. af fødslerne, hvor moderen er indvandrer fra et ikke-vestligt land, og hvor faderen er enten af dansk oprindelse eller efterkommer med dansk statsborgerskab, jf. Tabel 3.

**Tabel 3. Fordelingen af barnets oprindelse givet moderens oprindelse, gennemsnit 2008-2010**

<i>Barn</i>	<i>ixn</i>	<i>ixd</i>	<i>iwn</i>	<i>iwd</i>	<i>dxn</i>	<i>dxd</i>	<i>dwn</i>	<i>dwd</i>	<i>da</i>	<i>Række- sum</i>
<i>Moder</i>										
<i>ixn</i>	0	0	0	0	0,4825	0,2414	0	0	0,2761	1
<i>ixd</i>	0	0	0	0	0,0006	0,8169	0	0	0,1826	1
<i>iwn</i>	0	0	0	0	0	0	0,4300	0,0431	0,5269	1
<i>iwd</i>	0	0	0	0	0	0	0	0,2435	0,7565	1
<i>dxn</i>	0	0	0	0	0,5034	0,1586	0	0	0,3380	1
<i>dxd</i>	0	0	0	0	0	0	0	0	1	1
<i>dwn</i>	0	0	0	0	0	0	0,3567	0,0194	0,6239	1
<i>dwd</i>	0	0	0	0	0	0	0	0	1	1
<i>da</i>	0	0	0	0	0	0	0	0	1	1

Hvis moderen er indvandrer fra et vestligt land, er der derimod mellem 53 og 76 pct. sandsynlighed for, at faderen er enten efterkommer med dansk statsborgerskab eller af dansk oprindelse. Dette svarer nogenlunde til niveauet anvendt i sidste års fremskrivning.

Tilsvarende er der meget stor forskel for efterkommergrupperne. Hvis moderen er efterkommer fra et ikke-vestligt land uden dansk statsborgerskab, er der 50 pct. sandsynlighed for, at faderen enten er indvandrer eller efterkommer uden dansk statsborgerskab. Hvis moderen derimod er efterkommer fra et vestligt land uden dansk statsborgerskab, er der 36 pct. sandsynlighed for, at faderen er indvandrer eller efterkommer uden dansk statsborgerskab. De relativt store udsving i sandsynlighederne for efterkommerne såvel indbyrdes mellem grupperne og fra år til år skal i stor udstrækning tilskrives, at der er tale om små befolkningsgrupper.

## 6 Indvandring, udvandring og statsborgerskab

Nærværende kapitel behandler særskilt de gjorte antagelser med henblik på fastlæggelse af den fremtidige udvikling i hhv. indvandring, udvandring og statsborgerskabsskifte. Udviklingen i visse typer af indvandring er særdeles afhængig af den gældende lovgivning på området, hvorfor der nedenfor gives et overblik over de tiltag, der historisk har haft en markant betydning. De øvrige afsnit tjener som en uddybning af de i afsnit 1.2.5-1.2.7 nævnte antagelser.

### 6.1 Lovgivning med betydning for indvandringen<sup>30</sup>

Som beskrevet i afsnit 1.2.5 antages indvandringen fra vestlige og ikke-vestlige lande uden dansk statsborgerskab at blive fastlagt eksogent i befolkningsfremskrivningsmodellen. Indvandringen til disse to befolkningsgrupper er reguleret og kræver opholdstilladelse. Således er ændringer i lovgivningen særdeles afgørende for den fremtidige udvikling i indvandringen baseret på visse opholdsgrundlag. Der gives opholdstilladelser til følgende forskellige grupper: Flygtninge m.v., familiesammenførte, EU/EØS borgere, personer med beskæftigelse i Danmark og personer med andre grunde til ophold i Danmark. Omfanget og sammensætningen af indvandringen reguleres ved udlændingeloven.

Den oprindelige udlændingelov stammer fra 1952. Ifølge denne kunne udlændinge relativt frit indvandre til Danmark. I sammenhæng med den stigende arbejdsløshed i 1973 indførtes et indvandringsstop, hvilket betød, at udlændinge som hovedregel kun kunne opnå opholdstilladelse af beskæftigelseshensyn i særlige tilfælde. Indvandringstoppet betød ikke, at der ikke var indvandring, idet familiesammenføring og tildeling af asyl fortsat gav anledning til indvandring. Indvandringsstoppet omfattede ikke nordiske statsborgere og statsborgere i EU-landene, der kunne indrejse og arbejde i Danmark.

I 1983 blev udlændingeloven revideret. Den ny lovgivning havde blandt andet til formål at styrke retsstillingen for familiesammenførte og flygtninge. Der blev indført retskrav på familiesammenføring af børn, ægtefæller og forældre. Herudover indførtes begrebet "de facto flygtninge" i lovgivningen. Der skulle være tale om personer, der ikke umiddelbart var flygtninge efter FN's Flygtningekonventionen, men hvor det af lignende eller andre tungtvejende grunde, der medførte begrundet frygt for forfølgelse eller tilsvarende overgreb, ikke burde kræves, at vedkommende vendte tilbage til sit hjemland.

I 2002 gennemførtes endnu en større lovændring af udlændingeloven (lov nr. 365 af 6. juni 2002). Formålet med denne ændring var at begrænse adgangen til familiesammenføring og at indsnævre betingelserne for adgang til asyl.

---

<sup>30</sup> Nærværende afsnit er i vidt omfang en gengivelse af s. 85-88 i Velfærdskommissionen (2005).

Lovændringen i 2002 afskaffede muligheden for at få asyl med status som de facto flygtning. I stedet indførtes en mulighed for asyl til personer med beskyttelsesstatus. Det er personer, som ikke umiddelbart opfylder definitionen i FN's Flygtningekonvention, men som Danmark har forpligtiget sig til at modtage gennem andre konventioner, herunder Den Europæiske Menneskerettighedskonvention fra 1950 samt 6. tillægsprotokol fra 1983. Der gives opholdstilladelse til flygtninge, der risikerer dødsstraf, tortur, umenneskelig eller nedværdigende behandling eller straf, hvis de vender tilbage til deres hjemland.

Lovændringen i 2002 betød også en række regelændringer vedrørende familiesammenføring af ægtefæller. Der indførtes bl.a. en 24 års regel, der indebærer, at begge parter skal være 24 år eller derover. Indtil lovændringen var aldersgrænsen 18 år. Yderligere skete der en stramning i reglen om, at familiesammenføring med ægtefælle sker til det land, hvor parrets samlede tilknytning er størst ("tilknytningskravet"). Efter lovændringen gælder dette for alle personer, mens det før lovændringen ikke var gældende for danske statsborgere. Kravet om, at tilknytningen til Danmark skal være størst bortfalder, hvis den ægtefælle, der bor i Danmark, har haft dansk statsborgerskab i over 28 år. Det samme gælder for personer, som har opholdt sig lovligt i Danmark i over 28 år, og som er født og opvokset i Danmark eller kommet hertil som mindre børn. Familiesammenføring af ægtefælle kan opnås af personer med dansk eller nordisk statsborgerskab, flygtninge med ophold i Danmark eller personer, som har haft permanent opholdstilladelse i Danmark i over 3 år og som bor fast i Danmark. Ud over de ovennævnte krav er det en betingelse, at det kan godtgøres, at den herboende kan forsørge den familiesammenførte. Frem til lovændringen var der ikke noget forsørgelseskrav for personer med dansk eller nordisk statsborgerskab eller flygtninge med konventions- eller de facto status.

Endelig blev muligheden for familiesammenføring af forældre afskaffet med lovændringen i 2002. Indtil lovændringen kunne personer i Danmark med dansk eller nordisk statsborgerskab eller flygtningestatus opnå familiesammenføring med forældre over 60 år. Det var dog en betingelse, at personen i Danmark rådede over en bolig af en vis størrelse og at personen påtog sig at forsørge sine udenlandske forældre.

Reglerne for familiesammenføring i Danmark er strammere end de tilsvarende fælles EU regler, der åbner mulighed for en 21-års-regel for familiesammenføring af ægtefæller, men ikke for et "tilknytningskrav".

Pr. 1. maj 2004 oprettedes den første Østtåle med det formål at sikre, at arbejdstagere fra de nye østeuropæiske EU-lande har adgang til at søge arbejde i Danmark på lige vilkår med arbejdstagere fra andre EU-lande. Den første aftale indeholdte en national overgangsordning med krav om opholds- og arbejdstilladelse betinget af tilbud om fuldtidsarbejde på overenskomst- eller overenskomstækvivalente vilkår. Aftalen revideres i april 2006 for at smidiggøre ansættelsen af arbejdskraft, hvilket bl.a. skete gennem lempelse af kravene i

overgangsordningen. Eksempelvis blev det muligt for østeuropæere at opnå opholdstilladelse betinget af deltidsarbejde. Som led i en løbende vurdering revideres aftalen igen juni 2007. Østaftalen var generelt tænkt som en midlertidig ordning og bortfaldt således helt pr. 1. maj 2009.

## 6.2 Udviklingen i indvandringen 1992 - 2010

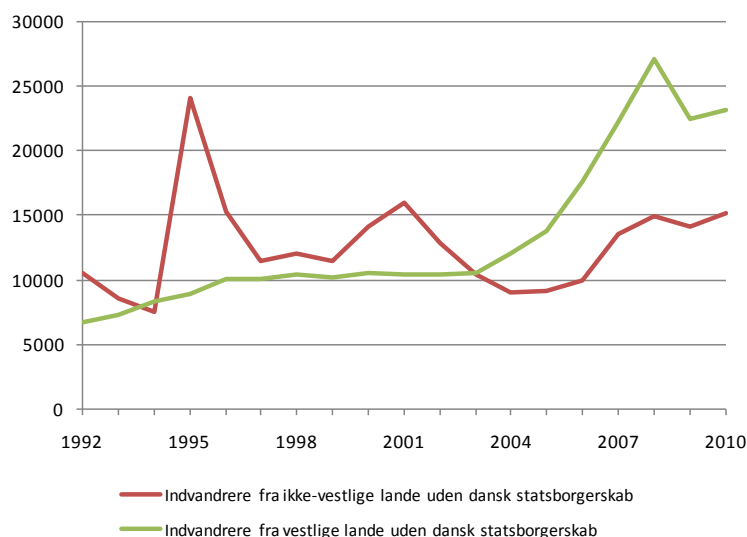
Indvandringsopgørelserne i dette afsnit omfatter ikke genindvandring, men udelukkende tilstrømning af indvandrere uden dansk statsborgerskab fra vestlige og ikke-vestlige lande. Denne type indvandring udgør langt størstedelen af indvandringen i Danmark, hvorfor fastlæggelse af netop dette mønster har central betydning for fremskrivningen. En person er opgjort som indvandret, såfremt personen ikke var til stede primo året, men er det ultimo. Personer, der når både at indvandre og udvandre inden for samme år, er således ikke omfattet af opgørelsen.

Der synes umiddelbart at være en sammenhæng mellem lovgivningsgrundlaget og udviklingen i indvandringen. Efter ændringen af lovgivningen i 1983 skete der et permanent løft i den årlige indvandring fra ikke-vestlige lande, som mere end fordobledes fra 1984 til 1985 og nåede et midlertidigt højdepunkt i 1986 (ikke illustreret). Frem til midten af 1990'erne lå indvandringen fra ikke-vestlige lande på ca. 7.000 til 10.000 personer, men steg herefter markant som følge af særlovgivningen vedr. midlertidige opholdstilladelser til personer flygtet fra krigen i det tidligere Jugoslavien og toppede således med 24.000 i år 1995<sup>31</sup>. Stramningen af udlændingeloven i 2002 ledte til et fald i indvandringen allerede i løbet af året og faldet fortsatte i perioden fra 2003 til 2005. Fra 2005 til 2008 steg indvandringen med ca. 6.000 personer til et niveau på 15.000. Fra 2008 til 2009 observeres det første fald i indvandringen fra ikke-vestlige lande gennem en længere årrække således, at niveauet i 2009 er på godt 14.000 personer. Efter 2009 observeres der igen en stigning i indvandringen, som i 2010 lå på godt 15.000 personer.

---

<sup>31</sup> Balkanstaterne tilordnes ikke-vestlige lande, men registreres som mere udviklede i den tidligere anvendte oprindelsesfordeling.

**Figur 25. Indvandringen af indvandrere fra vestlige og ikke-vestlige lande uden dansk statsborgerskab, 1992-2010**



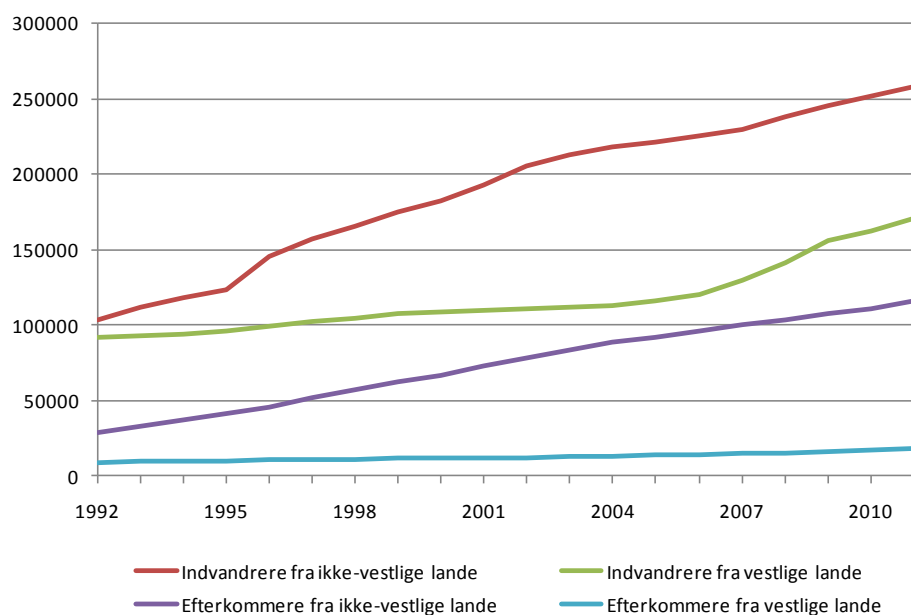
*Kilde: Udtræk fra PSD. Bemærk, at data i illustrationen ikke er korrigeret for forsinkede indberetninger.*

Indvandringen fra vestlige lande steg gradvist fra 1992 til 1996 til et niveau på ca. 10.300 personer omkring hvilket niveauet har ligget nogenlunde stabilt frem til år 2003. Siden 2004 har der været betydelig vækst i indvandringen fra vestlige lande og i år 2008 nåede den et niveau på ca. 27.000 personer, men faldt i 2009 til ca. 22.500 personer, jf. Figur 25. Fra 2009 til 2010 observeres igen en stigning i indvandringen fra vestlige lande, der i 2010 nåede et niveau på godt 23.000 personer. Tidspunktet for starten af vækstperioden kan konstateres at være sammenfaldende med tidspunktet for oprettelsen af Østaftalen.

