

## Forventet levetid for forskellige aldersgrupper

### *En anvendelse af Lee-Carter metoden på danske data*

14. juni 2006

Marianne Frank Hansen, Lars Haagen Pedersen og Peter Stephensen

#### 1. Indledning

Igennem det 20. århundrede steg middellevetiden med 24,1 år for mænd og 25,1 år for kvinder. Det svarer til en gennemsnitlig stigning i middellevetiden på omkring 2,9 måneder om året for mænd og 3,0 måneder om året for kvinder. Både danske mænd og kvinder havde ved indgangen til det 21. århundrede en middellevetid, der var steget med omkring 47 pct. i forhold til middellevetiden ved indgangen til det 20. århundrede. Middellevetiden var i 2005 75,5 år for mænd og 80,1 år for kvinder

Set over hele det 20. århundrede er væksten i den danske middellevetid omtrent fulgt med udviklingen i grænsen for middellevetiden målt som væksten i middellevetiden i det land, hvor denne er højest. Over den 160-årige periode, som der findes data for, har der været en bemærkelsesværdig stabil vækst i denne grænse på 3,0 måneder om året for kvinder og 2,7 måneder om året for mænd, jf. Oeppen & Vaupel (2002). Det er en central pointe hos forfatterne, at der ikke er nogen tendens til afbøjning i væksten over tid, hvilket fortolkes som et udtryk for, at der ikke er tegn på, at den befolkningsgruppe, der har den højeste middellevetid, nærmer sig en øvre biologisk grænse.

Både danske kvinder og mænd har ved indgangen til det 21. århundrede en middellevetid, som ligger noget lavere end den højst observerede middellevetid. I 2003 toppede japanske kvinder med en middellevetid på 85,3 år, mens danske kvinders var 79,5 år. Hos mændene toppede Island med en middellevetid på 79,0 år, mens danske mænds middellevetid var 74,9 år, jf. OECD (2005). Der er således grund til at forvente, at væksten i den danske middellevetid kan fastholdes i en lang tidshorisont uden, at der burde opstå tegn på, at en biologisk grænse nås.

Udviklingen i middellevetiden i Danmark i det 20 århundrede kan i store træk opdeles i tre delperioder karakteriseret ved henholdsvis relativ høj vækst frem til starten af 1950'erne, lavere vækst eller stort set ingen vækst i perioden fra midten af 1950'erne til starten af 1990'erne og til sidst en periode, hvor væksten igen er relativ høj. Dateringen af perioderne er lidt forskellig for mænd og kvinder, men det er fælles for de to køn, at der i perioden op til omkring 1995 stort set ikke var vækst i den aldersbetingede restlevetid.

De tre perioder med forskellige vækstrater i den aldersbetingede restlevetid genfindes i en række af de øvrige europæiske lande, herunder de øvrige nordiske lande (bortset fra Finland). Der er dog den forskel, at perioden, hvor der er lav eller ingen vækst i den aldersbetingede restlevetid er markant længere i Danmark end i de øvrige lande, hvor perioden med de højere vækstrater begynder i slutningen af 1970'erne eller i starten af 1980'erne.

De tre perioder med forskellig udvikling i den aldersbetingede restlevetid betyder, at en fremskrivning baseret på den historiske udvikling ikke bliver robust, men tværtimod bliver afhængig af valget af historisk periode, som lægges til grund. Den relative høje vækst i den aldersbetingede restlevetid i perioden efter 1995 kan ikke reproduceres med udgangspunkt i den historiske udvikling, der ligger forud for denne, fordi der i en lang periode forud for 1995 var meget lav vækst i restlevetiden for de fleste grupper i befolkningen. For en række af de øvrige lande, som har en tilsvarende udvikling, er problemet søgt løst ved i høj grad at basere fremskrivningen på data for perioden efter at stigningen i væksten i restlevetiden er indtruffet.

Samme metode anvendes i denne fremskrivning. Imidlertid betyder det senere vendepunkt for væksten i levetiden i Danmark end i de øvrige lande, at det bliver særligt usikkert at skønne over den fremtidige udvikling i Danmark, fordi den dataperiode, der kan lægges til grund bliver meget kort. Der er således en afvejning mellem usikkerheden ved kun at lægge en kort periode til grund og risikoen for, at en fremskrivning baseret på en længere dataserie vil blive domineret af den lange periode med meget lav vækst i restlevetiden. Hvis det sidste er tilfældet vil fremskrivningen udvise en vækstrate, der dels er meget lavere end den, der er observeret gennem de seneste 10 år, og dels føre til, at vækstraten i Danmark såvel på kort som på langt sigt vil ligge et godt stykke under den robuste trend i grænsen for middellevetidsudviklingen.

Baseres fremskrivningen omvendt på en kort dataserie fra omkring 1990 og frem, vil den kortsigtede udvikling i den aldersbetingede restlevetid svare bedre til den, der er observeret i den seneste periode (med høj vækst). I denne fremskrivning vil væksten have en tendens til at aftage, således at den langsigtede vækst svarer til den lange dataserie. Set over fremskrivningsperioden som gennemsnit betyder dette, at også i dette tilfælde vil væksten i Danmark ligge under trenden i grænsen for middellevetidsudviklingen. Ulempen ved denne fremskrivning er, at den anvendte dataserie er så kort, at estimerne på udviklingen i den aldersbetingede dødelighed bliver usikre og udviklingen for de enkelte aldre påvirket af udsving i den seneste periode.

På grund af denne usikkerhed gennemføres to fremskrivninger baseret henholdsvis på data for perioden 1990 – 2005 og på data fra perioden 1965 – 2005. Førstnævnte benævnes hovedfremskrivningen, fordi denne fremskrivnings egenskaber forventes bedst at kunne beskrive den umiddelbare fremtid. Fremskrivningen baseret på data fra 1965-2005 benævnes referencefremskrivningen. Formålet med at vise begge fremskrivninger er at illustrere den betydelige usikkerhed, der er forbundet med fremskrivningen af den aldersbetingede restlevetid. De to fremskrivninger sammenlignes med den fremskrivning af levetidsudviklingen for forskellige aldersgrupper, som Haldrup (2004) udførte for Velfærdskommissionen, jf. Velfærdskommissionen (2004). Denne fremskrivning afviger blandt andet ved at være baseret på data for perioden fra 1904 – 2002.

Såvel nærværende fremskrivning som Velfærdskommissionens fremskrivning anvender den af Lee & Carter (1992) foreslåede metode til fremskrivning af aldersbetingede dødeligheder. Denne metode har i de senere år opnået en betydelig udbredelse, og anvendes af internationale institutioner som f.eks. FN's Population Division, der foretager sammenhængende langsigtede befolkningsfremskrivninger for alle verdens lande, jf. UN (2003). Lee-Carter metoden anvendes herudover til befolkningsfremskrivninger i en række lande, f.eks. Sverige, jf. Statistiska centralbyrån (2003,2005).<sup>1</sup> Herudover anvendes metoden af en række forskergrupper. Den stigende udbredelse har betydet, at der er en ganske omfattende afprøvning af metoden på forskellige lande og generelt findes, at metoden har gode fremskrivningsegenskaber.<sup>2</sup>

Den internationale diskussion om metoder til fremskrivning af befolkningsudviklingen foregår derfor i høj grad inden for rammerne af Lee-Carter metoden, som søges videreudviklet med henblik på at forbedre fremskrivningsegenskaberne i de enkelte lande. Dette papir er i tråd med denne udvikling. I næste afsnit præsenteres hovedtrækene i Lee-Carter metoden og de mest udbredte typer af videreudviklinger af metoden. Dernæst præsenteres udviklingen i dødeligheden og restlevetiden for danske mænd og kvinder i afsnit 3. For hvert af de to køn foretages en sammenligning med udviklingen i de øvrige nordiske lande med henblik på at fastlægge, om der er tale om generelle eller landespecifikke forhold, der ligger til grund for den observerede udvikling i Danmark. Analysen af danske og internationale data anvendes som udgangspunkt for fastlæggelse af datagrundlaget for en fremskrivning af restlevetiden for henholdsvis mænd og kvinder fordelt på alder. Estimationsresultater og metode beskrives i afsnit 4, mens de justeringer i Lee-Carters fremskrivningsmetode som indføres i denne analyse præsenteres i afsnit 5. Den resulterende fremskrivning af restlevetiden fordelt på alder og køn præsenteres i det afsluttende afsnit 6 og sammenlignes med fremskrivningen i Velfærdskommissionen (2004).

Hovedresultaterne i fremskrivningen er, at middellevetiden frem til år 2100 stiger med 13,3 år for mænd og 9,8 år for kvinder. 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 4,7 år mens væksten var på 6,4 år for kvinder. I fremskrivningen fås, at den tilsvarende vækst frem til 2100 er 10,0 år for mænd og 8,1 år for kvinder. Dette afspejler det generelle resultat, at fremtidige reduktioner i dødelighederne forventes at få relativ større effekt på restlevetiden for ældre.

---

<sup>1</sup> Norge anvender andre statistiske estimationsmetoder, som har en mere kompleks struktur end Lee-Carter metoden, jf. Keilman, Pham & Hetland (2001).

<sup>2</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)

## 2. Lee-Carter metoden og udvidelser af denne

Den grundlæggende ide 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. På danske data laves dette opdelt på mænd og kvinder.

For at reducere dimension i problemet antages, at der er en betydelig regularitet til udviklingen af de aldersspecifikke dødeligheder, således at de kan beskrives ved hjælp af 3 elementer: En aldersspecifik del, der er uafhængig af tidspunktet, en tidsafhængig udvikling i et mål for den samlede dødelighed (kaldet mortalitetsindekset) og et (tidsafhængigt) gennemslag af udviklingen i mortalitetsindekset på dødeligheden i den enkelte aldersgruppe. Konkret antages følgende sammenhæng:

$$\ln(m(x,t)) = a(x) + b(x)k(t) + e(x,t)$$

hvor  $x$  er alderen,  $t$  er tiden

$m(x,t)$  er det alders- og tidsafhængige niveau for dødeligheden

$a(x)$  er den aldersafhængige del som i praksis er et gennemsnit over de logaritmiske dødeligheder i estimationsperioden

$k(t)$  er den tidsafhængige udvikling i dødeligheden og

$b(x)$  er gennemslaget af denne på dødelighed for aldersgruppe  $x$ .

$e(x,t)$  er et alders- og tidsafhængigt fejl-led.<sup>3</sup>

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

De udvidelser der i de senere år er lavet af Lee-Carters oprindelige artikel tager udgangspunkt i antagelserne eller de afledte implikationer.

Den første type analyser tager udgangspunkt i, at der antages, at vækstraten i de aldersbetingede dødeligheder er konstante og at forholdet mellem vækstraten i dødelighederne for aldersgrupperne er uændret over tid. Det undersøges derfor om data for det pågældende land har denne egenskab. Det generelle resultat er, at denne beskrivelse er dækkende for ganske lange del-perioder, men at der typisk vil være et eller flere tidspunkter med strukturelle brud, hvor sammenhængen ophører. Herefter følger endnu en periode, hvor antagelserne er opfyldt, men med en anden samlet vækstrate og/eller et andet forhold mellem de aldersbetingede vækstrater. Dette leder til en diskussion af længden af den dataserie, der indgår i estimationen, som le-

<sup>3</sup> For den konkrete estimationsprocedure henvises til f.eks. Haldrup (2004) eller dokumentationsrapporten til dette papir i Hansen (2006).

der frem til fremskrivningen. Der findes for en række lande, at begrænsning af dataserien forbedrer fremskrivningsegenskaberne i hvert fald på det kortere sigt.

Lee & Miller (2001) finder på amerikanske data en tendens til brud udviklingen i dødeligheden og i alderssammensætningen af denne omkring 1950 ligesom i Danmark og en række andre vestlige lande. De analyserer fremskrivningsegenskaberne ved lade modellen fremskrive den historiske periode frem til 1998. De finder, at en dataserie startende i 1950 har klart bedre fremskrivningsegenskaber end en serie startende i 1920.

Carter & Prskawetz (2001) sammenligner for Østrig egenskaberne ved fremskrivninger baseret på henholdsvis perioden fra 1947-1999 og 1976-1999 og at særligt for udviklingen i dødeligheden i de højere aldersgrupper er den korte periode at foretrække fordi der indtræder en ændring i den relative reduktion af dødelighederne mellem de forskellige aldersgrupper imod slutningen af dataperioden.

Booth, Maindonald & Smith (2002) videreudvikler denne ide ved at introducere et test for den optimale datalængde. Udgangspunktet er, at dataperioden skal ende i data for det senest offentliggjorte tidspunkt. Metoden skal derfor udpege startpunktet for den dataserie, der lægges til grund for fremskrivningen. Det kriterium, der anvendes, er, at udviklingen i mortalitetsindekset,  $k(t)$  skal kunne beskrives ved en lineær trend i den historiske periode. Datas afvigelse fra denne trend måles, og hvis reduktion af dataperioden leder til en "substantially smaller" afvigelse vælges at forkorte perioden. Metoden udpeger på australske data for det 20. århundrede et brud i trenden for  $k(t)$  i slutningen af 1940'erne (lidt før 1950) og et omkring 1968. Bruddet i 1950 reducerer faldet i dødelighederne, mens bruddet i 1968 forøger reduktionen i dødelighederne, og svarer således til det brud der ses i danske data i omkring 1995. Analysen viser, at fremskrivningsegenskaberne for Lee-Carter modellen baseret på data fra 1968 dominerer fremskrivninger baseret på en længere periode, dvs. enten fra 1950 eller hele perioden.

Booth, Tickle, & Smith (2005) sammenligner for 10 lande herunder Danmark ovenstående metode med henholdsvis Lee & Millers forslag om at starte dataserien i tidspunktet for det generelle brud omkring 1950 og en Lee-Carter estimation baseret på hele perioden. Ved anvendelse af data frem til 1985 og fremskrivninger på den historiske periode fra 1986-2000. For alle landene udpeger metoden i Booth, Maindonald & Smith (2002) et brud der ligger efter 1950. For Danmark udpeges året 1968 for mænd og 1967 for kvinder. For hovedparten af landene herunder for Danmark findes, at den kortere periode reducerer fejlen på fremskrivningen.

Lundström & Qvist (2004) undersøger udviklingen i svenske dødelighedsdata for det 20. århundrede og finder to større brud i udviklingen både i alderssammensætningen af dødelighedsreduktionen og i den gennemsnitlige størrelse af denne. Første brud dateres til 1951 og andet brud til 1981, jf. nedenfor for en nærmere gennemgang af den svenske udvikling. Baseret på denne analyse har de officielle svenske befolkningsfremskrivninger fra 2003 og frem været baseret på en udvikling i dødelighederne fremskrevet ved Lee-Carter metoden ud fra en dataserie for perioden fra 1990 til seneste data år, jf. Statistiska centralbyrån (2003,2005).

Den næste type analyser indfører flere forklarende variable i beskrivelsen af udviklingen i (logaritmen til) de aldersbetingede dødeligheder. Det mest åbenlyse er at øge dimensionen af leddet  $b(x)k(t)$ , således at udviklingen beskrives ved to eller flere stokastiske processer over tid, som har hver sit gennemslag på de aldersbetingede dødeligheder. Det betyder, at der ikke – som i den oprindelige formulering – nødvendigvis er et fast forhold mellem dødelighedsvækstraten i de enkelte aldre. Det vises i Girosi & King (2005), at dette forbedrer relationens forklaringskraft mærkbart blandt andet for en række skandinaviske lande (Danmark er ikke undersøgt).

En tilsvarende ide er at opdele den samlede aldersbetingede dødelighed efter en række hoveddødsårsager og således opfatte den samlede aldersbetingede dødelighed som summen af disse. For hver hoveddødsårsag beskrives den aldersbetingede dødelighed ved en "standard" Lee-Carter relation. Da de enkelte hoveddødsårsager har en forskellig aldersfordeling indebærer dette, at der heller ikke i dette tilfælde nødvendigvis er et fast forhold mellem dødelighedsvækstraterne i de enkelte aldre. Wilmoth (1995, 1996)

En tredje variant af dette er at opdele de aldersbetingede dødeligheder i overordnede aldersgrupper og fastlægge udviklingen i hvert af de overordnede aldersgrupper ved en standard Lee-Carter relation. En empirisk begrundelse for dette er, at de aldersbetingede dødeligheder har tendens til at udvise en opdeling i tre faser: Spæd- og småbørnsdødelighed op til ca. 10 år, dødelighed for ungdomsgruppen op til 30-40 års alderen og de resterende aldersgrupper. Man kan derved opnå noget af den effekt som kan opnås ved opdeling efter dødsårsager uden at have disse data til rådighed. Lundström & Qvist (2004) anvender således 40 år som en grænse.

En sidste mulighed som diskuteres i litteraturen er at søge at estimere en mere kompleks stokastisk proces for  $k(t)$  som f.eks. tillader en længere lag-struktur.

I nærværende analyse ses alene på den første type af ændringer dvs. på betydningen for estimation og fremskrivning af at opdele den samlede historiske dataserie i delperioder.

### 3. Aldersbetinget dødelighed og restlevetid

Der er en udpræget tendens til, at de aldersbetingede dødeligheder på et givet tidspunkt vokser eksponentielt med alderen. Det svarer til at risikoen for at dø vokser med en fast procent for hvert alderstrin. Fænomenet kaldes Gompertz lov, jf. Gompertz (1825) og har siden offentliggørelsen været anvendt til aktuarisk fastsættelse af f.eks. præmier ved livsforsikring. I Danmark anvendes sammenhængen i det forsikringstekniske grundlag G82<sup>4</sup>, hvor dødsintensiteten beskrives som

---

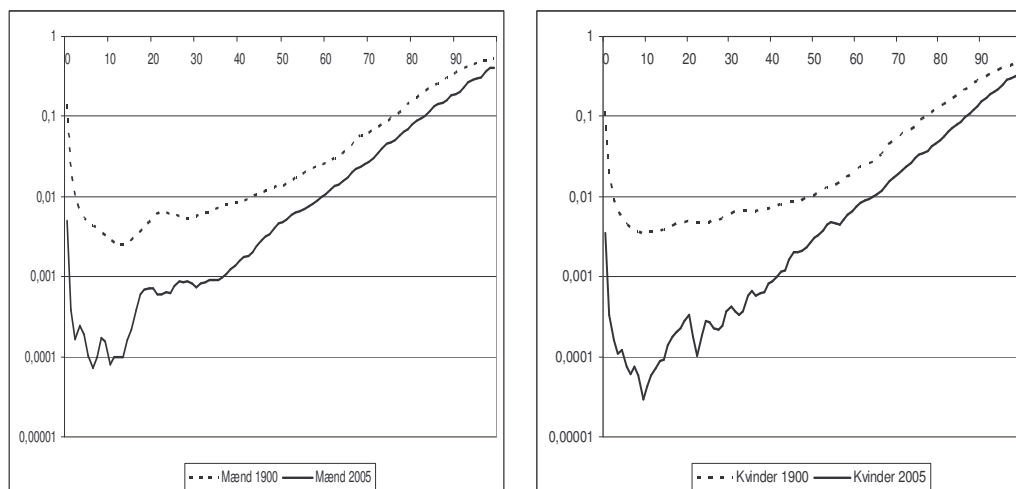
<sup>4</sup> Bemærk, at dødelighederne i G82 kun er aldersafhængige og således ikke tager højde for at den aldersbetingede restlevetid vokser med tiden. De skal derfor løbende opdateres for at tage højde for denne udvikling

$$m(x) = 0,0005 + 10^{-4,12+0,038x} \quad \text{for mænd}$$

$$m(x) = 0,0005 + 10^{-4,272+0,038x} \quad \text{for kvinder}$$

Tendensen, til at dødsrisikoen vokser med en fast procent for hvert alderstrin, gælder i Danmark fra omkring 30-årsalderen og indtil omkring 80 års alderen, jf. Figur 1, hvor dødsrisikoen for de enkelte aldersgrupper er vist med en logaritmisk skala, således at der er tendens til, at disse vokser lineært med 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. I data for 2005 er dødsrisikoen for de omkring 10-årige på niveau med en forlængelse af den (logaritmiske) lineære sammenhæng, mens de øvrige unge aldersgrupper fortsat ligger over.

**Figur 1. Aldersbetinget dødshyppighed for mænd og kvinder 1900 og 2005 (logaritmisk skala)**



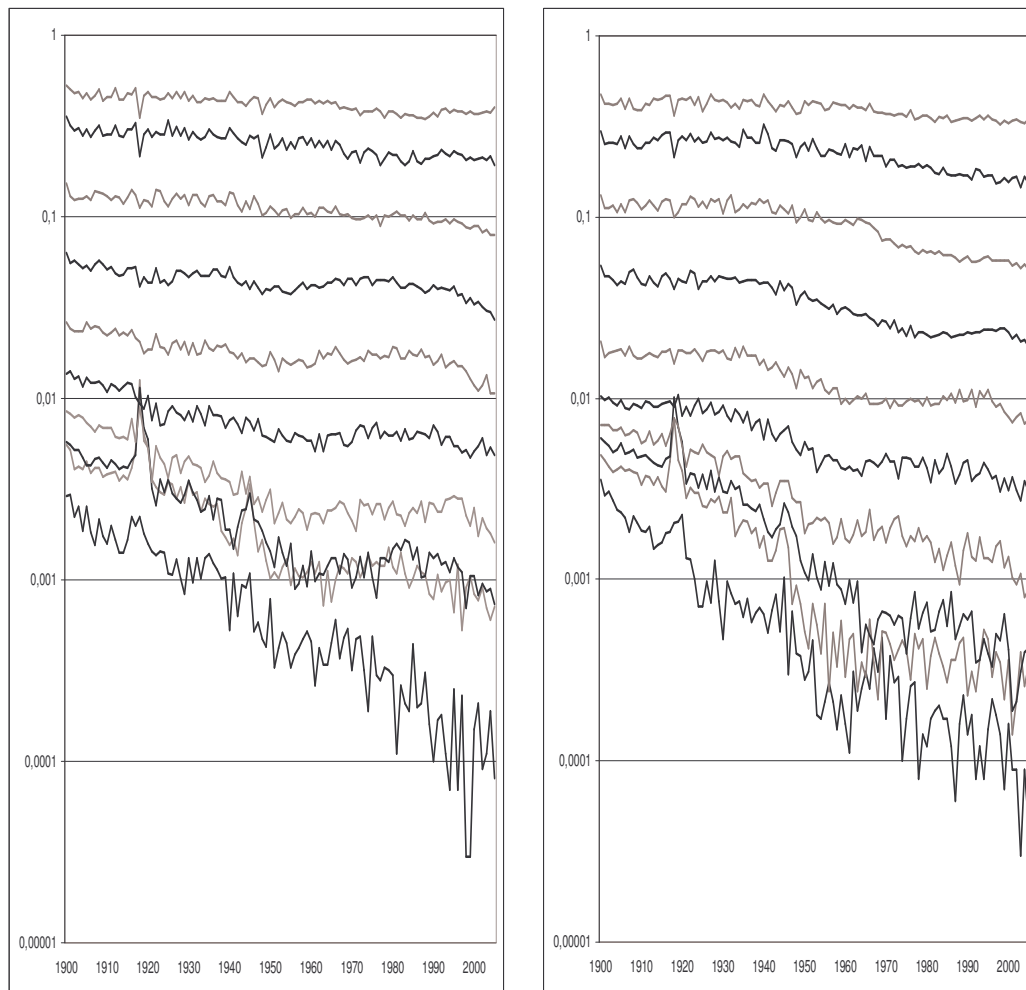
Den historiske udvikling i de aldersbetingede logaritmiske dødeligheder gennem det 20. århundrede har betydet, at logaritmen til dødelighederne er faldet mest for de yngre aldersgrupper (spæd- og småbørnsdødelighed, samt fald i dødeligheden for unge herunder død under barsel for kvinder). Faldet i de logaritmiske dødeligheder 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.