Udviklingen i den årlige indvandring giver sammen med udvandringstilbøjeligheden anledning til udviklingen i antallet af herboende indvandrere, jf. Figur 26.

Stigningen i antallet af herboende indvandrere fra ikke-vestlige lande var forholdsvis stabil indtil primo 1996, hvor antallet relativt til det foregående år var vokset fra 123.000 til 145.000. Den ekstraordinære indvandring i 1995 betød, at der var en kraftig vækst frem til 1996, og herefter førte det højere niveau for den årlige indvandring til en fortsat vækst i antallet af herboende indvandrere fra ikke-vestlige lande. Primo 2011 er antallet nået op på næsten 260.000 personer.

**Figur 26. Antallet af herboende indvandrere og efterkommere fra vestlige og ikke-vestlige udviklede lande, 1992– 2011**



Kilde: Udtræk fra PSD.

Antallet af herboende indvandrere fra vestlige lande har været svagt stigende fra 1992-2006 og har her opnået et niveau på omkring 120.000 personer. I de seneste år er væksten grundet den stigende indvandring tiltaget og i 2011 er antallet af herboende indvandrere fra vestlige lande godt 170.000.

Den høje vækst i antallet af herboende indvandrere fra ikke-vestlige lande sammenlignet med væksten i antallet af herboende indvandrere fra vestlige lande skyldes først og fremmest, at udvandringstilbøjeligheden er væsentligt lavere for førstnævnte gruppe. Dog ses væksten i antallet af indvandrere fra vestlige lande de seneste år at have overhalet væksten i antallet fra ikke-vestlige lande. Dette må i stort omfang kunne tilskrives lempelserne i Østtalen og de samtidigt gunstige konjunkturforskel i Danmark. Således ses væksten i antallet af herboende indvandrere fra vestlige lande grundet de allerseneste års afmattende konjunkturforskel også at aftage igen efter 2009. Væksten i antallet af indvandrere fra vestlige lande minder således nu igen mere om væksten i antallet af herboende indvandrere fra ikke-vestlige lande.

Antallet af efterkommere fra ikke-vestlige lande er vokset med en betydelig og nogenlunde konstant hastighed frem til i dag, hvor antallet af herboende har nået et niveau på næsten 116.000. Stigningen skal delvist tilskrives den lave udvandringstilbøjelighed for indvandrergruppen. Yderligere indebærer alderssammensætningen i indvandrergruppen, at der er mange i den fødedygtige alder, hvilket kombineret med en historisk høj fertilitet for indvandrere fra ikke-vestlige lande relativt til andre grupper, jf. Figur 19, også har bidraget til det stigende antal efterkommere.

Antallet af herboende efterkommere fra vestlige udviklede lande er lavt og kun svagt voksende fra 1992-2011, således at der næsten skete en fordobling fra 9.200 til 18.000 over hele perioden.

## 6.3 Fremskrivning af indvandringen

### 6.3.1 Førstegangsvandring – indvandrere uden dansk statsborgerskab

Den historiske udvikling i indvandringen og i sammensætningen af denne på oprindelseslande peger på, at reguleringen af indvandringen har meget stor betydning. Det er endnu vanskeligt at vurdere langsigteffekterne af de ret betydelige ændringer, der gennemførtes i 2002. Det er derimod oplagt, at ændringerne i hvert fald på det korte sigt har haft en afgørende indflydelse på indvandringen fra specielt ikke-vestlige lande, som er reduceret meget betydeligt i perioden. Dog noteres for denne befolkningsgruppe en stigning i indvandringen fra 2005 til 2008, som man ikke kan afvise er delvis konjunkturbetinget.

På grund af kombinationen af de store ændringer, som lovgivningen har medført, og den korte tid den har virket, antages det, at der ikke sker ændringer i indvandringen i fremtiden. Det betyder, at det årlige antal af indvandrere uden dansk statsborgerskab fra henholdsvis vestlige og ikke-vestlige lande antages at være konstant i fremskrivningen. Niveauet fastlægges ud fra en række overvejelser om de lovgivnings- og konjunkturmæssige forhold.

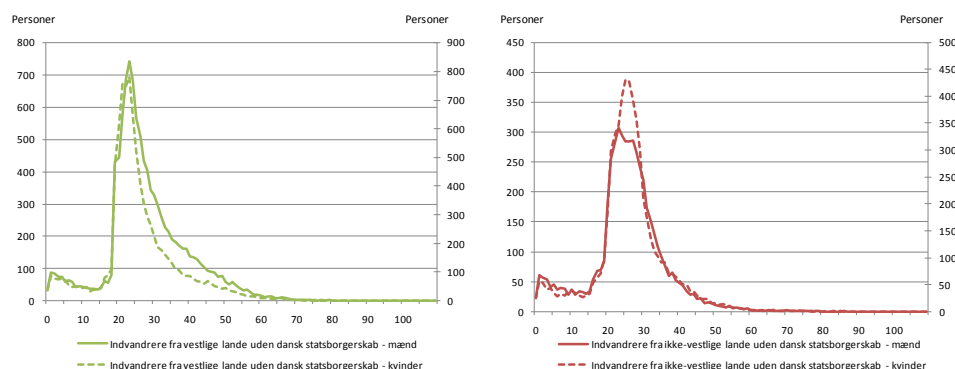
Ophævelse af overgangsordningen i Østtalen bør isoleret set lempeligøre tilgangen til det danske arbejdsmarked og sandsynliggør dermed en yderligere stigning i indvandringen fra de østlige EU-lande. Dog er lempelserne i overgangsordningen fra 2006 og 2007 direkte affødt af den på de tidspunkter store efterspørgsel efter arbejdskraft og dermed må en betragtelig andel af stigningen i indvandringen siden aftalens oprettelse anses som konjunkturafhængig. Til trods for den nylige afskaffelse af overgangsordningen er der med tanke på den økonomiske afmatning ikke grund til at tro, at det høje indvandringsniveau fra vestlige lande vil fortsætte på sit nuværende niveau i fremtiden. Således kan netop også observeres et fald i denne type indvandring på næsten 5.000 personer fra 2008 til 2009, der dog fra 2009 til 2010 efterfølges af en mindre stigning på ca. 700 personer.

Disse overvejelser sandsynliggør dermed ikke, at de seneste års stigning i indvandringen, vil afstedkomme en markant og vedvarende ændring i det fremtidige skøn for befolkningstilvæksten. Globalisering og politiske tiltag som eksempelvis Østtalen eller stramning af udlændingelovgivningen kan med rimelighed påvirke indvandringen på såvel kort, mellemlangt som langt sigt. Men det er selvsagt ikke rimeligt, at konjunkturer påvirker det langsigtede skøn. På baggrund af ovenstående er det vurderet, at udviklingen siden år 2002 er stærkt påvirket af konjunkturforsvarende, men at lovgivningen og globaliseringens rolle

også bør medtages i overvejelserne ved fastlæggelse af det fremtidige skøn for indvandringen. En mindre opjustering i det eksogene bruttoindvandringsskøn fra vestlige lande er vurderet hensigtsmæssig med det formål at opnå en fremtidig nettoindvandring og dermed befolkningstilvækst, der ikke afviger markant fra sidste års skøn. Skønnet for nettoindvandringen er endvidere i overensstemmelse med den gennemsnitlige historiske tendens. Den årlige indvandring fra vestlige lande fastlægges dermed til et omfang af 20.000 personer gennem hele fremskrivningen, mens indvandringen fra ikke-vestlige lande antages at være på 11.000 personer om året. Sammensætningen på køn og alder er fastlagt ud fra et gennemsnit af tendenserne for årene 2008-2010 i de respektive befolkningsgrupper.

For begge befolkningsgrupper gælder, at indvandringens aldersfordeling er stort set sammenfaldende for de to køn. For gruppen af indvandrere fra ikke-vestlige lande uden dansk statsborgerskab gælder, at indvandringen er koncentreret i alderen fra omkring 20 til 35 år. Indvandringen af personer i de øvrige aldersgrupper er begrænset. For gruppen af indvandrere fra vestlige lande uden dansk statsborgerskab er der mere spredning i aldersprofilen, men også for denne gruppe gælder, at indvandringen er størst for aldersgruppen fra 20 til 35 år, jf. Figur 27.

**Figur 27. Årlig fremtidig indvandring af personer uden dansk statsborgerskab fordelt på køn, alder og oprindelse.**



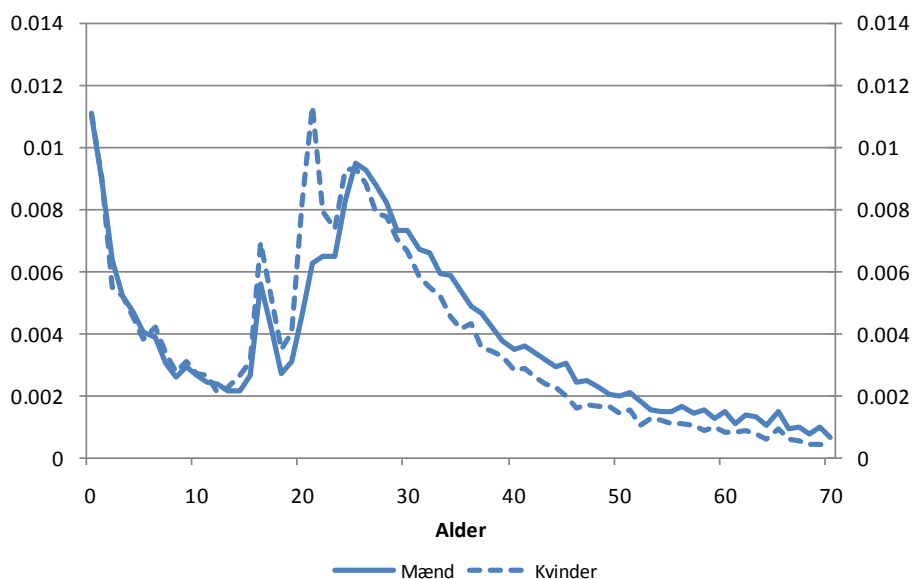
Kilde: *Befolkningsfremskrivning 2011. Fordelingen på køn og alder er fastlagt ud fra den gennemsnitlige fordeling i årene 2008-2010.*

### 6.3.2 Genindvandring

Indvandring til alle andre befolkningsgrupper end de to, der er nævnt i ovenstående afsnit betegnes genindvandring. Dette vedrører således grupperne omfattende indvandrere med dansk statsborgerskab, efterkommere og personer med dansk oprindelse. Fremskrivningen af denne type indvandring fastlægges ud fra en antagelse om, at indvandringen for givet køn og alder udgør en fast andel

af den relevante befolkningsgruppe.<sup>32</sup> Denne andel kaldes en genindvandringkvotient og er beregnet som et gennemsnit over data korrigeret for forsinkede indberetninger for perioden 2008-2010, jf. Figur 28 og Figur 29.

**Figur 28. Genindvandringkvotienter for mænd og kvinder af dansk oprindelse, gennemsnit 2008-2010**

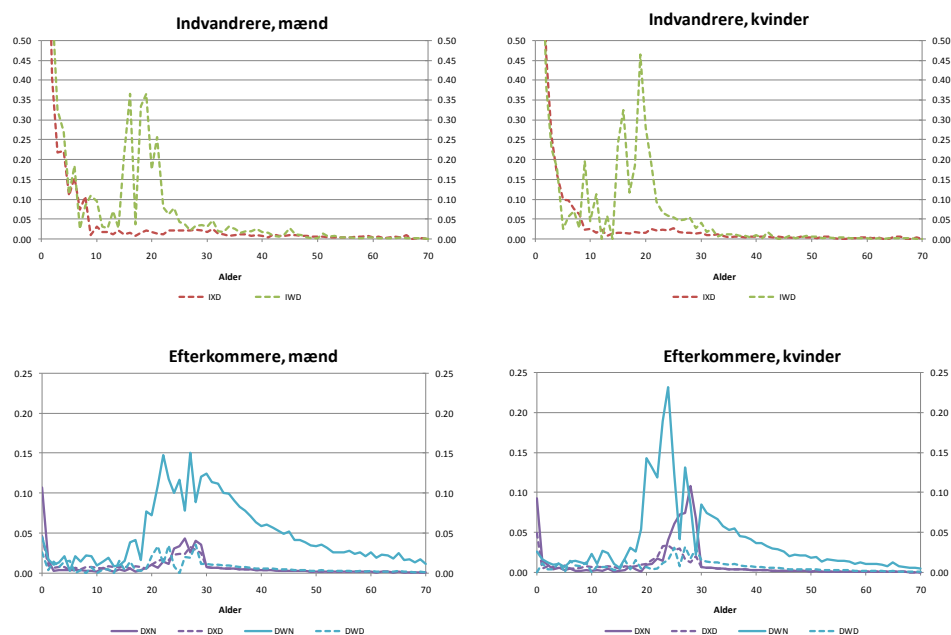


Kilde: Befolkningsfremskrivning 2011.