Opdeles den historiske udvikling gennem det 20. århundrede i delperioder, er der imidlertid betydelig forskel på de relative størrelser af faldet i den aldersbetingede dødelighed. Lee-Carter metoden indebærer, at den historiske tendens forlænges således, at den fremskrevne fordeling af faldet i dødelighederne svarer til den historiske udvikling. Det er derfor ikke uden betydning for den fremtidige struktur i dødeligheden, hvilken periode, der lægges til grund for fremskrivningen. Anvendes en lang da-



taserie vil der ved lange fremskrivninger med Lee-Carter være en tendens til, at de yngre aldersgruppers dødelighed kommer under den lineære tendens, der ses i de nuværende data.<sup>5</sup>

**Figur 2. Aldersbetinget døds hyppighed 1900-2005, mænd (venstre) og kvinder (højre) Aldersgrupper 10 år, 20 år, ..., 90 år, 99 år (logaritmisk skala)**



Udviklingen i de aldersbetingede dødeligheder for henholdsvis mænd og kvinder gennem det 20. århundrede er vist for 10-årige, 20-årige, ..., 90-årige og 99-årige i Figur 2. Dødelighederne for 99-årige ligger øverst, herefter følger de øvrige aldre i faldende orden.

<sup>5</sup> Dette har mindre betydning for befolkningsfremskrivningen, fordi de absolutte dødeligheder er begrænsede i udgangspunktet.



Figurene for begge køn viser som en overordnet tendens, at (logaritmen til) dødeligheden er faldende for alle aldersgrupper, og at der er en tydelig tendens over det lange sigt til, at faldet er størst for de yngre aldersgrupper. Der er således langt større spredning i logaritmen til dødelighederne ved indgangen til det 21. århundrede end ved indgangen til det 20. århundrede.

Der er herudover for begge køn en tendens til, at faldet i dødelighederne for alle aldersgrupper på nær 10-årige flader ud omkring 1950 og først for alvor begynder at falde igen omkring 1995. Især for mænd kommer de to skift over en kort periode, jf. Figur 2 (venstre side).

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.

### Udviklingen beskrevet ved restlevetid

Som alternativ til udviklingen i de aldersbetingede dødeligheder kan udviklingen beskrives ved udviklingen i restlevetiden for de enkelte aldersgrupper. Fra en økonomisk-politisk synsvinkel vil dette ofte være mere relevant på grund af den betydelige omfordeling mellem aldersgrupperne, som følger af den sociale kontrakt i velfærdssamfundet, jf. f.eks. Andersen & Pedersen (2005).

Restlevetiden for en given aldersgruppe afhænger af dødeligheden i alle højere aldersgrupper. Middellevetiden (= restlevetiden for en 0-årig) indeholder dermed information på en sammenvejet kompakt form om dødeligheden for alle aldersgrupper. Udviklingen i middellevetiden er derfor måde at udtrykke udviklingen i den samlede dødelighed.

For ethvert år,  $t$ , er restlevetiderne for 1-99-årige beregnet efter nedenstående formel, hvor restlevetiden for en 100-årig er sat til et halvt år og  $x$  angiver alder:

$$\text{restlevetid}(x,t) = m(x,t) \cdot 0,5 + (1 - m(x,t)) \cdot (1 + \text{restlevetid}(x+1,t))$$

Indholdet i denne formel forstås givetvist bedst ved at se på de ekstreme tilfælde, hvor dødelighedsraten  $m(x,t)$  måtte være henholdsvis 0 og 1. Såfremt  $m(x,t) = 1$  ville restlevetiden for den pågældende aldersgruppe være et halvt år, idet det antages, at man dør ligeligt fordelt over året. Er  $m(x,t) = 0$ , kan man se frem til at leve yderligere et år tillagt restlevetiden for en et år ældre aldersgruppe.

Givet at døden indtræder inden for det første leveår, vil det generelt forekomme tidligere end efter et halvt år. Som følge heraf anvendes en speciel og kønsafhængig korrektion for 0-årige. Her anvendes formlen

$$\text{restlevetid}(0,t) = m(0,t) \cdot G + (1 - m(0,t)) \cdot (1 + \text{restlevetid}(1,t))$$

hvor  $G$  er en kønsafhængig konstant, der er mindre end 0,5 og således netop angiver, at spædbørn, der ikke overlever, har tendens til at dø inden for et halvt år af deres første leveår. For mænd er anvendt  $G = 0,111$ , mens størrelsen for kvinder er  $0,112^6$ .

Formlen for beregning af restlevetiden illustrerer tydeligt den ovenfor omtalte afhængighed mellem restlevetid for aldersgruppe  $x$  og dødelighedsrater for aldersgrupperne  $x$  og ældre. Samtidig kan det bemærkes, at udviklingen i restlevetiden i kraft af værende en ikke-lineær transformation af dødelighedsraterne, vil være præget af antagelsen om en eksponentiel udvikling i sidstnævnte.

Bemærk, at en konstant årligt 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, fordi 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(2001). Dette betyder, at Lee-Carter metoden ikke reproducerer den lineære stigning i middellevetiden, som findes i den nævnte reference af Oeppen & Vaupel (2002).

### Restlevetidsudviklingen for mænd

Restlevetidsudviklingen for mænd i aldersgrupperne 0, 10, 20, ..., 90 år gennem de sidste hundrede år fremgår af Figur 3. Da restlevetiden netop er en sammenvæjning af dødelighederne på et givet tidspunkt, vil brud i trenden i dødelighederne på et givet tidspunkt indebære et tilsvarende brud i udviklingen i restlevetiden. Derfor er der i den historiske udvikling i restlevetiden for mænd to tydelige brud svarende til de brud, der blev identificeret i udviklingen i dødelighederne, sammenhold Figur 2 (venstre side) med Figur 3 (venstre side).

Den historiske udvikling i restlevetiden for mænd kan derfor opdeles i 3 adskilte perioder karakteriseret ved forskellige vækstrater for alle aldersgrupper.

- Den første periode fra 1904 til 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. 15 pct. over den 50-årige periode (svarende til 0,3 pct. pr. år). Målt absolut er væksten i restlevetiden derfor aftagende med alderen for aldersgrupperne. Middellevetidsudviklingen skiller sig ud på grund af et betydeligt større fald i spædbørnsdødeligheden end i dødeligheden for de øvrige aldersgrupper.

---

<sup>6</sup> Disse tal er i overensstemmelse med Danmarks Statistiks, som disse er angivet i 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.

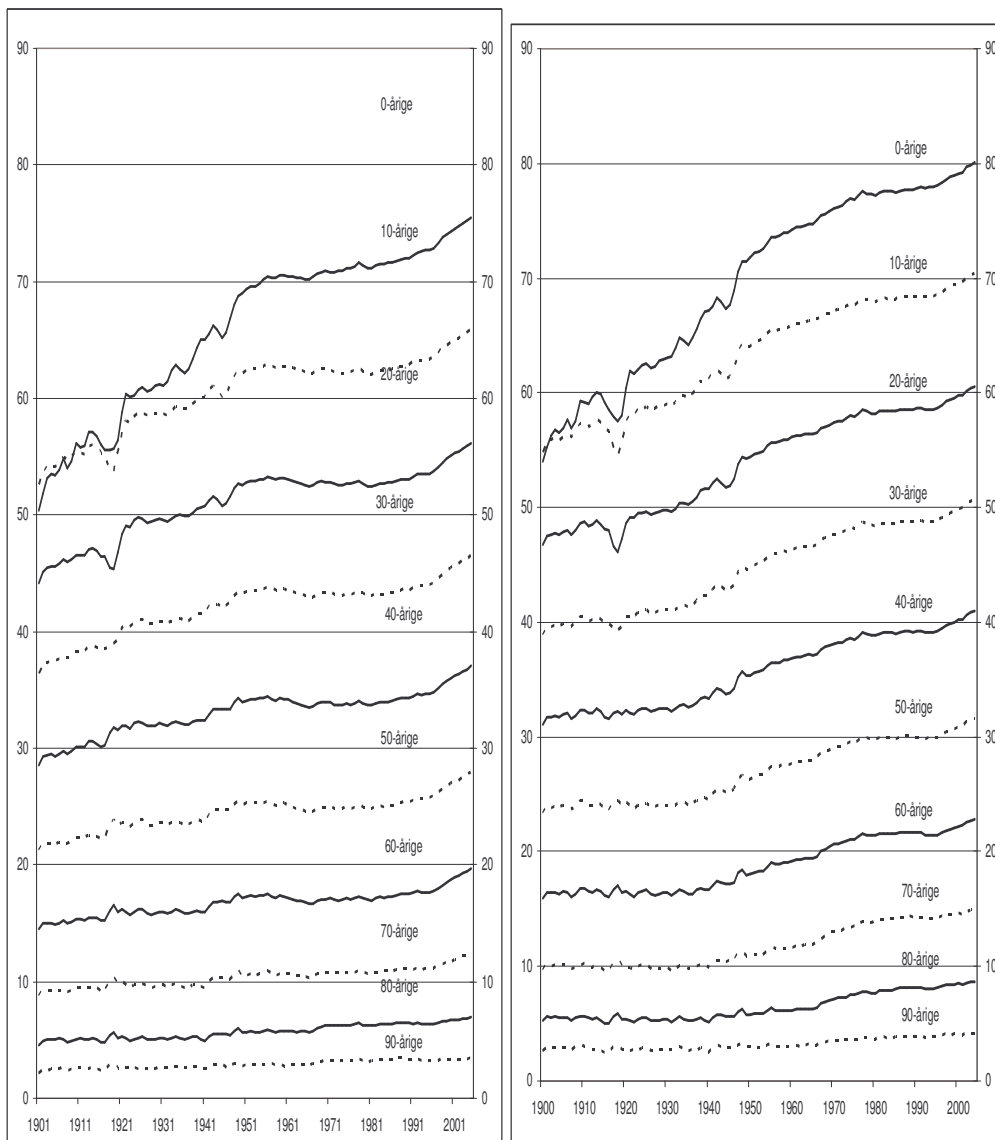
per. Stigningen i middellevetiden er derfor omkring 30 pct. i perioden fra 1904 til 1953, eller omtrent dobbelt så høj en vækst som de øvrige aldersgrupper. Udviklingen svarer til en gennemsnitlig årlig stigning i middellevetiden på 0,5 pct. Det bemærkes, at der for de yngre aldersgrupper er et betydeligt knæk i væksten omkring 1918. Det er konsekvenserne af den spanske syge i 1918, der er årsagen til dette.

- 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 70 år. Vækstraten for den 40-årige periode ligger samlet på omkring 1,5 pct. for disse aldersgrupper (svarende til en gennemsnitlig årlig vækst på 0,04 pct.). Selvom der også er en meget betydelig opbremsning i væksten i middellevetiden er vækstraten i denne på omkring 5 pct. over den godt 40-årige periode (svarende til en gennemsnitlig årlig vækst på 0,1 pct.). For de ældste aldersgrupper er faldet i vækstraten mindre og ligger f.eks. for 80-årige på 8 pct. over perioden
- Den tredje og sidste periode er den nuværende og strækker sig foreløbigt fra 1995 til 2005. I denne periode er restlevetiden for alle aldersgrupper (bortset fra middellevetiden) vokset mere end i den foregående periode selvom den foregående periode strækker sig over en årrække, som er ca. 4 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 3,8 pct. siden 1995, restlevetiden for 20-årige mænd med 4,8 pct., restlevetiden for 50-årige mænd med 8,5 pct. og restlevetiden for 60-årige mænd med 11,2 pct. Forskellene i vækstraten betyder, at der kun er begrænsede forskelle i den absolutte vækst for de forskellige aldersgrupper. Således er restlevetiden for en 10-årig dreng vokset med 2,6 år over perioden, mens restlevetiden for en 60-årig mand er vokset med 2,0 år.

### **International sammenligning af udviklingen for mænd**

Den bemærkelsesværdige udvikling, hvor der er en betydelig reduktion i væksten i middellevetiden omkring 1950 er ikke enestående for Danmark. Denne tendens genfindes i en række af de vestlige lande som på dette tidspunkt har en høj middellevetid, jf. Figur 4, der viser udviklingen for de nordiske lande. For både Danmark, Norge, Sverige og Island ses reduktionen i middellevetidsvæksten omkring 1950. For Norge, Sverige og Island er der en tendens til øget vækst igen fra omkring 1980. For Danmark indtræder denne tendens først omkring 1995. De samme tendenser gør sig gældende for restlevetiden for 60-årige mænd, jf. Figur 5.

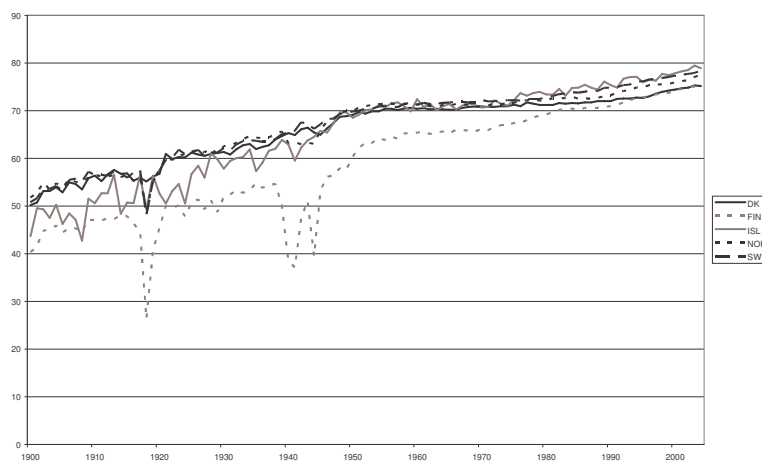
**Figur 3. Restlevetidsudviklingen for udvalgte aldersgrupper mænd (venstre) og kvinder (højre).**



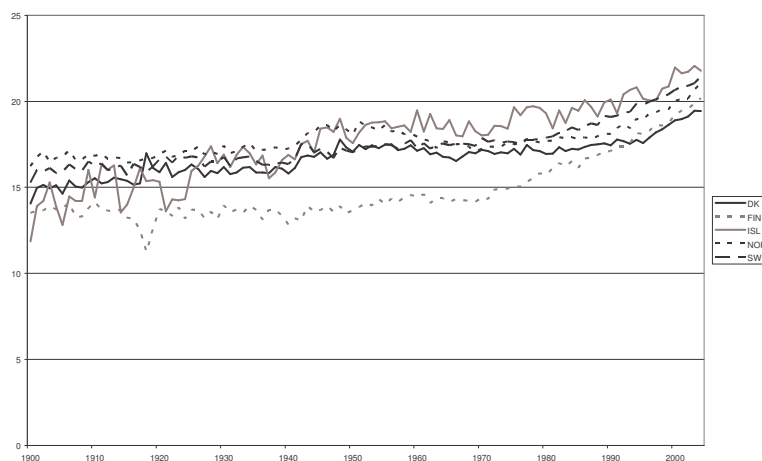
I det følgende sammenlignes den danske udvikling med den svenske, jf. Lundstöm og Qvist (2004). Udviklingen i den svenske restlevetid for mænd falder ligesom den danske i 3 tydeligt afgrænsede delperioder. Lundstöm og Qvist daterer den første – hvor der er en betydelig vækst i restlevetiden for alle aldersgrupper – fra århundredeskiftet til 1951. Altså stort set samme periode som i Danmark. Fordelingen af væksten fordelt på aldersgrupper svarer også til den danske, således at vækstraten er størst for middellevetiden, mens der er en lavere, men omtrent ens vækstrate for aldersgrupperne fra 10 år og derover.

Denne periode efterfølges ligesom i Danmark af en lavvækstperiode, som i det svenske tilfælde strækker sig frem til 1981. Det er karakteristisk for de fleste vestlige lande at lavvækstperioden er kortere end i Danmark og typisk ophører i sidste halvdel af 1970'erne eller i starten af 1980'erne. Ligesom i Danmark er den svenske udvikling i denne lavvækstperiode endvidere karakteriseret ved, at spædbørnsdødeligheden fortsat falder mere end de øvrige dødeligheder, og at opbremsningen i væksten i middellevetiden derfor er lavere end faldet i væksten i restlevetiden for de øvrige aldersgrupper.

**Figur 4. Middellevetidsudviklingen for mænd i de nordiske lande, 1900 -2005**



**Figur 5. Restlevetidsudviklingen for 60-årige mænd i de nordiske lande, 1900 -2005**



Perioden fra 1981 svarer til den periode, der i Danmark starter i 1995, og er således karakteriseret ved, at vækstraten er betydeligt højere end i den foregående periode,

og at vækstraten er højere for de ældre aldersgrupper end for de yngre. Sverige har således oplevet en højvækstperiode i restlevetiden, der foreløbig har varet i omkring 25 år, mens den tilsvarende danske endnu kun er på ca. 10 år.

På basis af de internationale sammenligninger er der således 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 forsat, 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.

Disse tre ret klart afgrænsede perioder i udviklingen i restlevetiden for mænd giver anledning til, at der kan anlægges forskellige strategier ved fremskrivningen af den forventede udvikling i restlevetiden, jf. nedenfor.

### **Restlevetidsudviklingen for kvinder**

Svarende til beskrivelsen for mænd fremgår restlevetidsudviklingen for kvinder i aldersgrupperne 0, 10, 20, ..., 90 år gennem de sidste hundrede år af Figur 3 (højre side). Som beskrevet ovenfor kan udviklingen i dødelighederne for kvinder gennem det 20. århundrede opdeles i 4 adskilte perioder, som ikke er helt så markante som i mændenes tilfælde. De 4 perioder genfindes (på samme måde som hos mændene) i udviklingen i den aldersbetingede restlevetid for kvinder.

- Den første periode varer fra 1904 - 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 vækst i middellevetiden. Vækstraten i middellevetiden for kvinder over perioden fra 1904 til 1936 er på 11 pct., mens den tilsvarende vækstrate i restlevetiden for 10-årige piger er 7 pct., for 30-årige kvinder er den 5 pct. og for 50-årige er den 2 pct. For de alder ældste aldersgrupper er der tale om et fald i restlevetiden gennem perioden. Således reduceres restlevetiden med 1 pct. for 70-årige kvinder og med 4 pct. for 80-årige kvinder.
- I den anden 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 dog med undtagelse af middellevetiden, hvor vækstraten er høj som følge af forsat faldende spæd- og småbørnsdødelighed. I perioden fra 1936 til 1978 vokser middellevetiden for kvinder med 20 pct., mens restlevetiden for 10-årige piger vokser med 14 pct., restlevetiden for 30-årige kvinder med 17 pct., for 60-årige kvinder med 32 pct. og for 80-årige med 45 pct.
- Tredje periode – 0-vækstperioden – er fra 1978 til 1995. I denne perioden er der stort set ikke vækst i restlevetiden for kvinder i nogen aldersgrupper.

- Fjerde og sidste periode er den nuværende der foreløbig dækker perioden fra 1995 til 2005. I denne periode er der atter en betydelig vækst i restlevetiden for kvinder i alle aldersgrupper. Som hos mændene 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 gældende for aldersgruppen på 60 år og op efter, som har forholdsvis ensartede vækstrater. Middellevetiden for kvinder vokser 2,8 pct. fra 1995 til 2005, mens restlevetiden for 10-årige piger vokser 2,9 pct. i samme periode og det tilsvarende tal for 30-årige er 3,9 pct. For 60-årige vokser restlevetiden med 6,5 pct. i perioden og noget tilsvarende gør sig gældende for de ældre aldersgrupper.

### **International sammenligning af udviklingen for kvinder**

Udviklingen i restlevetiden for danske kvinders adskiller sig fra de fleste andre lande ved at der er periode fra 1978 til 1995, hvor væksten i restlevetiden stort set går i stå. Bortset fra denne periode er der et betydeligt sammenfald mellem udviklingen i Danmark og de øvrige nordiske lande. Undtagelsen er Finland, som gennem hele det 20. århundrede oplever en større stigning i middellevetiden for kvinder end de øvrige nordiske lande og ved indgangen til det 21. århundrede har overhalet Danmark og opnået et niveau der svarer til de øvrige nordiske lande, jf. Figur 6.