Genindvandringkvotienterne for mænd og kvinder af dansk oprindelse er typisk under 1 pct. for alle årgange. Eneste undtagelse er genindvandringkvotienten for 21-årige kvinder, som er omkring 1,1 pct. Det modsvarer, at udvandringkvotienten for kvinder topes for 20-årige (jf. nedenfor). For mænd findes den højeste genindvandringkvotient blandt 25-26-årige. Også dette modsvarer af, at den højeste udvandringkvotient for mænd af dansk oprindelse er blandt 24-25-årige. I begge tilfælde afspejler kvotienterne, at en del yngre personer af dansk oprindelse tager et korterevarende ophold i udlandet, jf. Figur 28.

<sup>32</sup> Man kan således tale om en "indvandringssandsynlighed".

**Figur 29. Genindvandringskvotienter for indvandrere og efterkommere fordelt på køn og alder, gennemsnit 2008-2010**



Kilde: Befolkningsfremskrivning 2011.

Aldersprofilen for genindvandringskvotienterne for indvandrere med dansk statsborgerskab fra vestlige lande minder om de tilsvarende for personer af dansk oprindelse, men niveauet for både mænd og kvinder er væsentligt højere. Samtidig er der meget høje genindvandringskvotienterne for indvandrerbørn med dansk statsborgerskab. De høje kvotienter afspejler i høj grad, at der er meget få personer i de pågældende grupper (og i særdeleshed når disse fordeles på alderstrin). For indvandrere med dansk statsborgerskab fra ikke-vestlige udviklede lande er der også meget høje genindvandringskvotienter i de yngste aldre, mens niveauet i de højere aldersgrupper minder mere om genindvandringskvotienterne for personer med dansk oprindelse, jf. Figur 29.

For samtlige efterkommere gør det sig gældende, at antallet af registreringer knyttet til personer ældre end 30 år er mangelfulde. Til udbedring af dette er det således valgt at imputere indvandringstilbøjeligheder for de pågældende aldersgrupper som en skalering af den eksisterende aldersprofil for personer af dansk oprindelse<sup>33</sup>. Som det fremgår af afsnit 6.4 anvendes en tilsvarende tilgang til korrektion af udvandringsmønsteret. Indvandringstilbøjeligheden er ca. 10 gange større for efterkommere fra vestlige lande uden dansk statsborgerskab relativt til, hvad der gør sig gældende for personer af dansk oprindelse. Indvandringstilbøjelighederne for øvrige efterkommere er på et niveau, der i større udstrækning minder om det, der kan observeres for personer af dansk oprindelse.

<sup>33</sup> En tilsvarende metode anvendes til udbedring af indvandringsdata for indvandrere med dansk statsborgerskab ældre end 70 år (ikke illustreret).

Som nævnt i afsnit 1.2.5 har anvendelsen af PSD klarlagt, at indvandringen opgjort efter registrerede hændelser er en anelse undervurderet relativt til den faktiske indvandring grundet forsinkede indberetninger. For et givet år kan noteres nævneværdige ændringer i opgørelsen op til et år efter den første indberetning. Ud fra historisk data for indvandring opgjort i to på hinanden følgende år bestemmes derfor køns- og aldersgruppeafhængige skaleringsfaktorer, der pålægges de rå indvandringsfrekvenser for at korrigere for forsinkelser. Skaleringen korrigerende for forsinkede indberetninger er gennemført umiddelbart inden aldersprofilerne for efterkommere udbedres ved skalering til profilen for personer af dansk oprindelse.

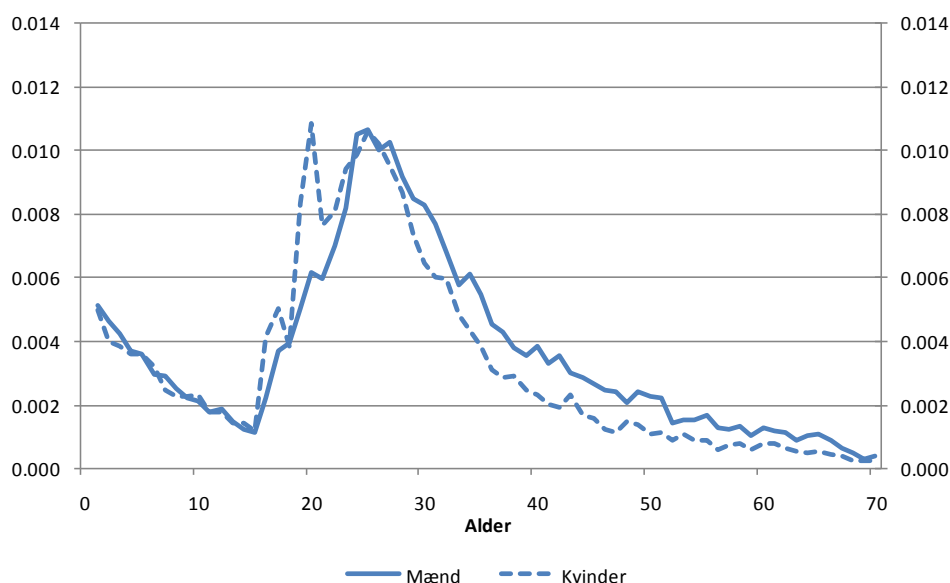
## 6.4 Udvandring

Udvandringen i fremskrivningen bestemmes for alle befolkningsgrupper ved at fastlægge en alders- og kønsspecifik udvandringssandsynlighed. Udvandringssandsynlighederne fastlægges for hver alder, køn og oprindelse som gennemsnittet af de historiske udvandringkvotienter for de seneste tre år, dvs. 2008-2010. Disse er definerede som antal udvandrede delt med antal personer i den relevante befolkningsgruppe, jf. Figur 30 og Figur 31.

Udvandringkvotienterne for mænd og kvinder af dansk oprindelse er generelt lavere end 1 pct. Dog er der en højere udvandringkvotient for 20-årige kvinder på 1,1 pct., mens også kvinder i alderen 24-26 år har en udvandringkvotient på lidt over 1 pct. For mænd findes den maksimale udvandringkvotient for 24-25-årige, hvor den er næsten 1,1 pct. Mænd i aldersgruppen 24-28 år har en udvandringkvotient på over 1 pct.

Som nævnt under afsnittet om genindvandringkvotienter er der en bemærkelsesværdig lighed i ud- og genindvandringprofilerne bortset fra en mindre aldersforskydning. Dette indikerer som nævnt, at en betydelig del af udvandringssaktiviteten skyldes kortere udenlandsophold.

**Figur 30. Udvandringskvotienter for personer af dansk oprindelse fordelt på alder og køn, gennemsnit 2008-2010**



Kilde: Befolkningsfremskrivning 2011.

De højeste udvandringskvotienter findes for gruppen af indvandrere fra vestlige lande uden dansk statsborgerskab. For både mænd og kvinder er der nogle markante toppe i aldersprofilen for udvandringskvotienterne. Den første har toppunkt omkring 5 års alderen og har et niveau på godt 15 pct. For alle aldersgrupper under 10 år observeres en udvandringssandsynlighed over 10 pct. for denne befolkningsgruppe. Det samme gælder omtrent for aldersgruppen fra 20-35 år for både mænd og kvinder. For mænd er der toppunkt i udvandringskvotienterne ved det 23. år og for kvinder ved det 22. år. For begge køn er niveauet i toppunktet på ca. 22 pct.

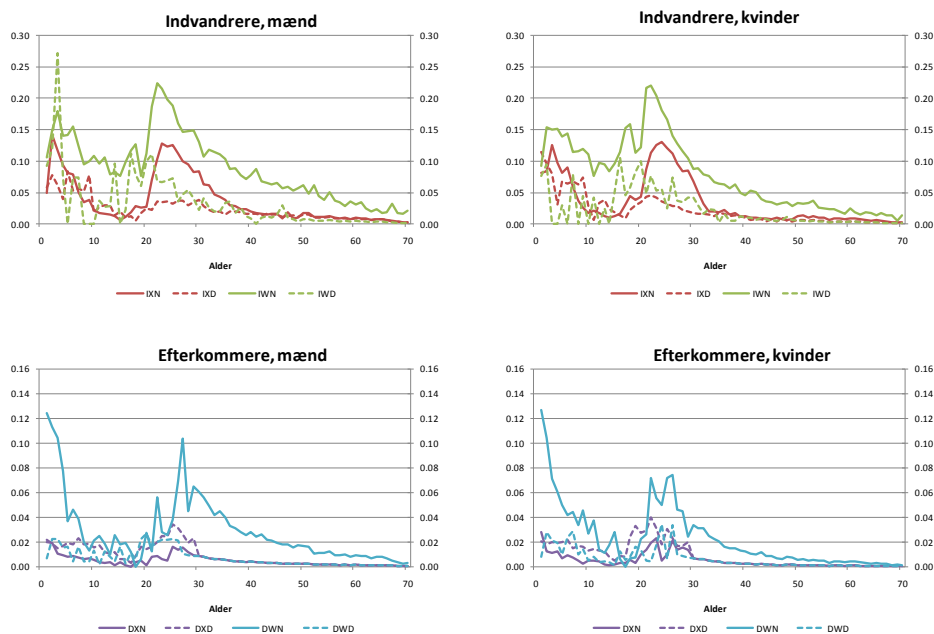
De øvrige indvandrergupper har væsentligt lavere udvandringskvotienter. Den store gruppe bestående af indvandrere fra ikke-vestlige lande uden dansk statsborgerskab har udvandringssandsynligheder, der primært kun overstiger 5 pct. for aldersgrupperne 20-30 år, jf. Figur 31.

Udvandringskvotienterne for efterkommere minder, som det ses, om dem for indvandrere, selvom niveauerne er noget lavere. Efterkommere fra vestlige lande har højere udvandringskvotienter end efterkommere fra ikke-vestlige lande. For at kompensere for mangelfuldt data, er aldersprofilen for efterkommere ældre end 30 år, som tilfældet også var det med indvandring, en skalering af udvandringsprofilen hørende til personer af dansk oprindelse. For indvandrere er der anvendt samme princip for personer ældre end 50 år.

Problemet med forsinkede indberetninger er som nævnt tidligere noget mere udtalt, når der ses på udvandring. Således ændres opgørelserne væsentligt indtil tre år efter opgørelsestidspunktet. Som tilfældet var det med indvandring er

dette forhold forsøgt udbedret gennem anvendelse af historiske køns- og aldersbestemte skaleringsfaktorer.

**Figur 31. Udvandringskvotienter for indvandrere og efterkommere fordelt på alder, køn og oprindelse, gennemsnit 2008-2010**



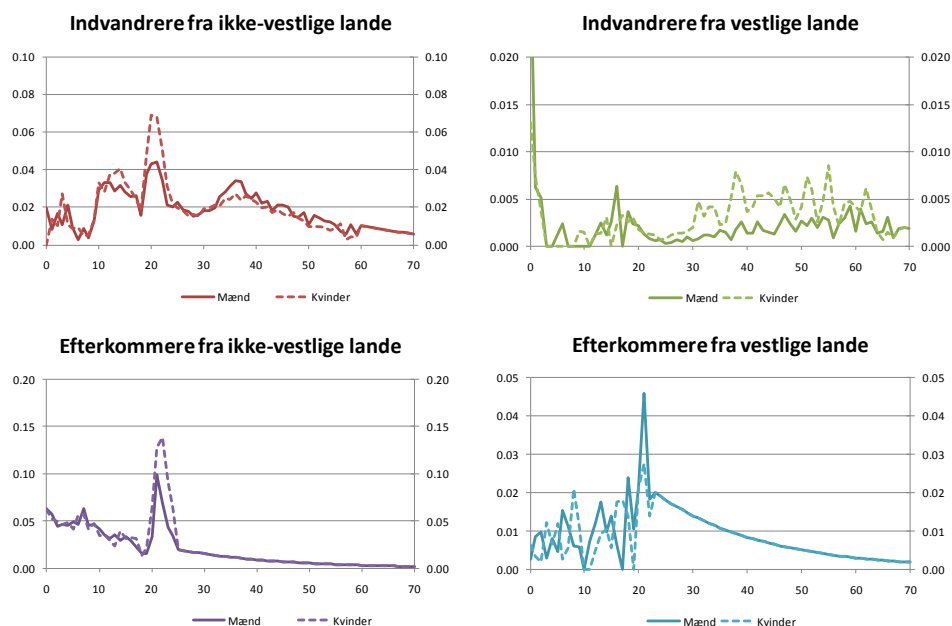
Kilde: Befolkningsfremskrivning 2011.

## 6.5 Statsborgerskabsskift

Statsborgerskabsskiftskvotienten angiver for givet køn, oprindelse og alder andelen af personer, der skifter til dansk statsborgerskab. Kvotienten er derfor kun defineret for grupper bestående af personer, der ikke er danske statsborgere. Statsborgerskabsskiftskvotienten defineres som andelen af personer i en given befolkningsgruppe, der bliver danske statsborgere. I fremskrivningen antages kvotienterne at være konstante og beregnes som et gennemsnit af de tre seneste år, dvs. her 2008 til 2010.

Umiddelbart kan observeres, at indvandrere fra vestlige lande generelt har lavere tilbøjelighed til at skifte statsborgerskab end indvandrere fra ikke-vestlige lande. Det samme indbyrdes forhold gør sig gældende mellem grupperne af efterkommere, jf. Figur 32.

**Figur 32. Statsborgerskabsskiftskvotienter for indvandrere og efterkommere fordelt på alder, køn og oprindelse, gennemsnit 2008-2010**



Kilde: Befolkningsfremskrivning 2011.

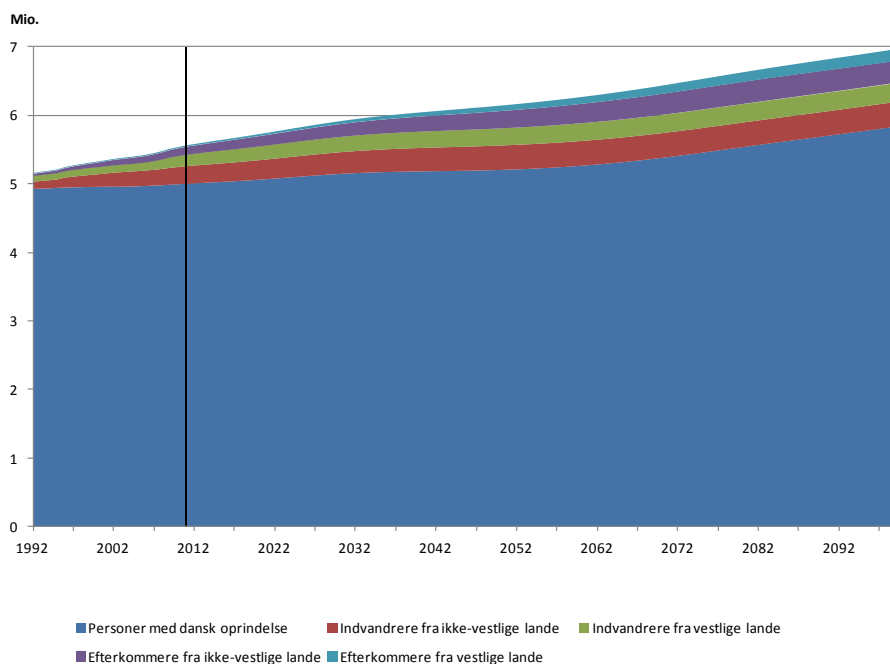
Den mest markante tilbøjelighed til at skifte statsborgerskab observeres for indvandrere og efterkommere omkring 20-års alderen, hvor niveauet specielt for efterkommere er væsentligt højere end for de omkringliggende aldre. For efterkommere ældre end 25 år og indvandrere ældre end 60 år udviser historisk data store udsving på tværs af alder, hvorfor en højere grad af stabilitet er indført gennem imputering ved en aftagende potensfunktion.

## 7 Resultater

### 7.1 Udviklingen i den samlede befolkning

Danmarks befolkning er vokset fra 2,4 mio. personer i 1901 til 5,56 mio. ved indgangen til 2011. Der har været positiv befolkningsvækst i alle årene bortset fra en kortere periode i begyndelsen af 1980'erne.

**Figur 33. Den samlede befolkning fordelt efter oprindelse, 1992-2100**



Kilde: Udtræk fra PSD (1992-2011) og Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste historiske år (2011).

2011-befolkningsfremskrivningen giver som resultat, at tendensen i befolkningsvæksten fortsætter – om end med formindsket styrke – gennem det 21. århundrede. Ved udgangen af århundredet forventes befolkningen at blive på næsten 7 mio. mennesker. Der forventes nogenlunde konstant vækst i befolkningen frem til 2024, hvorefter væksten aftager gennem den kommende 20-årige periode. Fra og med 2046 øges befolkningstilvæksten igen og i år 2074 passerer befolkningstallet 6,5 mio., jf. Figur 33.

Befolkningsvæksten er på godt 17.000 personer om året og svagt stigende i de første år af fremskrivningen. Efter år 2024 aftager tilvæksten dog gradvist indtil der nås et niveau på ca. 9.700 personer i år 2046. Herefter noteres igen en monotont stigende tilvækst frem mod år 2076, hvor befolkningen øges med ca. 19.500 personer årligt. Efterfølgende falder tilvæksten igen og når et niveau på godt 17.500 personer mod slutningen af århundredet.

Nettoindvandringen udgør i fremskrivningens start ca. 10.000 personer om året. Niveautet aftager svagt indtil år 2020, hvorefter hovedtendensen er en svag stigning, men med udsving mellem 9.100 og 9.300 personer årligt.

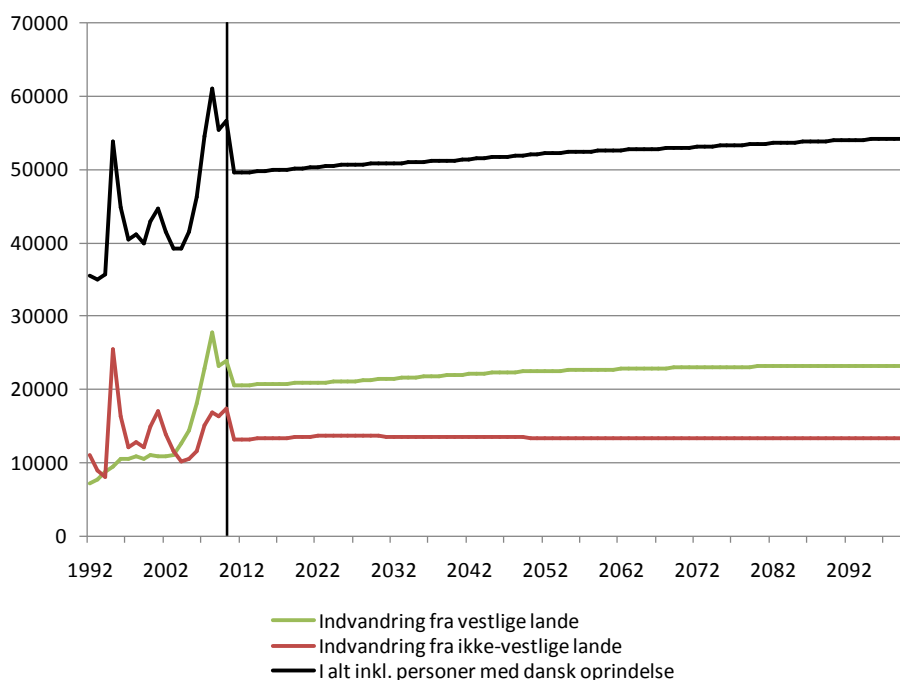
Fødselsoverskuddet (fødsler – døde) udgør den resterende del af befolkningsændringen og varierer mellem ca. 10.400 og 500 personer om året. Antallet af fødsler svinger mellem 62.000 og 75.000 om året i fremskrivningsperioden og udsvingene er betydeligt mindre end i det seneste århundrede. Antallet af døde pr. år er ligeledes rimeligt stabilt og ligger i fremskrivningsperioden typisk mellem 54.000 og 68.000 personer.

## 7.2 Indvandring og udvandring

Den samlede bruttoindvandring til Danmark var næsten 57.000 personer i 2010. En meget betydelig del af disse personer var personer, som genindvandrede til Danmark efter mere end et års ophold i udlandet. Antallet af genindvandrere domineres dog af antallet af førstegangsindvandrere, der i 2010 udgør ca. 38.400 personer.

Betragtes alene indvandringen af personer med oprindelse fra enten vestlige eller ikke-vestlige lande var den i 2010 på ca. 41.000 personer. I befolkningsfremskrivningen antages bruttoindvandringen til de fire grupper på kort sigt at falde således, at den i 2011 er små 34.000 personer, hvorefter den er svagt stigende. Indvandringen fra vestlige lande er omtrent dobbelt så stor som indvandringen fra ikke-vestlige lande.

**Figur 34. Bruttoindvandring**

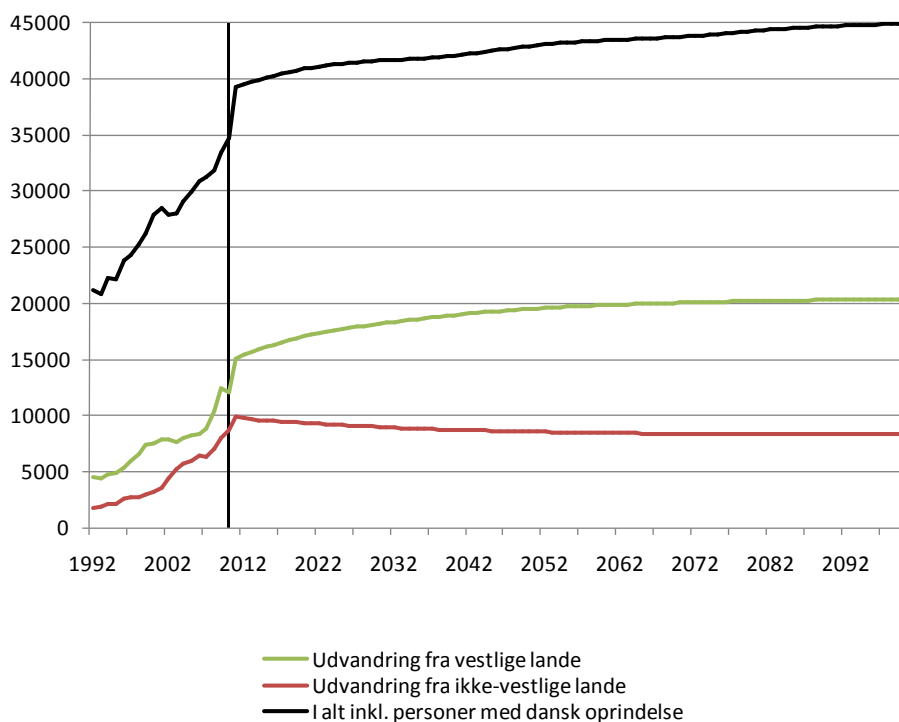


Kilde: Udtræk fra PSD (1992-2010) og Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste historiske år (2010). Bemærk, at historisk data ikke er korrigeret for forsinkede indberetninger.

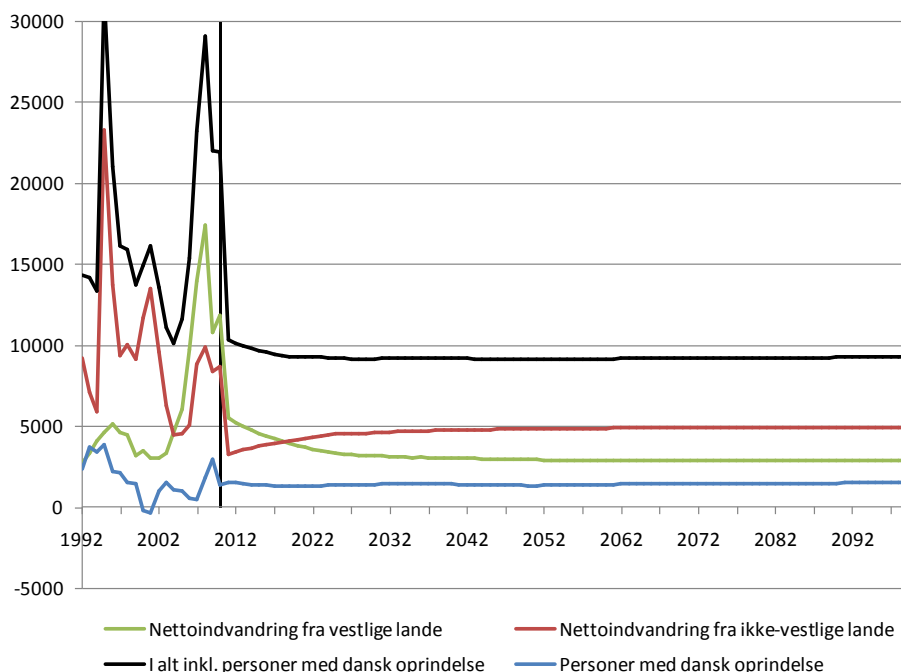
Udvandringen fra Danmark ligger på omkring 34.700 personer i år 2010 (opgjort pr. 1. januar 2011 og dermed ikke korrigeret for forsinkede indberetninger). Heraf var ca. 12.000 udvandring af personer fra vestlige lande, mens ca. 8.700 var udvandring af personer fra ikke-vestlige lande. Det er således under halvdelen af den samlede udvandring, der består af personer af dansk oprindelse.

I fremskrivningen vokser udvandringen gradvist og når 43.000 personer i år 2050. Denne stigning skyldes først og fremmest en stigning i udvandringen af efterkommere fra vestlige lande og af personer med dansk oprindelse, jf. Figur 35. Idet hastigheden med hvilken den samlede udvandring stiger er stort set identisk med stigningstakten for den samlede indvandring, vil nettoindvandringen være stort set konstant i fremskrivningsperioden. Det ses, at denne samlede tendens også kan genfindes for hver af de enkelte befolkningsgrupper, jf. Figur 36. På sigt er nettoindvandringen for personer med dansk oprindelse ca. 1.500 personer årligt, mens nettoindvandringen fra vestlige og ikke-vestlige lande har et omfang på henholdsvis 2.900 og 4.900 personer årligt.

**Figur 35. Udvandring**



Kilde: Udtræk fra PSD (1992-2010) og Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste historiske år (2010). Bemærk, at historisk data ikke er korrigeret for forsinkede indberetninger, hvilket delvist forklarer niveauforskellen mellem 2010 og 2011.

**Figur 36. Nettoindvandring**

Kilde: Udtræk fra PSD (1992-2010) og Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste historiske år (2010). Bemærk, at historisk data ikke er korrigeret for forsinkede indberetninger.