Den danske og den svenske udvikling i restlevetiden for 60-årige kvinder har store lighedstræk gennem det 20. århundrede. De to eneste større forskelle er at stigningen i vækstraten sætter ind i midten af 1930'erne i Danmark, mens det først sker i starten af 1940'erne i Sverige. Denne forskel betyder at restlevetiden i Danmark når det svenske niveau omkring 1940. Udviklingen i de to lande er derefter helt parallel frem til 1978, hvor væksten ophører i Danmark men forsætter i Sverige. Omkring 1995 vender væksten tilbage i Danmark, men i den mellemliggende periode er forskellen i restlevetiden vokset til 2,7 år, Figur 7.

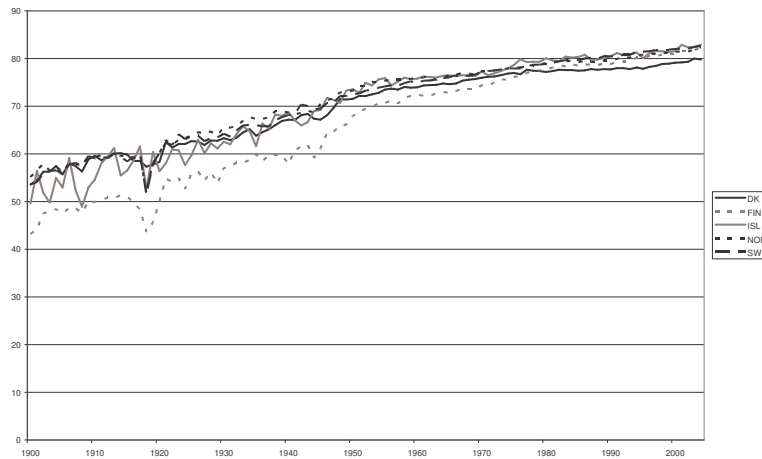
For Island og Norge, hvor restlevetiden er højest, er der en tendens til at væksten aftager i perioden fra 1950 til 1970 således, at udviklingen minder om den, der ses for alle de nordiske lande for mænd. Fra 1970 stiger væksten igen, og der er stort set sammenfald i restlevetidsudviklingen i Norge, Sverige, og Island igennem de sidste 35 år. Finland har gennem hele perioden en højere vækst men starter på et lavere niveau. Ved udgangen af århundredet har den finske restlevetid samme niveau som de tre øvrige nordiske lande – og Danmark ligger som det eneste nordiske land under de øvrige, jf. Figur 7.

Sammensætningen af væksten i middellevetiden for kvinder er stort set identisk i Danmark og Sverige i første halvdel af det 20. århundrede. Væksten i middellevetiden skyldes i denne periode i høj grad reduktion i spæd- og småbørnsdødeligheden. Der er kun en begrænset vækst i restlevetiden for voksne kvinder i Sverige frem til omkring 1940, hvorefter der er vækst i restlevetiden for alle aldersgrupper. Væksten forsætter med uændret styrke frem til de seneste år, hvor der synes at være en tendens til at vækstraten aftager, jf. Lundström og Qvist (2004).

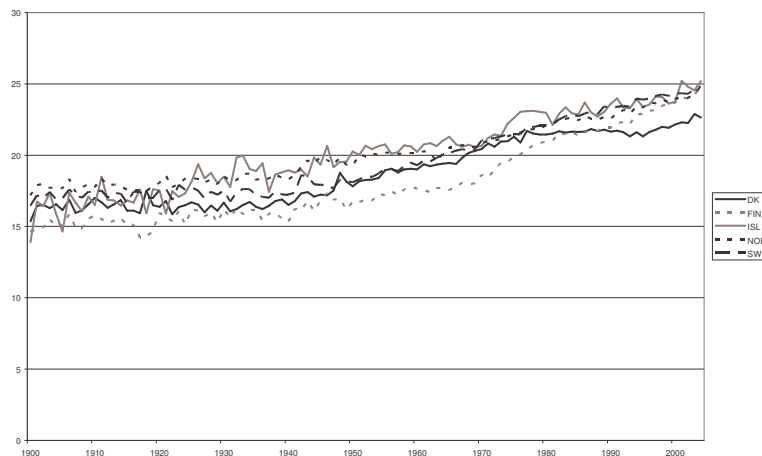


Internationale sammenligninger peger derfor på, at opbremsningen i væksten i levetiden for kvinder i 1970'erne, 80'erne og begyndelsen af 1990'erne er et særskilt dansk fænomen, og at den seneste periodes udvikling tyder på, at middellevetidens udvikling igen svarer til udviklingen i de omkringliggende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

**Figur 6. Middellevetidens udvikling for kvinder i de nordiske lande, 1900 - 2005**



**Figur 7. Udviklingen i restlevetiden for 60-årige kvinder i de nordiske lande 1900 - 2005**



#### 4. Estimation af Lee-Carter modellen

##### Data

Data for de årlige dødeligheder for aldersgrupper og køn offentliggøres årligt af Danmarks Statistik. Senest offentliggjorte tal dateres 2004/2005. Dateringen skyldes,

at Danmarks Statistik til offentliggørelsesformål anvender data, der er "udglattet" ved anvendelse af to år og to aldersgrupper i hvert år.<sup>7</sup> DREAMs database indeholder (ikke-udglattede) data for hver aldersgruppe og hvert år for perioden 1981-2005. Til analyse af den historiske udvikling og undersøgelse af estimationsresultaternes afhængighed af dataperioden er der imidlertid brug for en længere dataperiode. Disse er tilgængelige i en stor international database: Human Mortality Database (HMD). Databasen vedligeholdes af forskere ved University of California, Berkely og Max Planck Institute for Demographic Research og er tilgængelig via Internettet.<sup>8</sup> Data for Danmark foreligger her for hvert år og hver aldersgruppe i perioden 1835-2004.<sup>9</sup>

Til nærværende formål er anvendt variabelen  $q(x)$  fra de periodiske dødelighedstavler i HMD. Nyeste tilgængelige data er her fra 2004, men for at inddrage den nyeste tilgængelige information konstrueres et skøn for dødshyppighederne for 2005 ved anvendelse af data fra DREAMs befolkningsdatabase som gøres sammenlignelige med HMD data. Disse dødeligheder føjes til HMD data, så der for såvel mænd som kvinder slutteligt fremstår data frem til 2005 til brug for fremskrivningen.

Den tekniske estimation foregår på disse data. Den endelige rapportering af resultater i det foregående og det efterfølgende kapitel er imidlertid korrigeret og "udglattet", således at de er umiddelbart sammenlignelige med de af Danmarks Statistik offentliggjorte værdier med den konvention, at dødeligheder som Danmarks Statistik benævner 2004/2005, kaldes 2005 – svarende til Danmarks Statistiks egen konvention i forbindelse med befolkningsfremskrivning.

### Estimation af modellens parametre og mortalitetsindekset

Som tidligere nævnt tager nærværende estimation og fremskrivning af dødsrater udgangspunkt i Lee-Carters oprindelige fremstilling Lee-Carter (1992) i hvilken de logaritmiske dødsrater antages at kunne beskrives ved sammenhængen

$$\ln(m(x,t)) = a(x) + b(x)k(t) + \varepsilon(x,t)$$

hvor  $m(x,t)$  angiver dødelighedsraten på tidspunkt  $t$  for aldersgruppe  $x$ .

Parameteren  $a(x)$  angiver som tidligere antydte den grundliggende form på aldersprofilen over tid og beregnes i praksis som gennemsnittet over tid af de logaritmiske dødsrater. Den aldersafhængige parameter  $b(x)$  angiver afvigelser fra den generelle profil over tid. Centralt for fremskrivningen er den aldersuafhængige parameter  $k(t)$ , der beskriver udviklingen over tid i den generelle dødelighed og benævnes derfor

<sup>7</sup> Årsagen til denne praksis er ifølge Danmarks Statistik historisk betinget og opretholdes for at sikre sammenlignelighed med historiske data. Der er tilsyneladende ikke en tilsvarende praksis i andre lande. Der er derfor en bias i danske middelevetidsopgørelser i forhold til opgørelser i andre lande. Estimationerne i Haldrup (2004) baseres på data på offentliggørelsesniveau.

<sup>8</sup> [www.mortality.org](http://www.mortality.org) som også indeholder dokumentation metoden til konstruktion af dataserierne.

<sup>9</sup> Data på enkelt år findes for perioden fra 1963. For den foregående periode er der foretaget en (avanceret) beregning af data i årene mellem to tællinger.

mortalitetsindekset. Når det fælles niveau ændres over tid indikerer parameteren  $b(x)$  således effekten på de aldersafhængige logaritmiske dødelighedsrater. I det indekset vil være aftagende over tid, indikerer en positiv indgang i vektoren  $b(x)$ , at dødeligheden for den pågældende aldersgruppe er aftagende, mens en negativ indgang indikerer det modsatte. Aftager mortalitetsindekset lineært, dvs.  $dk(t)/dt$  er konstant vil de logaritmiske dødelighedsrater også aftage lineært over tid med proportionalitetsfaktoren  $b(x)$ . Dette medfører, at dødelighedsraterne er eksponentielt aftagende.

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

$$\sum_{t=t_1}^{t_n} k(t) = 0$$

$$\sum_x b(x) = 1$$

Indledningsvist estimeres parameteren  $a(x)$  som gennemsnittet over de logaritmiske dødelighedsrater i dataperioden, dvs.

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

Efterfølgende beregnes en matrix bestående af de centrerede logaritmiske rater som

$$\ln(m(x,t)) - \hat{a}(x)$$

hvorefter denne påføres en Singular Value Decomposition fra hvilken estimerer for henholdsvis  $b(x)$  og  $k(t)$  kan udtrages.

Således er fundet et initialt bud på modellens parametre. Der er imidlertid ikke er nogen garanti for, at denne estimation resulterer i, at der for et givet år opnås, at det antal døde der kan beregnes fra estimationen svarer til det faktiske samlede antal døde. Derfor indeholder den oprindelige Lee-Carter formuleringe 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 den faktiske levealder. 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 den mest relevante målvariabel.

Ved anvendelse af de initiale parameterestimer for  $a(x)$  og  $b(x)$  bestemmes mortalitetsindekset ved anvendelse af en Newton-Raphson algoritme således, at forventet levealder i hele dataperioden opnås. Som det er 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  $a(x)$  og  $b(x)$ . Dette kan opfattes som første trin i en iterativ proces, der leder til værdier af  $a(x)$  og  $b(x)$ , som er indbyrds konsistente med  $k(t)$ . I denne analyse suppleres dette med andet trin i den iterative proces, idet der med de opdaterede værdier for  $a(x)$  og  $b(x)$  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. Parametrene i Lee-Carter modellen er nu fastlagt og dødelighedsrater kan estimeres inden for dataperioden.

Den centrale variabel for fremskrivningen af udviklingen i levetiden er udviklingen i mortalitetsindekset,  $k(t)$  over tid. Det er derfor centralt, at processen beskrivende udviklingen i dette indeks fastlægges og er statistisk velspecificeret. Denne fastlæggelse baseres på statistiske test og suppleres med grafisk inspektion af serien.

### Afgrænsning af dataperioden

Gennemgangen af den historiske udvikling i 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.

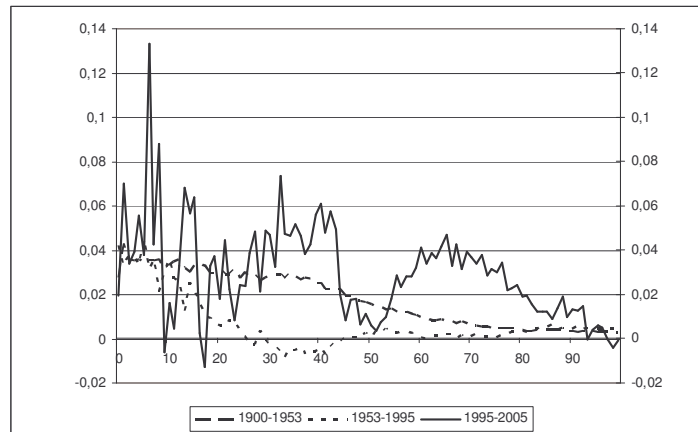
Lee-Carter metoden fører til, at fremskrivningen vil have tendens til at forlænge det gennemsnitlige observerede årlige fald i (logaritmen til) dødelighederne igennem estimationsperioden til fremskrivningsperioden. Samtidig vil alderssammensætningen af faldet i dødelighederne også have tendens til at afspejle den gennemsnitlige aldersfordeling af dødelighedsfaldet.

For Danmark identificeredes 3 delperioder for mænd og 4 del perioder for kvinder, jf. afsnit 2. Svarende til analyserne i den i afsnit 2 refererede litteratur er Lee-Carter modellen blevet estimeret for hver af disse 7 delperioder.

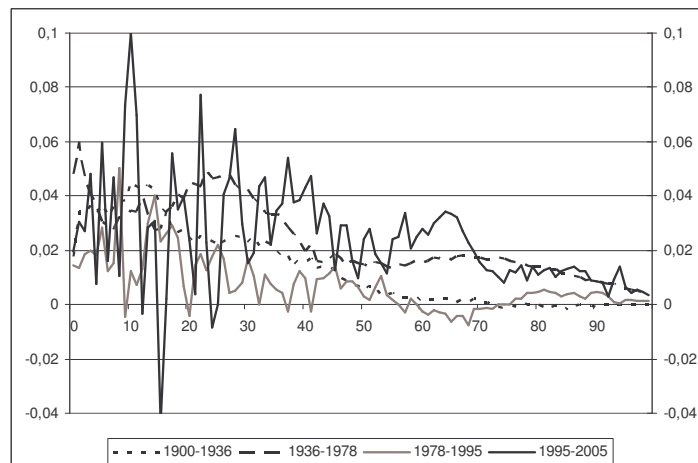
Resultaterne af estimationerne af  $k(t)$  og  $b(x)$  er anvendt til at beregne det estimerede årlige procentvise fald i dødelighederne de enkelte aldre, jf.

Figur 8 og Figur 9.<sup>10</sup>

**Figur 8. Estimerede årlige procentvise fald i de aldersbetingede dødeligheder, mænd**



**Figur 9. Estimerede årlige procentvise fald i de aldersbetingede dødeligheder, kvinder**



Det årlige fald i de aldersafhængige dødeligheder for mænd er meget afhængigt af hvilken af de 3 delperioder, der lægges til grund. Lægges perioden fra 1900-1953 til grund, ses, at den procentvise reduktion i dødeligheden er aftagende med alderen,

<sup>10</sup> Det procentvise årlige fald fremkommer som  $b(x) \Delta k(t)$ . Da  $k(t)$  er en  $I(1)$  proces vil fremskrivningen lede til et lineært fald i  $k(t)$  efter nogle perioder.  $\Delta k(t)$  er derfor konstant og man kan se bort fra tidsafhængigheden i det årlige fald.

idet den årlige reduktion i spædbørnsdødeligheden er 4 pct., mens reduktionen i dødeligheden for en 45-årig er 2 pct. og reduktionen herefter gradvist reduceres til 0 for de ældste aldre.

Lægges perioden 1953-1995 til grund fås også årlige fald i spædbørnsdødeligheden på ca. 4 pct., men for de øvrige aldersgrupper er der væsentlig mindre fald. For mænd over 25 år er der intet fald.

Endelig ses en estimation på basis af perioden fra 1995 at lede til de største procentvise fald i de aldersbetingede dødeligheder, men samtidig er det klart, at den korte periode leder til betydelige fluktuationer for en aldersgruppe til den næste. Dette er en indikation af, at dataperioden er for kort til, at der kan opnås stabile aldersbetingede estimater, fordi disse bliver påvirkelige af tilfælde udsving i enkelte år. I denne forbindelse er særligt de systematisk lave reduktioner i dødelighederne for aldersgrupperne fra ca. 45 til 55 år et betydeligt problem for strukturen i dødeligheden i fremskrivningsperioden.

Det er dog på trods af usikkerheden vedrørende de 45-55 årige forholdsvis klart, at aldersstrukturen i dødelighedsfaldet er ændret, når denne periode lægges til grund. De største procentvise fald opnås for gruppen i alderen 30 til 40 år, hvor gennemsnittet er på ca. 5 pct. Imidlertid er det procentvise fald for gruppen fra 60 til 80 år med et årligt niveau på 3 – 4 pct. også betydeligt højere end det ses ved estimation baseret på de øvrige perioder. Det sidste er afgørende, fordi dødeligheden for disse aldersgrupper er høj sammenlignet de unge aldersgrupper.

Analysen af den estimerede effekt på faldet i de aldersafhængige dødeligheder for mænd viser således klart, at estimaterne af  $b(x)$  og  $k(t)$  ikke er robuste over de tre dataperioder. Valg af dataperiode vil derfor have afgørende effekt på fremskrivningen.

De samme tendenser til forskelle i både det gennemsnitlige niveau for det procentvise fald i dødelighederne og aldersfordelingen af disse findes for kvinder. Forskellene er dog generelt mindre end for mændene. Også i dette tilfælde ses, at anvendelse af korte tidsperioder for estimationen giver anledning til, at estimaterne for de enkelte aldre svinger meget betydeligt og derved indikerer, at den korte periode giver anledning til betydelig usikkerhed. Samtidig gælder også i dette tilfælde, at estimationen baseret på den seneste periode giver anledning til de største fald, og at dette særligt er markant for aldersgruppen fra 30 til 70 år, hvor de årlige fald er på mellem 5 og 3 pct. Som hos mændene er estimaterne baseret på data for perioden forud for den sidste karakteriseret ved stort set at føre til uændrede aldersbetingede dødeligheder.

Også for kvinder er konklusionen derfor, at estimaterne af  $b(x)$  og  $k(t)$  ikke er robuste over de fire dataperioder, og at dataperiodens længde derfor må forventes at have betydelig effekt på estimaterne.

Tilsvarende resultater har i en række af de øvrige vestlige lande, jf. referencerne i afsnit 2 ledt til, at estimationen alene baseres på data fra den seneste periode, hvor estimaterne er robuste, hvilket svarer til, at der kun anvendes data fra efter det seneste strukturelle brud. For Danmark indebærer dette, som vist ovenfor, at estimaterne ikke bliver robuste, fordi den periode der lægges til grund er for kort. På grund af det

strukturelle skift i udviklingen i dødelighederne i 1995 vil udvidelse af perioden føre til, at parameterestimerne for  $b(x)\Delta k(t)$  vil falde systematisk for hvert år perioden udvides med. Dette giver anledning til, at der er meget betydelig usikkerhed på fremskrivningerne af den forventede levetid.

Baseres fremskrivningen på den seneste periode, dvs. fra 1995 og frem, vil det indebære, at den vil have tendens til at forlænge den seneste udvikling. Da data indikerer, at strukturen i dødelighedsudviklingen fastholdes i længere perioder, er der grund til at forvente, at en fremskrivning baseret på denne dataserier vil have de bedste fremskrivningsegenskaber på kortere sigt.

Årsagen, til at det kan være relevant at inddrage længere dataperiode, er både den betydelige usikkerhed med hensyn til estimerne og hensynet til det længere sigt. Som vist er den seneste periode 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 endelig valg af længden af dataserien er derfor et kompromis mellem den under vurdering 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.

Der er derfor gennemført analyser, hvor estimationen er baseret på dataperioder af forskellig længde. Resultatet af disse er, at der kun er begrænset effekt på estimerne af at forlænge dataserien ud over perioden fra 1965-2005. Estimer baseret på denne dataserie anvendes derfor som reference, og der laves en fremskrivning baseret på denne. Fremskrivningen indebærer, at der fra første fremskrivningsår er en markant lavere vækst i middellevetiden end den, der er blevet observeret i de seneste ti år.

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 estimer er det nødvendigt at udvide antallet af observationer. Valget af dataserie er efter denne procedure faldet på perioden fra 1990-2005. En fremskrivning baseret på denne tidsserie har for både mænd og kvinder den egenskab, at den langsigtede vækstrate i middellevetiden konvergerer mod vækstraten for fremskrivningen baseret på perioden fra 1965-2005, og at de to vækstrater er ens inden udgangen af det 21. århundrede. På det korte sigt indebærer denne fremskrivning en vis opbremsning i væksten i middellevetiden i forhold til den historiske trend.

### Estimationsresultater

I det følgende gennemgås estimationsresultaterne for både dataserien 1990-2005, der anvendes til fremskrivningen for både mænd og kvinder og den langsigtede reference dataserie fra 1965-2005 også for både mænd og kvinder.



Nedenstående modeller for indekset for dødelighederne,  $k(t)$  findes efter test at være velspecificerede for de fire fremskrivninger af  $k(t)$ . Indeks frembragt på baggrund af data fra 1990-2005 findes at kunne beskrives ved følgende, hvor spredningen på de estimerede koefficienter er angivet i parentes

$$\text{Mænd:} \quad \Delta k(t) = -2,4820_{(0,6087)}$$

$$\text{Kvinder:} \quad \Delta k(t) = -4,5374_{(1,3046)} - 0,6735_{(0,2611)} \Delta k(t-1)$$

For fremskrivningen baseret på data fra 1965-2005 følger mortalitetsindekset for såvel mænd som kvinder en  $I(1)$  proces med en trend.

$$\text{Mænd:} \quad \Delta k(t) = -1,9951_{(0,4964)} - 0,3150_{(0,1805)} \Delta k(t-1)$$

$$\text{Kvinder:} \quad \Delta k(t) = -2,7285_{(0,6079)} - 0,4840_{(0,1553)} \Delta k(t-1)$$

Driften eller konstantleddet i ovenstående sammenhænge vil efter få perioder dominere således, at det fremskrevne indeks  $k(t)$  på sigt vil følge et lineært trendforløb. Sammen med de estimerede værdier af gennemslaget på de aldersbetingede dødeligheder,  $b(x)$  af et givet fald i dødelighedsindekset,  $k(t)$  kan dette anvendes til at finde de aldersbetingede procentvise reduktioner i dødelighederne i de to fremskrivninger. jf. Figur 10 og Figur 11.

For mænd indebærer en fremskrivning baseret på perioden 1990-2005, at den årlige reduktion i dødeligheden bliver større for alle aldersgrupper bortset fra unge i visse aldre mellem 10 og 20 år end, hvis fremskrivningen baseres på data for perioden 1965-2005.