### 7.3 Befolkningsgrupperne

Ved indgangen til 2011 boede der i alt 562.517 indvandrere og efterkommere i Danmark, et antal, der stiger til ca. 860.000 i 2040. I år 2100 vil antallet af herboende indvandrere og efterkommere være vokset til godt 1,1 mio. mennesker.

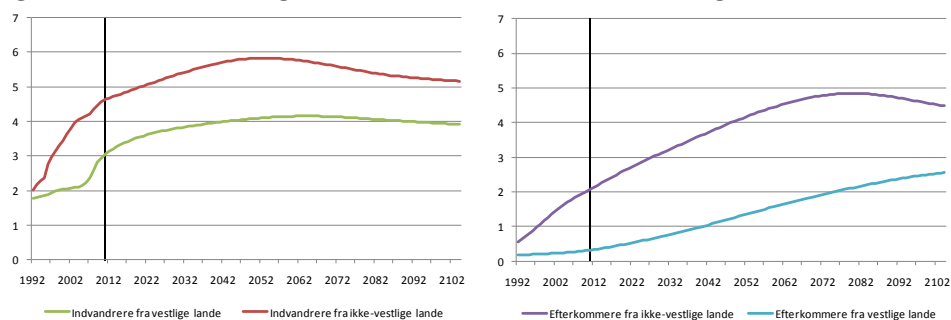
Primo 2011 boede der 4.998.111 personer af dansk oprindelse i Danmark. I fremskrivningen antages antallet af personer med dansk oprindelse fortsat at stige, men dog med formindsket vækstrate i perioden 2030-2060. Herefter genoptages den initiale stigningstakt den sidste halvdel af århundredet. Således udgøres befolkningsgruppen i år 2100 af ca. 5,85 mio. personer.

Målt som andel af den samlede befolkning udgjorde herboende indvandrere og efterkommere i 2011 i alt 10,1 pct. Denne andel vokser gradvist frem til omkring år 2080, hvor andelen når et niveau på 16,5 pct.. Herefter noteres en svagt aftagende tendens. Væksten i indvandreres og efterkommeres andel af befolkningen er aftagende over tid således, at andelen fra 2011 til 2040 er vokset med 4,1 pct. point, mens den de næste 29 år blot vokser med 2,1 pct. point.

Betragtes alene indvandrere og efterkommere fra ikke-vestlige lande udgjorde disse befolkningsgrupper tilsammen 6,7 pct. af befolkningen i 2011. Disse

befolkningsgruppers andel af den samlede befolkning er voksende frem til år 2070, hvor de tilsammen udgør omkring 10,4 pct. af befolkningen. Herefter er andelen svagt aftagende. Andelen af indvandrere fra ikke-vestlige lande topper i 2053, mens andelen af efterkommere fra ikke-vestlige lande topper i 2080. Begge befolkningsgruppers andel af den samlede befolkning aftager herefter, jf. Figur 37.

**Figur 37. Indvandrere og efterkommeres andel af befolkningen**



Kilde: Udtræk fra PSD (1992-2011) og Befolkningsfremskrivning 2011. Den lodrette streg angiver sidste historiske år (2011).

Med antagelserne om ind- og udvandring samt fertilitet fås således, at indvandreres og efterkommeres samlede andel af befolkningen forventes at blive forøget med 60 pct. ved udgangen af det 21. århundrede.

**Tabel 4. Befolkningen opdelt efter oprindelse**

	Personer med dansk oprindelse	Indvandrere fra ikke-vestlige lande	Indvandrere fra vestlige lande	Efterkommere fra ikke-vestlige lande	Efterkommere fra vestlige lande	I alt
	Personer					
1992	4 924 423	103 499	92 073	28 933	9 215	5 158 143
1993	4 928 766	112 018	92 996	32 825	9 494	5 176 099
1994	4 932 352	118 527	94 483	36 855	9 841	5 192 058
1995	4 939 124	123 622	96 621	40 936	10 180	5 210 483
1996	4 944 113	145 326	99 006	46 112	10 588	5 245 145
1997	4 946 776	157 494	102 172	51 448	10 934	5 268 824
1998	4 949 847	165 769	104 834	56 807	11 292	5 288 549
1999	4 952 931	174 738	107 254	61 917	11 597	5 308 437
2000	4 954 581	182 702	108 227	67 142	11 881	5 324 533
2001	4 956 443	193 229	109 746	72 959	12 120	5 344 497
2002	4 956 114	205 014	110 794	78 661	12 455	5 363 038
2003	4 955 992	213 399	112 278	83 811	12 824	5 378 304
2004	4 958 524	218 309	113 433	88 370	13 254	5 391 890
2005	4 962 508	221 758	116 332	92 325	13 709	5 406 632
2006	4 967 565	225 020	120 622	96 047	14 093	5 423 347
2007	4 973 200	230 057	129 291	99 908	14 670	5 447 126
2008	4 977 829	237 695	140 970	103 788	15 509	5 475 791
2009	4 985 415	245 983	155 788	107 768	16 497	5 511 451
2010	4 992 000	252 012	162 410	111 345	16 971	5 534 738
2011	4 998 111	258 146	170 758	115 597	18 016	5 560 628
2020	5 058 120	285 183	202 940	149 211	28 033	5 723 488
2030	5 142 224	316 668	224 834	185 267	42 497	5 911 490
2040	5 179 759	342 217	239 247	218 879	59 929	6 040 032
2050	5 201 114	357 559	250 630	250 818	79 495	6 139 616
2060	5 261 611	362 519	259 759	280 434	99 805	6 264 128
2070	5 376 658	361 838	266 071	305 224	120 309	6 430 099
2080	5 533 018	359 299	269 831	320 743	140 442	6 623 333
2090	5 689 757	359 192	272 414	323 022	159 013	6 803 398
2100	5 848 202	362 249	274 698	317 813	174 794	6 977 756

Kilde: Befolkningsfremskrivning 2011.

## 7.4 Aldersfordeling

Gennem de sidste hundrede år er der hele tiden blevet flere personer i den erhvervsaktive alder, som vi her definerer som personer mellem 15 og 64 år.<sup>34</sup> Ved indgangen til 1900-tallet var der således knap 1½ mio. personer i den erhvervsaktive alder, mens der i år 2000 var godt 3½ mio. personer. Dermed er der i dag over dobbelt så mange forsørgere som for hundrede år siden.

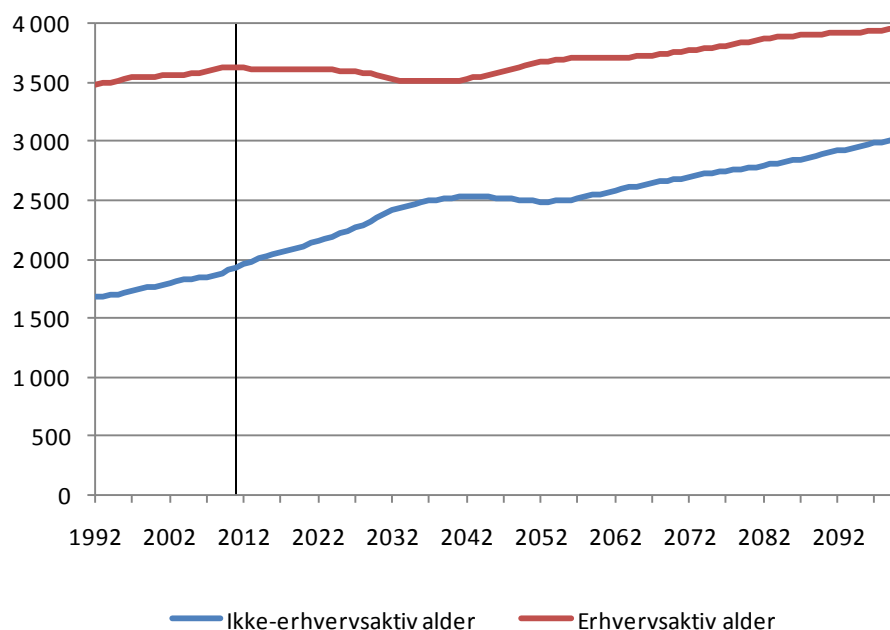
Der forventes en mindre gunstig udvikling i antallet af forsørgere i de kommende årtier, idet antallet af personer i den erhvervsaktive alder forventes at falde fra 2011. Fra et højdepunkt omkring 3,63 mio. personer i den erhvervsaktive alder i dette år, forventes antallet at falde støt til ca. 3,51 mio. personer omkring år 2038. Der forventes altså en reduktion i antallet af personer i den erhvervsaktive

<sup>34</sup> Anvendelsen af en fast aldersgrænse for den erhvervsaktive alder er næppe rimelig over så lang en periode, hvor middellevetiden er vokset betydeligt. Tallene skal derfor kun tages som en indikator for udviklingen.

alder på ca. 120.000 på en periode på omkring 30 år. I det omfang den erhvervsaktive alder udvides i perioden – bl.a. fordi levetiden stiger – kan faldet i antal personer i den erhvervsaktive alder blive mindre og eventuelt modvirkes.

I perioden efter 2038 er antallet af personer i den erhvervsaktive alder stigende og når i år 2100 op på ca. 3,96 mio. personer, jf. Figur 38. Samme år udgør de 15-64-åriges andel af befolkningen 57 pct. Til sammenligning udgør de 15-64-årige i 2011 65 pct. af befolkningen, jf. Figur 40.

**Figur 38. Antal personer hhv. i og uden for den erhvervsaktive alder, 1000 personer**



Kilde: Befolkningsfremskrivning 2011.

Samtidig er der i løbet af det seneste århundrede sket en stigning i antallet af børn og ældre, dvs. personer uden for den erhvervsaktive alder. Omkring år 1900 var der således omkring 1 mio. børn og ældre, hvilket er steget til ca. 1,93 mio. i 2011. Der er altså blevet flere personer i de aldersgrupper, som typisk skal forsyrges. Denne udvikling ventes at fortsætte i de kommende år, så der er næsten 2,53 mio. børn og ældre kort efter år 2040. Herefter forventes et midlertidigt fald i antallet af personer i disse aldersgrupper, men allerede omkring 2055 stiger antallet af personer igen. I år 2100 er antallet af personer uden for den erhvervsaktive alder vokset til 3,0 mio., jf. Figur 38.

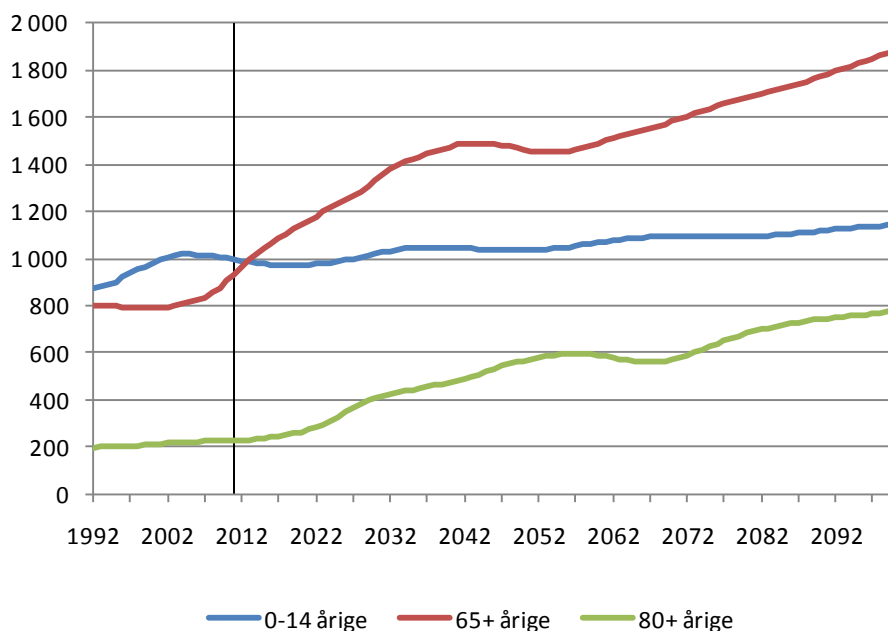
Hovedårsagen til, at antallet af personer uden for den erhvervsaktive alder er steget, er, at der er blevet langt flere ældre, mens antallet af børn – med undtagelse af en kort periode fra 1940 – ikke er steget. Således er antallet af personer over 64 år steget fra knap 200.000 i 1900 til ca. 930.000 i 2011. Denne udvikling ventes at fortsætte i de kommende år, således at antallet af ældre topper omkring 2045 med 1,49 mio. personer. Det vil sige, at der om godt 30 år

skønnes at være godt 550.000 flere ældre end i dag. I perioden fra 2045 til 2055 falder antallet af ældre lidt, hvorefter det igen er stigende. I år 2100 er antallet af ældre på over 64 år vokset til 1,88 mio. personer.

Udviklingen betyder, at mens personer over 64 år i 2011 udgør 16,8 pct. af befolkningen, vil andelen vokse til 24,6 pct. i 2042, hvor den topper midlertidigt. I 2100 udgør personer over 64 år 26,9 pct. af befolkningen, jf. Figur 40.

Antallet af børn under 15 år ventes at være nogenlunde uændret i fremskrivningsperioden. Frem mod 2020 falder antallet fra omkring 1,0 mio. i dag til 970.000, men herefter vil der være en gradvis stigning mod et niveau på godt 1,1 mio. personer i slutningen af århundredet, jf. Figur 39.

**Figur 39. Antal personer under 15 år og over hhv. 65 år og 80 år, 1000 personer**



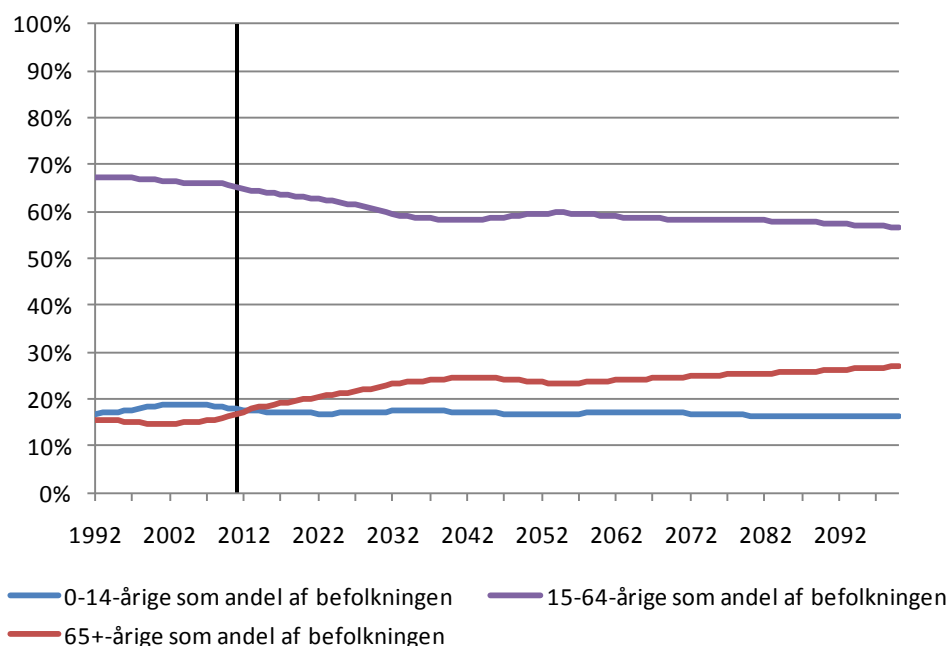
Kilde: Befolkningsfremskrivning 2011.

Der forventes en relativt større stigning i antallet af ældste ældre end i ældregruppen som helhed. Således skønnes antallet af personer på 80 år og derover, at blive mere end fordoblet fra omkring 229.000 til 479.000 personer i løbet af de kommende 30 år. Væksten i antallet af personer på 80 år og derover fortsætter med at vokse – dog ikke monotont – og antallet udgør i 2100 780.000 personer, jf. Figur 39. Denne stigning er specielt interessant, fordi de offentlige serviceudgifter pr. person er større for denne gruppe end for andre aldersgrupper.

I forhold til den fremtidige finansiering af velfærdssamfundet er udviklingen i forholdet mellem de forskellige aldersgrupper i befolkningen af afgørende betydning. Størrelsesforholdet mellem forskellige aldersgrupper i befolkningen måles ofte ved udviklingen i den demografiske forsørgerkvote, der sætter

antallet af børn og ældre i forhold til antallet af personer i den erhvervsaktive alder. Målet kan betragtes som et groft mål for antallet af personer, som skal forsørges, delt med antallet af potentielle forsørgere. Den kraftige stigning i sidste århundrede i antallet af personer i den erhvervsaktive alder, har historisk betydet et fald i den demografiske forsørgerkvote. Grundet aldring af de pågældende erhvervsaktive, er den nuværende tendens dog en stigende forsørgerkvote.

**Figur 40. Befolkningen fordelt efter alder i pct. af den samlede befolkning**



Kilde: Befolkningsfremskrivning 2011.

*Den demografiske forsørgerkvote* defineres som summen af antallet af personer i alderen 0-14 år og antallet af personer, der er 65 år eller derover, divideret med antallet af personer i alderen 15-64 år.

Målt ved dette begreb skulle godt 6 erhvervsaktive for hundrede år siden forsørge 4 personer, mens 6 erhvervsaktive i dag kun skal forsørge lidt over 3 personer uden for den erhvervsaktive alder. Alternativt kan man sige, at 20 personer skal forsørge lidt over 10 personer. Fremover forventes der at blive færre i den erhvervsaktive alder og flere uden for den erhvervsaktive alder. Det betyder, at det omkring år 2040 forventes at 20 erhvervsaktive skal forsørge godt 14 personer, der ikke er i den erhvervsaktive alder. Fra 2055 stiger forsørgerkvoten til næsten 0,8 frem mod år 2100, jf. Figur 41. Således skal der her 20 erhvervsaktive til at forsørge godt 15 ikke-erhvervsaktive.

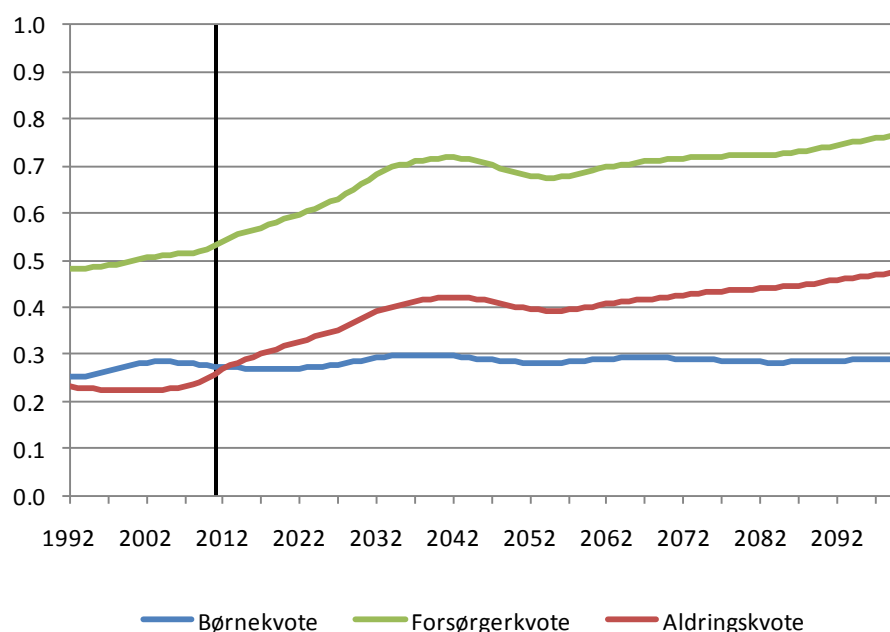
Sammensætningen af personerne uden for den erhvervsaktive alder er dog markant anderledes end ved det 20. århundredes begyndelse. Dette ses ved at opdele den demografiske forsørgerkvote i henholdsvis en demografisk børnekvote og en demografisk ældrekvote.

Den demografiske børnekvote er defineret som antallet af personer i alderen 0 – 14 år divideret med antallet af personer i alderen 15 – 64 år.

Den demografiske ældrekvote er defineret som antallet af personer på 65 år og derover divideret med antallet af personer i alderen 15 – 64 år.

I løbet af det 20. århundrede har antallet af børn ligget nogenlunde konstant, mens der er blevet flere i den erhvervsaktive alder. Således var der i år 1900 omkring lidt mere end et barn pr. 2 forsørgere, hvilket i dag er faldet til lidt under det halve. I fremskrivningen fastholdes dette niveau, således at børnekvoten stabiliseres omkring 0,3, jf. Figur 41.

**Figur 41. Demografiske kvoter**



Kilde: Befolkningsfremskrivning 2011.

Omvendt er antallet af ældre gennem det 20. århundrede vokset betydeligt mere end antallet af personer i den erhvervsaktive alder. I begyndelsen af århundredet var der 1 ældre pr. 10 personer i alderen fra 15 – 64 år. I 2011 skal de samme erhvervsaktive forsøge mere end 2 ældre. Frem til 2042 forventes den demografiske ældrekvote næsten at blive fordoblet i forhold til 2011, idet den forventes at nå et niveau på godt 0,4. Efter et midlertidigt fald i ældrekvoten stiger denne igen i århundredets sidste halvdel og nærmer sig 0,5 ved udgangen af århundredet.

Ændringen i sammensætningen af den demografiske forsørgerkvote – fra hovedsageligt at dække over forsørgelse af børn i starten af det 20. århundrede til i dag at repræsentere en stor set lige fordeling af børn og ældre og i fremtiden til hovedsageligt at være en forsørgelse af ældre – har betydning for de offentlige finanser. Det skyldes, at de offentlige udgifter til en person på 65 år eller derover

– med den nuværende udgiftssammensætning – er væsentlig højere end udgifterne til et barn.

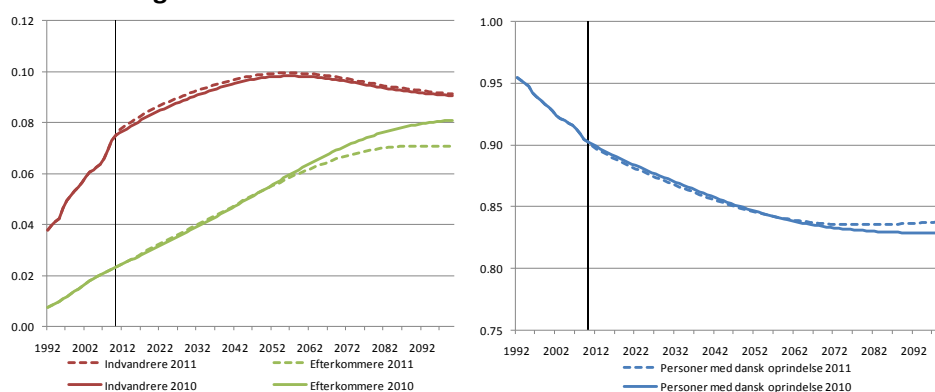
## 8 Sammenligning med Befolkningsfremskrivning 2010

I nærværende afsnit sammenholdes i hovedtræk resultaterne fra 2011-fremskrivningen med Befolkningsfremskrivning 2010. Sammenligningen vil begrænse sig til at omfatte tre grupper bestående af personer af dansk oprindelse, indvandrere samt efterkommere.

### 8.1 Udviklingen i den samlede befolkning

Befolkningens sammensætning på oprindelse ændrer sig ikke markant i 2011-fremskrivningen relativt til Befolkningsfremskrivning 2010, jf. Figur 42. Dog er det overordnet værd at bemærke, at efterkommeres andel af befolkningen er faldet på sigt, hvis årsag klarlægges i det følgende.

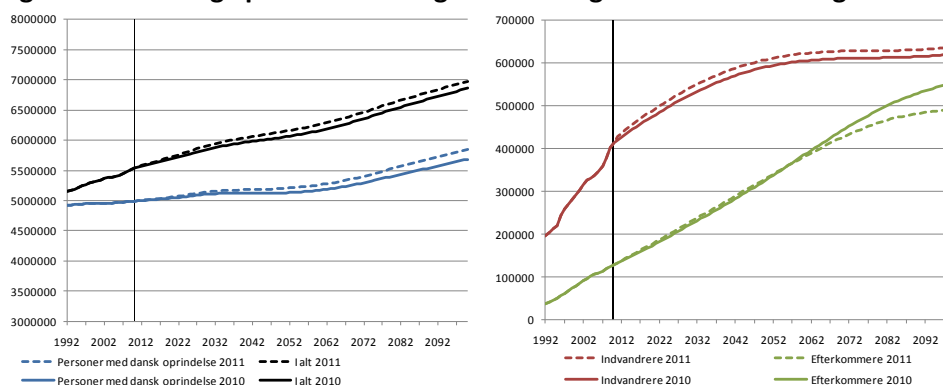
**Figur 42. Udvikling i befolkningens sammensætning i 2010- og 2011-fremskrivningen**



Kilde: DREAMs befolkningsfremskrivning 2010 og Befolkningsfremskrivning 2011. Bemærk, at historisk data varierer, idet data anvendt til 2011-fremskrivningen stammer fra et opdateret udtræk fra PSD. Lodret akse angiver seneste historiske år, her 2011.

Overordnet afstedkommer opdateret data samt metoderevisioner, at den samlede befolkning i år 2050 og år 2100 er øget med hhv. godt 90.000 og 115.000 personer relativt til sidste års fremskrivning, jf. Figur 43. Stigningen dækker, som det klarlægges i det følgende, over et vedvarende højere fødselsoverskud end i sidste års fremskrivning, der på sigt er tilstrækkelig til at dominere en lavere nettoindvandring. Frem til omkring 2050, er den samlede nettoindvandring dog højere end i sidste års fremskrivning.

Antallet af personer med dansk oprindelse er gennem det meste af fremskrivningen større end sidste år. Således er antallet i 2050 øget med ca. 72.000 og antallet i år 2100 er øget med ca. 163.000. Stigningen er foranlediget af en initialt højere fertilitet og en vedvarende højere nettoindvandring end sidste år.

**Figur 43. Udvikling i primobefolkningen i 2010- og 2011-fremskrivningen**

Kilde: DREAMs befolkningsfremskrivning 2010 og Befolkningsfremskrivning 2011. Bemærk, at historisk data varierer, idet data anvendt til 2011-fremskrivningen stammer fra et opdateret udtræk fra PSD. Lodret akse angiver seneste historiske år, her 2011.

Det samlede antal indvandrere i befolkningen er konsekvent højere i 2011-fremskrivningen relativt til sidste år. I år 2050 er der således ca. 17.000 flere indvandrere i befolkningen, mens forskellen i år 2100 er ca. 16.000. Dette dækker over en langsigtet stigning i antallet af indvandrere fra vestlige lande på ca. 33.000 personer og et fald i antallet af ikke-vestlige indvandrere på ca. 17.000 personer.

Årsagen til denne udvikling er en stigning i nettoindvandringen fra vestlige lande i 2011-fremskrivningen, der primært skal tilskrives, at skønnet for eksogen indvandring er opjusteret med 3.000 personer årligt. De pågældende indvandrere har dog også en relativt høj udvandringstilbøjelighed, der på sigt mindsker forskellen i nettoindvandringen relativt til sidste års fremskrivning. Den samlede nettoindvandring af indvandrere er initialt højere end i sidste års fremskrivning, men vil dog på sigt være lidt mindre, jf. Figur 49, hvilket tilskrives et fald i nettoindvandringen for ikke-vestlige indvandrere.