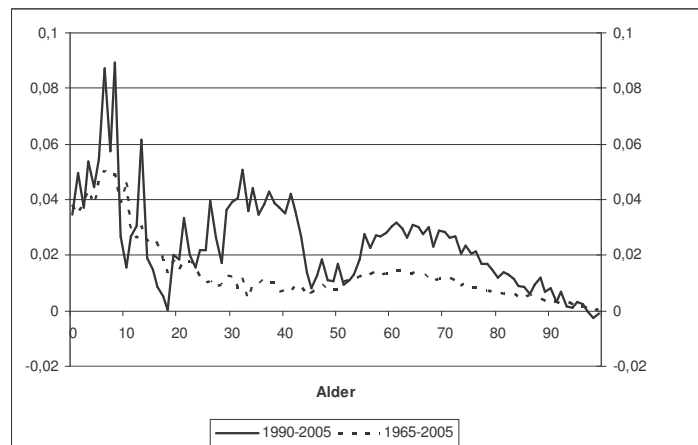
Fremskrivningen baseret på perioden 1965-2005 indebærer årlige fald i dødeligheden på omkring 4-5 pct. for børn faldende til omkring 2 pct. for 20-årige og 1 pct. for mænd i alderen 25-75 år. For de ældste er det årlige fald lidt mindre.

De tilsvarende fald i de aldersbetingede dødeligheder baseret på data for 1990-2005 afviger især for aldersgruppen 25-40 år, hvor det årlige fald i fremskrivningen bliver på omkring 4 pct. For aldersgruppen fra 55-70 år bliver det årlige fald på omkring 3 pct. mens faldet for de ældre aldersgrupper gradvist reduceres fra dette niveau til 0 pct. for personer over 95 år.

De aldersbetingede årlige reduktioner i dødeligheden er for det første mere usikre når den korte dataperiode lægges til grund. Dette fremgår af kurvens mere takkede struktur. For det andet og vigtigere er tendensen til, at de procentvise årlige fald er størst for de laveste aldre og gradvist faldende med alderen, ikke systematisk til stede med den korte dataserie. Markante undtagelser er aldersgrupperne fra omkring

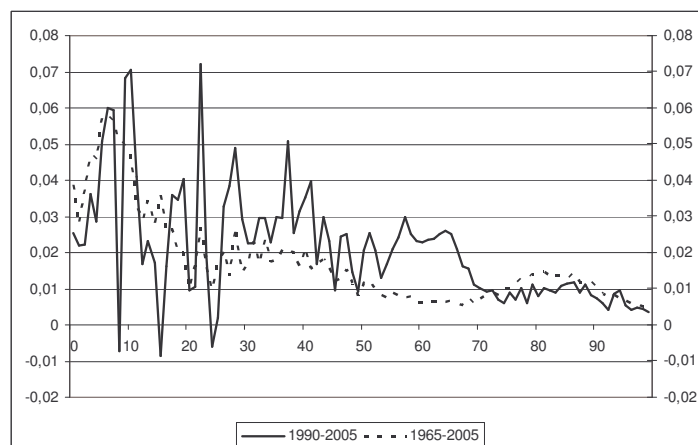
17 til 32 år og fra 52 til 60 år, hvor reduktionen i dødeligheden er voksende med alderen.

**Figur 10. Estimerede årlige procentvise fald i de aldersbetingede dødeligheder, mænd**



Det sidste punkt er kritisk, fordi fremskrivningen over tid vil afvige stadigt mere markant fra Gompertz lov om, at dødelighederne vokser eksponentielt over med alderen. For de to nævnte aldersgrupper vil dødelighederne over tid få tendens til at være aftagende med alderen.

**Figur 11. Estimerede årlige procentvise fald i de aldersbetingede dødeligheder, kvinder**



For kvinder er forskellen mellem de to sæt estimerede reduktioner i de aldersbetingede dødeligheder mindre end for mændene. Der er dog også for kvinderne en tendens til reduktionen er størst i det tilfælde, hvor den baseres på data fra perioden 1990-2005. Der er dog en tendens til at dette ikke er tilfældet både for de yngre aldre op til 25 år og de ældre dvs. fra ca. 72 år og herover.

Svarende til resultaterne for mænd haves, at den kortere dataperiode fører betydelige udsving i reduktionerne i de aldersbetingede dødeligheder, men i modsætning til hos mændene er der ikke længere sammenhængende aldersgrupper for reduktionen i dødelighederne er stigende med alderen. Tendensen til brud på Gompertz lov er derfor ikke udtalt for kvinderne

## 5. Korrektioner af fremskrivningsmetoden

Som angivet ovenfor fremkommer fremskrivningen ved, at udviklingen i det estimerede mortalitetsindeks,  $k(t)$  fremskrives og udviklingen i de aldersspecifikke dødeligheder beregnes ud fra de estimerede  $a(x)$  og  $b(x)$ . Herudfra kan udviklingen i den aldersspecifikke restlevetid i fremskrivningen beregnes.

Der foretages dog to yderligere korrektioner inden dette gøres. 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 den aldersbetingede dødelighed er estimeret kan den seneste observation indeholde en ikke ubetydelig residual i forhold til den estimerede værdi. Dette er en betydelig praktisk ulempe, som Lee-Carter metoden har tilfælles med andre metoder, der baserede sig på estimerede værdier. Bell (1997) foreslår en metode til at forhindre disse spring. Der anvendes en modificeret udgave af denne metode i denne fremskrivning.

For det andet kan Lee-Carter metoden som det fremgik ovenfor føre til uhensigtsmæssige forskelle i de årlige reduktioner af de aldersbetingede dødelighed i aldersgrupper, der ligger i nærheden af hinanden. Det skyldes, at effekten på den enkelte aldersspecifikke dødelighed fremkommer ved at udviklingen i mortalitetsindekset  $k(t)$  multipliceres med den aldersspecifikke vektor  $b(x)$ . Tilfældige udsving imellem  $b(i)$  og  $b(i+1)$  forstærkes derfor uhensigtsmæssigt. Problemet er størst hvis metoden baseres på en kortere dataserie og hvis metoden samtidig anvendes til længere fremskrivninger. De Jong & Tickle (2005) foreslår en udglatningsmekanisme for dette tilfælde. I nærværende fremskrivning er der anvendt en udglatningsmekanisme, som afviger fra denne.

Inden resultaterne af fremskrivningen præsenteres sidst i afsnittet gennemgås disse to justeringer.

### Bell's metode samt modifikationen 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, fore-

slår Bell (1997) anvendelse af en korrektionsmetode. I henhold til Bells fremstilling kan problemet med spring/- knæpunkter 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

$$Bias(x) = \ln(m(x, t)) - \ln(\hat{m}(x, t))$$

Sidste led på højre side angiver den estimerede værdi for de logaritmiske dødeligheder det sidste historiske år.  $x$  angiver som hidtil alder. Bell's foreslåede bias-justering indebærer, at den beregnede bias tillægges samtlige logaritmiske dødeligheder i fremskrivningen. Den bias-justerede fremskrivning bliver for fremskrivningsperiode  $s$  dermed

$$\ln(\hat{m}(x, t + s))_{biascorrected} = \ln(\hat{m}(x, t + s)) + Bias(x), \quad s > 0$$

Idet de fremskrevne dødeligheder er beregnet på Lee-Carter metoden, kan ovenstående formuleres som

$$\begin{aligned} \ln(\hat{m}(x, t + s))_{biascorrected} &= \ln(\hat{m}(x, t + s)) + \ln(m(x, t)) - \ln(\hat{m}(x, t)) \\ &= a(x) + b(x)k(t + s) + \ln(m(x, t)) - a(x) - b(x)k(t) \\ &= b(x)[k(t + s) - k(t)] + \ln(m(x, t)) \end{aligned}$$

Biaskorrektionen påvirker dødshyppighederne som følger

$$\begin{aligned} \ln(\hat{m}(x, t + s))_{biascorrected} &= \ln(\hat{m}(x, t + s)) + Bias(x) \Leftrightarrow \\ \hat{m}(x, t + s)_{biascorrected} &= \exp(a(x) + b(x)k(t + s)) \cdot \exp(Bias(x)) \end{aligned}$$

hvorfor det kan konstateres, at biaskorrektionen fremkommer ved multiplikation af de oprindelige fremskrevne dødeligheder med en aldersafhængig konstant. Som det ses af første linie i systemet ovenfor, er forskellen mellem de biaskorrigerede estimerede logaritmiske dødeligheder og de ikke-biaskorrigerede estimerede logaritmiske dødeligheder per definition konstant over tid, idet differencen jo netop er den aldersafhængige biaskorrektion  $Bias(x)$ .

Tages udgangspunkt i dødshyppighederne kan differencen mellem det biaskorrigerede og det ikke-biaskorrigerede estimat udtrykkes som

$$\hat{m}(x, t + s)_{biascorrected} - \hat{m}(x, t + s) = \hat{m}(x, t + s) \cdot (\exp(Bias(x)) - 1)$$

Eftersom de fremskrevne dødshyppigheder aftager eksponentielt over tid, vil forskellen mellem de biaskorrigerede og ikke-biaskorrigerede dødshyppigheder 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 nærliggende at undersøge, hvorvidt fremskrivningspræcisionen påvirkes af anvendelse af metoden. Til dette formål anvendes data fra 1835-2004 fra HMD. På baggrund af en dataperiode omfattende 40 år fremskrives dødelighederne for de efterfølgende 10 år, hvorefter den gennemsnitlige kvadratiske fremskrivningsfejl over alder for det 10. fremskrivningsår beregnes. Dataperiodens begyndelses- og sluttidspunkt forskydes et år, og eksperimentet gentages. Første data- og fremskrivningsperiode bliver dermed hhv. 1835-1874 og 1875-1884, mens de sidste bliver 1955-1994 og 1995-2004. Eksperimentet gennemføres både med og uden anvendelse af Bells korrektion. I 57 % af tilfældene er fejlen i det 10. fremskrivningsår mindst, når Bells korrektion anvendes. Dette taler umiddelbart for en ukritisk anvendelse af metoden. På den anden side er den numeriske værdi af fremskrivningsfejlen markant større i de tilfælde hvor fremskrivning uden korrektion er mest præcis. Den gennemsnitlige fremskrivningsfejl er derfor kun marginalt mindre, når Bell korrektion anvendes.

En udtømmende sammenligning af fremskrivningsegenskaberne med og uden Bells korrektion, bør i princippet omfatte et bredt udvalg af dataperiodelængder og fremskrivningshorisonter. Registreres fremskrivningsfejlen således i et eksperiment identisk med ovenstående, men med en fremskrivningsperiode på 20 år, vil den gennemsnitlige fejl være mindst, såfremt Bells korrektion udelades, til trods for, at fejlen ved anvendelse af korrektionen stadig er mindst i over halvdelen af tilfældene, her 55 %.

En tredje type eksperiment er gennemført ved at variere dataperiodens længde fra 15 til 100 år for et givet slutår. For hvert af disse slutår registreres den mindste fremskrivningsfejl i det 20. fremskrivningsår henholdsvis med og uden Bell's korrektion. I dette tilfælde vil fremskrivning uden Bell's korrektion være at foretrække i 60 % af tilfældene. Den mindste fremskrivningsfejl er således ikke nødvendigvis frembragt af den samme periodelængde i eksperimenterne med og uden anvendelse af Bells korrektion og formålet er dermed ikke at afgøre metodens indflydelse på en bestemt periodelængde, men derimod på den mindste fremskrivningsfejl.