Befolkningsvæksten for indvandrere forstærkes yderligere af et marginalt fald i antallet af dødsfald, jf. Figur 48. Den primære årsag til dette skal ikke findes i et fald i restlevetiden, men derimod i ændret omfang af nettoindvandringen. For indvandrere fra vestlige lande gælder, at der typisk er en ret høj udvandringssandsynlighed, hvorfor stigningen i antallet af herboende hovedsageligt vil være lokaliseret i de aldersgrupper, hvor ændringer i dødeligheden ikke er markante.

Antallet af efterkommere er med undtagelse af det lange sigt marginalt højere i 2011-fremskrivningen relativt til sidste år. Kombinationen af en initialt højere indvandring fra ikke-vestlige lande og en kortsigtet højere fertilitet afstedkommer en initial stigning i fødselsoverskuddet for efterkommere fra ikke-vestlige lande relativt til sidste års fremskrivning. Trods en højere nettoindvandring af efterkommere fra ikke-vestlige lande i årets fremskrivning, vil faldet i antallet af indvandrere fra ikke-vestlige lande dog betyde et fald i fødselsoverskuddet for den tilhørende efterkommergruppe gennem det meste af

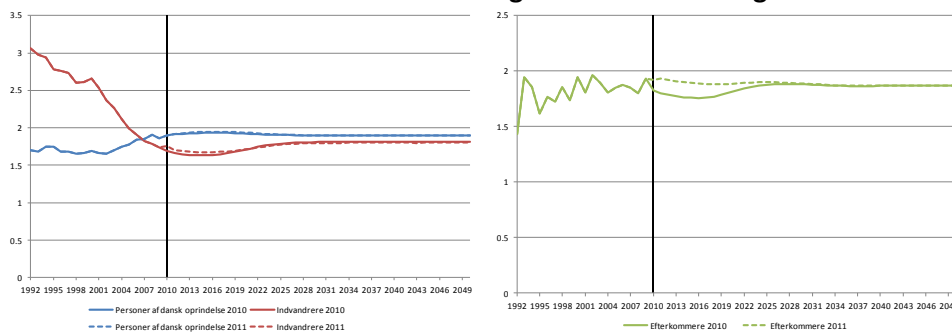
fremskrivningen. På sigt er der i årets fremskrivning ca. 6.000 færre efterkommere fra ikke-vestlige lande end tilfældet var sidste år.

Det permanent højere antal vestlige indvandrere i årets fremskrivning vil sammen med den kortsigtede stigning i fertiliteten betyde et højere antal fødte, jf. Figur 46, og i forlængelse heraf også et højere fødselsoverskud af vestlige efterkommere. På sigt kan dette dog ikke opveje en markant nedskrivning af nettoindvandringen, jf. Figur 49 og Figur 50. Den langsigtede nedskrivning af efterkommeres nettoindvandring hidrører primært efterkommere fra vestlige lande og skal i stor udstrækning tilskrives indførelse af en restriktion af efterkommeres genindvandring, der afhjælper en tidligere mangel i modellen, der fejlagtigt over tid tillod antallet af medlemmer af en generation at overstige det antal, der oprindeligt var født ind i generationen.

## 8.2 Fødsler og fertilitet

Som forklaret i afsnit 5.6, anvendes med undtagelse af fordeling på statsborgerskab samme langsigtetsniveauer i 2011-fremskrivningen som i 2010-fremskrivningen. Nedenfor er sammenholdt udviklingen i den samlede fertilitet for hhv. personer af dansk oprindelse, indvandrere og efterkommere i de to fremskrivninger, jf. Figur 44.

**Figur 44. Udviklingen i den samlede fertilitet for personer af dansk oprindelse, indvandrere samt efterkommere i 2010- og 2011-fremskrivningen**



*Kilde: DREAMs befolkningsfremskrivning 2010 og Befolkningsfremskrivning 2011. Variationer i historisk data skal tilskrives opdatering af udtræk fra PSD relativt til sidste år. Lodret akse angiver seneste historiske år, her 2010.*

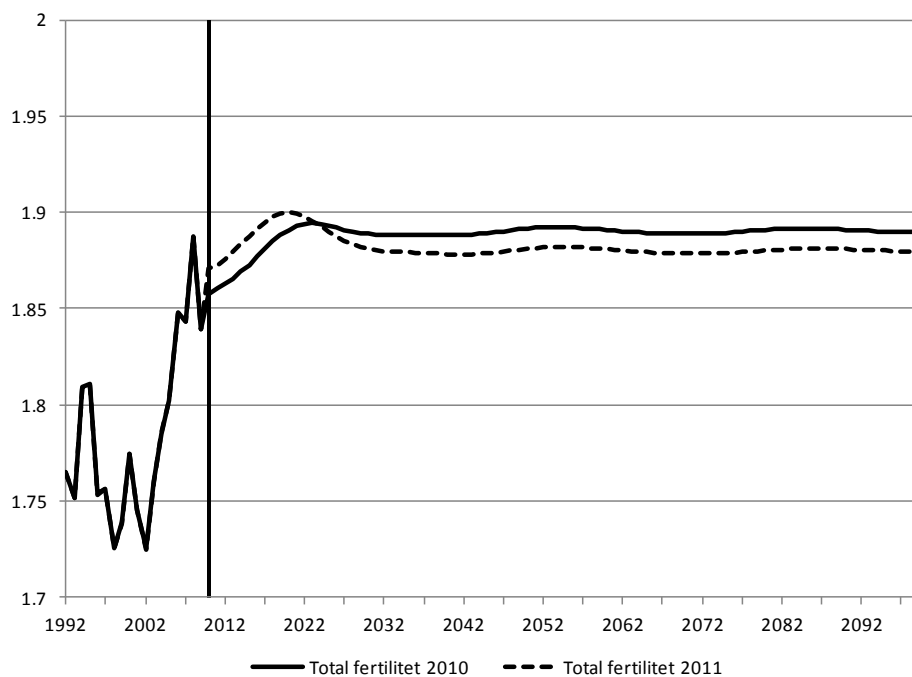
Som det ses, vil den nyeste udvikling tillades at indvirke på fertiliteten på det korte sigt, mens tilpasning til langsigtetsniveauet efterfølgende indfinder sig rimeligt hurtigt. Specielt for efterkommere ses betydningen af tilgængeligheden af et ekstra dataår. At den faktiske fertilitet i år 2010 for efterkommere er på højde med niveauet i år 2009, betyder at trendretningen for fremskrivningen vil ændre sig, hvorfor det kortsigtede niveau nu er på niveau med det langsigtede skøn. For alle oprindelsesgrupper ses forskellen mellem de to fremskrivninger at være stort set elimineret omkring år 2030.

Justering af de langsigtede statsborgerskabsfordelte skøn for indvandrere betyder, at den løbende ændring i statsborgerskabs sammensætningen vil

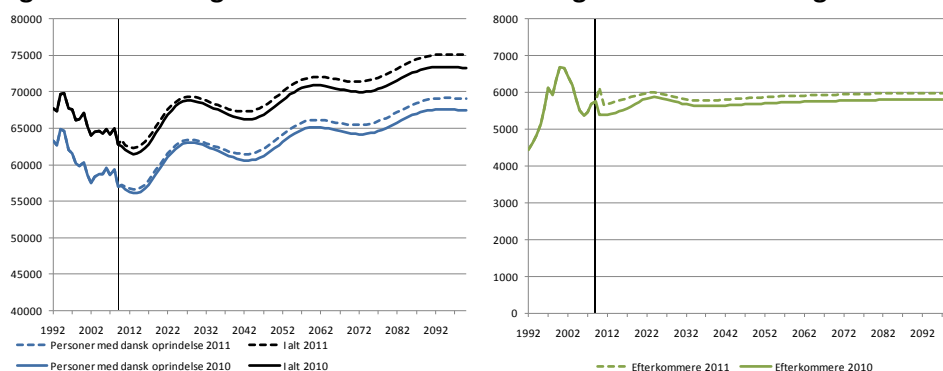
foranledige, at den samlede fertilitet for indvandrere i 2011-fremskrivningen på sigt er 1,80 mod 1,82 i sidste års fremskrivning. Dette dækker over en lille nedjustering af fertiliteten for vestlige indvandrere og en opjustering af fertiliteten for ikke-vestlige. På det korte sigt kan som tilfældet også var for efterkommere konstateres en stigning i fertiliteten for både indvandrere og personer af dansk oprindelse.

Kombineres befolkningens sammensætning på oprindelse med de enkelte oprindelsesgruppers fertilitet på alle tidspunkter i fremskrivningen, giver dette anledning til udviklingen i den samlede fertilitet for hele befolkningen. Denne er i Figur 45 sammenholdt for 2010- og 2011-fremskrivningen. Nedjusteringen af den samlede fertilitet fra 1,89 til 1,88 skal tilskrives kombinationen af en marginalt øget andel af indvandrere og en lidt lavere langsigtet fertilitet for denne gruppe relativt til sidste års fremskrivning. Udviklingen i antallet af fødsler fordelt på oprindelse er sammenholdt i Figur 46.

**Figur 45. Udviklingen i den totale fertilitet i 2010- og 2011-fremskrivningen**



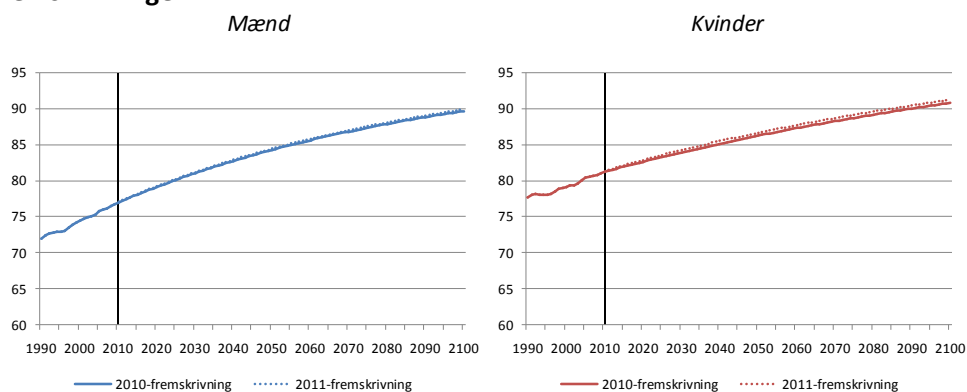
*Kilde: DREAMs befolkningsfremskrivning 2010 og Befolkningsfremskrivning 2011. Variationer i historisk data skal tilskrives opdatering af udtræk fra PSD relativt til sidste år. Lodret akse angiver seneste historiske år, her 2010.*

**Figur 46. Udvikling i antallet af fødsler i 2010- og 2011 fremskrivningen**

Kilde: DREAMs befolkningsfremskrivning 2010 og Befolkningsfremskrivning 2011. Variationer i historisk data skal tilskrives opdatering af udtræk fra PSD relativt til sidste år. Lodret akse angiver seneste historiske år, her 2010.

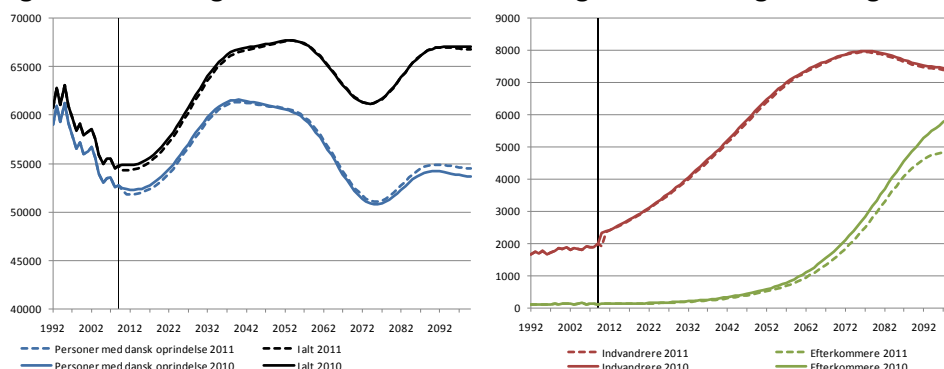
### 8.3 Dødsfald og middellevetid

For såvel mænd som kvinder kan noteres en stigning i middellevetiden fra 2009 til 2010, hvilket har indvirket på udviklingen relativt til sidste år. I 2050 er middellevetiden for mænd lidt under 0,2 år højere i dette års fremskrivning end sidste år og vil dermed ligge på 84,5 år, mens forskellen i år 2100 er øget til godt 0,2 år, hvor middellevetiden vil nå et niveau på 89,9 år.

**Figur 47. Udviklingen i middellevetiden for mænd og kvinder i 2010- og 2011-fremskrivningen**

Kilde: DREAMs befolkningsfremskrivning 2009 og Befolkningsfremskrivning 2010. Variationer i historisk data skal tilskrives opdatering af udtræk fra PSD relativt til sidste år. Lodret akse angiver seneste historiske år, her 2009.

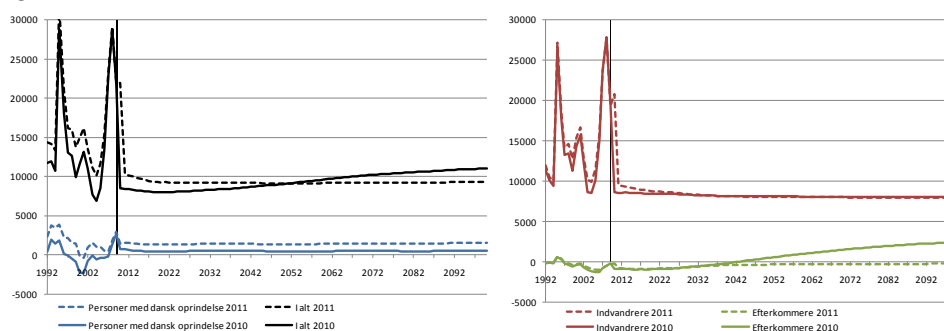
For kvinder er middellevetiden i år 2050 og 2100 henholdsvis 86,7 år og 91,2 år, hvilket er 0,4 år højere end i sidste års fremskrivning. Udviklingen i antallet af dødsfald, jf. Figur 48, vil være relativt robust overfor ændring af dødelighedsforecastet mellem 2010 og 2011. Niveauforskellene vil derfor primært være afstedkommet af den ændrede udvikling i primobefolkningen.

**Figur 48. Udvikling i antallet af døde i Befolkningsfremskrivning 2010- og 2011**

Kilde: Befolkningsfremskrivning 2010 og 2011. Variationer i historisk data skal tilskrives opdatering af udtræk fra PSD relativt til sidste år. Lodret akse angiver seneste historiske år, her 2010.

## 8.4 Vandringer

Udviklingen i nettoindvandringen er illustreret i Figur 49 og allerede omtalt ovenfor. Udviklingen er afledt af forskellen mellem ind- og udvandring illustreret i Figur 50 og Figur 51. De historiske data er i figurerne ikke korrigerede for forsinkelser og betydningen ses som tidligere nævnt at være størst for udvandringsdata. Fremskrevne værdier for ind- og udvandring er dog baseret på hyppigheder korrigeret for forsinkelser.

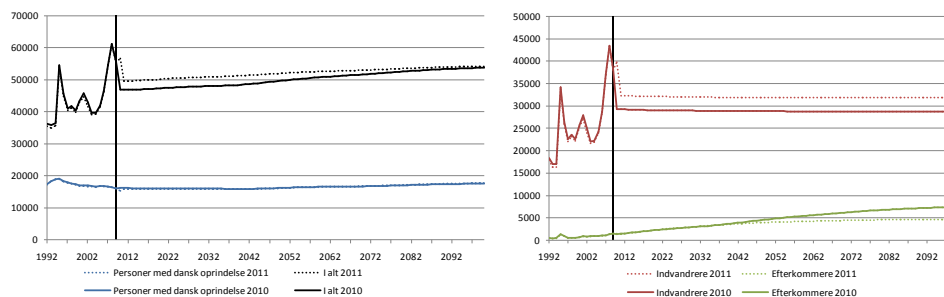
**Figur 49. Udvikling i nettoindvandringen i Befolkningsfremskrivning 2010- og 2011**

Kilde: DREAMs befolkningsfremskrivning 2010 og 2011. Variationer i historisk data skal tilskrives opdatering af udtræk fra PSD relativt til sidste år. Desuden er historisk data i figuren ikke korrigeret for forsinkede indberetninger. Lodret akse angiver seneste historiske år, her 2010.

Forskellen mellem 2010- og 2011-fremskrivningen skal tilskrives en kombination af opdatering af datagrundlaget, herunder opdatering af korrektionen for forsinkede indberetninger samt tilføjelse af et nyt dataår. Eksogen indvandring af indvandrere uden dansk statsborgerskab var i 2010-fremskrivningen konstant 28.000 personer om året (heraf 11.000 fra ikke-vestlige lande og 17.000 fra

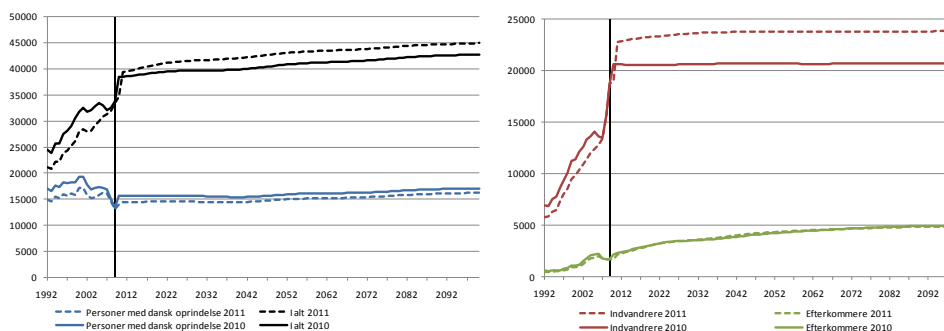
vestlige lande), mens det tilsvarende niveau i 2011-fremskrivningen er fastsat til 31.000 personer grundet tidligere nævnt opjustering af vestlig indvandring med 3.000 personer. Indvandringen til gruppen bestående af indvandrere med dansk statsborgerskab ses således at have mindre betydning, jf. Figur 50.

**Figur 50. Udvikling i indvandringen i Befolkningsfremskrivning 2010- og 2011**



Kilde: Befolkningsfremskrivning 2010 og 2011. Variationer i historisk data skal tilskrives opdatering af udtræk fra PSD relativt til sidste år. Desuden er historisk data i figuren ikke korrigeret for forsinkede indberetninger. Lodret akse angiver seneste historiske år, her 2010.

**Figur 51. Udvikling i udvandringen i Befolkningsfremskrivning 2010- og 2011**



Kilde: Befolkningsfremskrivning 2010 og 2011. Variationer i historisk data skal tilskrives opdatering af udtræk fra PSD relativt til sidste år. Desuden er historisk data i figuren ikke korrigeret for forsinkede indberetninger. Lodret akse angiver seneste historiske år, her 2010.

Nedenstående tabel sammenholder i oversigtsform de kvantitative forskelle mellem 2010- og 2011-fremskrivningerne.

**Tabel 5. Beholdninger og strømme i 2010- og 2011-fremskrivningerne i år 2050 og år 2100.**

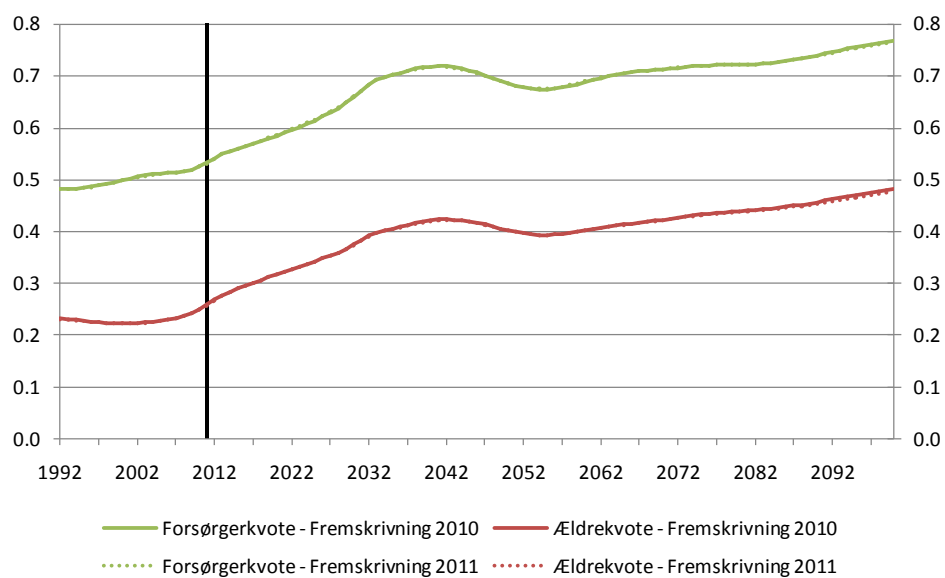
	2010-fremskrivning, 2050	2011-fremskrivning, 2050	2010-fremskrivning, 2100	2011-fremskrivning, 2100
<b>Befolkning</b>				
Personer af dansk oprindelse	5 129 254	5 201 114	5 685 536	5 848 202
Indvandrere	590 873	608 190	620 981	636 947
Efterkommere	326 524	330 313	555 642	492 607
<b>Indvandring</b>				
Personer af dansk oprindelse	16 258	16 212	17 628	17 792
Indvandrere	28 826	31 869	28 778	31 815
Efterkommere	4 660	3 975	7 423	4 674
<b>Udvandring</b>				
Personer af dansk oprindelse	15 824	14 877	17 094	16 252
Indvandrere	20 689	23 782	20 720	23 840
Efterkommere	4 175	4 285	4 976	4 880
<b>Fertilitet</b>				
Total	1.89	1.88	1.89	1.88
Personer af dansk oprindelse	1.90	1.90	1.90	1.90
Indvandrere	1.82	1.80	1.82	1.80
Efterkommere	1.87	1.87	1.86	1.86
<b>Restlevetid</b>				
Middellevetid, mænd	84.3	84.5	89.7	89.9
Restlevetid for 60-årige, mænd	26.1	26.2	30.5	30.7
Middellevetid, kvinder	86.3	86.7	90.8	91.2
Restlevetid for 60-årige, kvinder	27.3	27.6	31.2	31.5

Kilde: *Befolkningsfremskrivning 2010 og 2011.*

## 8.5 Udvikling i forsørger- og ældrekvote

Både forsørger- og ældrekvoten er stort set uforandrede i forhold til 2010-fremskrivningen. For både forsørger- og ældrekvotens vedkommende spores dog mod fremskrivningens slutning marginale fald relativt til sidste års fremskrivning, jf. Figur 52. Indledningsvist afslører data endvidere, at forsørgerkvoten er højere end sidste år og da ældrekvoten i samme periode ikke er ændret nævneværdigt, kan ændringen i forsørgerkvoten henføres til stigningen i antallet af børn, der er direkte foranlediget af det højere antal fødsler i fremskrivningens begyndelse, jf. Figur 46.

**Figur 52. Udviklingen i forsørger- og ældrekvoten i hhv. 2010- og 2011-fremskrivningen**



*Kilde: Befolkningsfremskrivning 2010 og 2011. Variationer i historisk data skal tilskrives opdatering af udtræk fra PSD relativt til sidste år. Den lodrette akse angiver seneste historiske år, her 2011.*

## 9 Referencer

Alho, J. M. & Spencer B. D. (2005): *Statistical Demography and Forecasting*, Springer.

Andreev, K.F. and J. Vaupel (2006): *Forecasts of Cohort Mortality After Age 50, MPIDR Working Paper WP 2006-012*, Max Planck Institute for Demographic Research, Rostock, Germany.

Bell, W. R. (1997): Comparing and Assessing Time Series Methods for Forecasting Age-Specific Fertility and Mortality Rates, *Journal of Official Statistics* vol. 13 no. 3 279-303.

Burrage, K, A. Williams, J. Erhel & B. Pohl (1994): The implementation of a Generalized Cross Validation algorithm using deflation techniques for linear systems, *Research Report no. 94-05*, Seminar für Angewandte Mathematik, Eidgenössische Technische Hochschule, Switzerland.

Chatfield, C. (2004): *The Analysis of Time Series – An Introduction*, Chapman & Hall/CRC, 6. udgave.

D'Addio, A. C. & M. M. d'Ercole (2005): Policies, Institutions and fertility rates: A panel data analysis for OECD countries, *OECD Economic Studies* No. 41, 2005/2.

Danmarks Statistik (1978): *Befolkningsprognoser 1977-2000*, *Statistiske undersøgelser* nr. 36.

Danmarks Statistik (2003): *Befolkningens bevægelser*.

Danmarks Statistik (2010a): *Indvandrere i Danmark 2010*.

Danmarks Statistik (2010b): *NYT fra Danmarks Statistik, nr. 124, marts 2010*.

De Jong, P. & L. Tickle (2005): *Extending Lee-Carter Mortality Forecasting*, Working Paper Institute of Actuaries of Australia.

Gompertz, B. (1825): On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society of London, Vol. 115 (1825)*, 513-585.

Hansen M. F. & Barington M. L. (2009): *Danmarks fremtidige befolkning. Befolkningsfremskrivning 2009. DREAM rapport 2009*. [www.dreammodel.dk](http://www.dreammodel.dk).

Haldrup, N. (2004): Estimation af middellevetider for mænd og kvinder i Danmark 2002-2100 baseret på Lee-Carter metoden. Arbejdsrapport 2004:3, Velfærdskommissionen.

Hyndman, King & Billah (2002): Local Linear Forecasts Using Cubic Smoothing Splines, Monash University Australia, Working Paper 10/2002.

Keilman, N., D. Q. Pham, & A. Hetland (2001): Norway's Uncertain Demographic Future, *Statistisk sentralbyrå*, Social and Economic Studies.

Lee, R. (2000): The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications, *North American Actuarial Journal*, vol. 4 no. 1, 80-93.

Lee, R., and L. Carter (1992a): Modelling and Forecasting the Time Series of U.S. Mortality, *Journal of the American Statistical Association*, 87, 659-671.

Lee, R., and L. Carter (1992b): Modelling and forecasting US sex differentials in mortality, *International Journal of Forecasting* vol. 8, 393-411.

Lee, R. & T. Miller (2001): Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality, *Demography* vol. 38 no. 4, 537-549.

Oeppen, J. & J. W. Vaupel (2002): Broken Limits to Life Expectancy, *Science* vol. 296 May 2002.

Statistiska centralbyrån (2003): Sveriges framtida befolkning, befolkningsfraskrivning för åren 2003-2050, Demografiske rapporter 2003:4.

Statistiska centralbyrån (2005): SCB's modell för befolkningsprognoser – En dokumentation, Bakgrundsmaterial om demografi, barn och familj 2005:1.

United Nations (2004): World Population in 2300. UN, New York.

Velfærdskommissionen (2006): Fremtidens Velfærd - vores valg, Analyserapport, København.