På basis af disse resultater anvendes her en modificeret version af Bell's metode. Modifikationen indebærer for det første, at korrektionen udelukkende anvendes på 50-99-årige. Idet dødshyppigheden 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 længere sigt reduceret ved at lade effekten af korrektionen aftage over tid. Biaskorrektionen  $Bias(x)$  multipliceres således med potensfunktionen  $f(t) = \gamma^t$ , der for  $\gamma < 1$  er aftagende over tid, samtidig med at potensfunktionen sikrer, at biaskorrektionen i sidste dataår ( $t = 0$ ) er uændret.

Der er to yderpunkter for potensfunktionen. For det første tilfælde  $\gamma = 0$  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 i denne fremskrivning valgt at fastlægge parameteren  $\gamma$  således halveringstiden for korrektionen,  $T_{1/2}$ , bliver 10 år. I det halveringstiden defineres som

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

er

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

som i nærværende tilfælde får værdien  $\gamma = 0,93303$ . For fuldstændighedens skyld bliver den modificerede biasjusterede fremskrivning for fremskrivningsperiode  $s$  dermed

$$\ln(\hat{m}(x, t + s))_{\text{biascorrected}} = \ln(\hat{m}(x, t + s)) + \gamma^s \text{Bias}(x), \quad s > 0$$

hvor  $x \geq 50$ .

### 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. Derfor modificeres de estimerede dødshyppigheder i fremskrivningen, hvilket også indebærer, at de ovenfor præsenterede biasjusterede rater korrigeres yderligere.

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

$$\begin{aligned} \ln(\hat{m}(x, t + s))_{\text{smooth}} &= a(x) + b(x)k(t) + \tilde{b}(x)(k(t + s) - k(t)) + \gamma^s \text{Bias}(x) \\ &= a(x) + b(x)k(t) + \tilde{b}(x)(k(t + s) - k(t)) + \gamma^s \text{Bias}(x) + b(x)k(t + s) - b(x)k(t + s) \\ &= \ln(\hat{m}(x, t + s)) + \tilde{b}(x)(k(t + s) - k(t)) + \gamma^s \text{Bias}(x) - b(x)(k(t + s) - k(t)) \\ &= \ln(\hat{m}(x, t + s))_{\text{biascorrected}} + (\tilde{b}(x) - b(x))(k(t + s) - k(t)) \end{aligned}$$

hvor  $t$  her angiver basisåret eller det sidste år for hvilket data er tilgængelige og  $Bias(x) = 0$  for  $x < 50$ . Parameteren  $\tilde{b}(x)$  er en udglattet version af estimatet  $b(x)$ . I praksis er  $\tilde{b}(x)$  en såkaldt *cubic spline* af  $b(x)$ , hvilket essentielt indebærer, at der mellem samtlige punkter i  $b(x)$  er fittet et polynomium af grad tre under antagelser om minimering af såvel kurvatur i  $\tilde{b}(x)$  som af kvadratisk afstand til indgangene i  $b(x)$ .

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

## 6. Resultater af fremskrivningen

Beskrivelsen af resultaterne af fremskrivningen følger præsentationen af den historiske udvikling og starter derfor med en gennemgang af udviklingen i den aldersbetingede dødelighed i såvel hovedfremskrivningen som referencefremskrivningen og Velfærdskommissionens fremskrivning. Herefter gennemgås udviklingen i den aldersbetingede restlevetid, og der fokuseres særligt på udviklingen i middellevetiden og restlevetiden for en 60-årig.

### Aldersstrukturen i dødelighederne

De estimerede årlige reduktioner i den aldersbetingede dødelighed er generelt positive for alle aldersgrupper, jf. Figur 10 og Figur 11. Det betyder, at fremskrivningen over tid vil føre til, at dødelighederne reduceres for alle aldersgrupper.

Logaritmen til de aldersbetingede dødeligheder er i Figur 12 vist både for det seneste år i dataserien, som er 2005 og fremskrevet til 2050. Der vises både selve fremskrivningen baseret på dataserien fra 1990 og referenceforløbet baseret på dataserien fra 1965. For både mænd og kvinder reduceres den aldersbetingede dødelighed frem mod 2050.

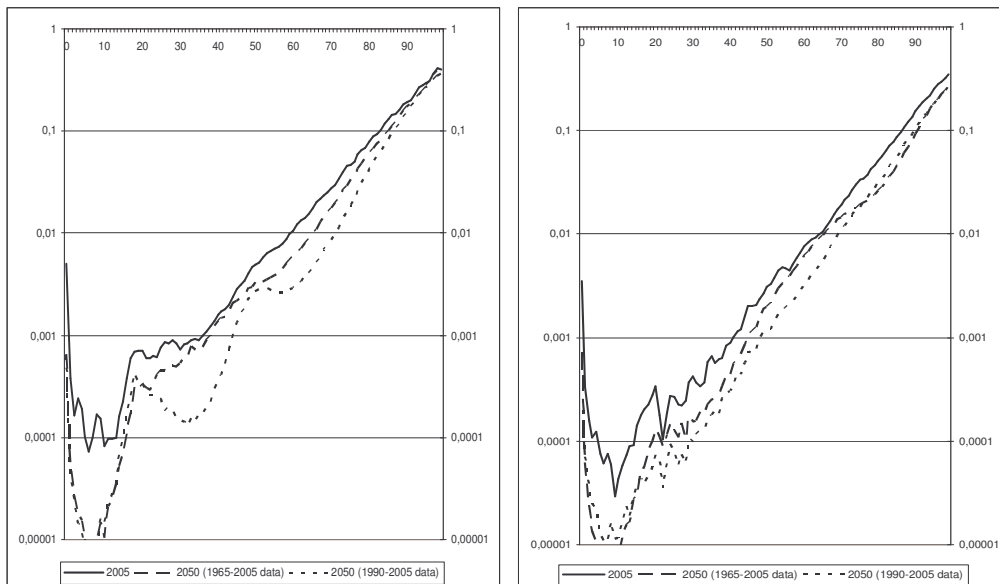
I fremskrivningen ses, at tendensen til, at de årlige reduktioner i den aldersbetingede dødelighed for mænd er voksende med alderen for aldersgrupperne 17-32 år og 52-60 år, som forventet fører til et ganske markant brud på Gompertz lov om, at dødeligheden vokser eksponentielt med alderen. For de to nævnte aldersgrupper er dødeligheden faldende med alderen i år 2050. For referencefremskrivningen baseret på dataserien fra 1965 er der tendenser til samme type problem for aldersgruppen fra 52 til 60-årige.

Samtidig fremgår det, at referencefremskrivningen indebærer, at reduktionen i dødeligheden vil være meget beskeden, særligt for aldersgrupperne fra 25 til 50 år, hvor der stort set ikke er nogen reduktion i dødelighederne.



For kvinder haves omvendt, at fremskrivningen leder til en stort set perfekt lineær vækst i logaritmen til dødeligheden med alderen i år 2050. Derimod leder reference-fremskrivningen baseret på dataserien fra 1965 til en tendens til brud på den lineære vækst som følge af, at de estimerede årlige reduktioner i dødeligheden blandt kvinder i aldersgrupperne fra 30-70 år er meget lave, hvis denne dataperiode lægges til grund.

**Figur 12. Aldersbetinget dødshyppighed for mænd (venstre figur) og kvinder (højre figur)**



Konklusionen er derfor, at for mænd har valget af den korte dataserie som udgangspunkt for fremskrivningen en negativ konsekvens for aldersstrukturen i dødeligheden sammenlignet med referenceserien, mens dette ikke er tilfældet for kvinder.

### Udviklingen i dødeligheden for de enkelte aldersgrupper

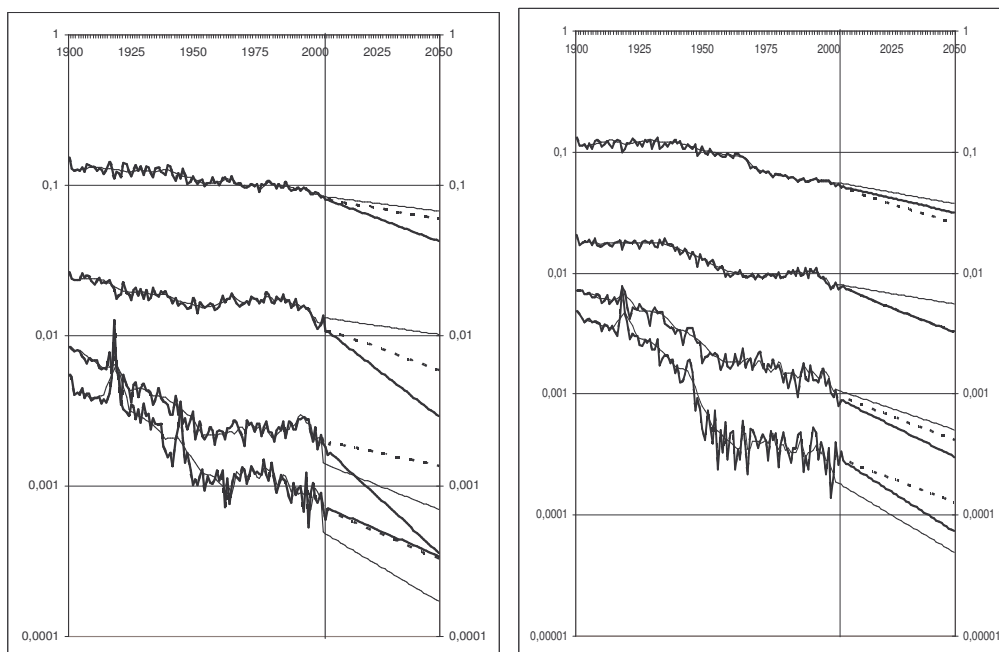
Betydningen af den valgte estimationsperiode illustreres klart ved at tage udgangspunkt i logaritmen til dødeligheden for en given alder. Lee-Carter metoden implicerer, at den fremskrevne udvikling i logaritmen til dødeligheden for en given aldersgruppe kan vurderes ud fra trenden i dødeligheden gennem den valgte estimationsperiode. I Figur 13 sammenlignes fremskrivningen (fed fuldt-optrukket kurve) med referencefremskrivningen baseret på dataserien fra 1965 (fed stiplede kurve) og fremskrivningen i Haldrup (2004), der er baseret på data for perioden 1904-2003 (tynd fuldt-optrukket kurve).

Som forventeligt er der især for gruppen af mænd i alderen 30 til 70 år en meget betydelig forskel på udviklingen i dødeligheden i de tre fremskrivninger. Dødeligheden i hovedfremskrivningen udviser en betydelig tendens til fald, fordi dødeligheden for disse aldersgrupper er faldet betydeligt gennem de seneste ti år. I perioden op til 1995 faldt dødelighederne stort set ikke for denne aldersgruppe af mænd. Referencefremskrivningen får derfor en væsentlig lavere reduktion i den fremtidige dødelighed, fordi denne fremskrives ud fra udviklingen i to historiske to delperioder med forskellig historisk reduktion. For aldersgrupperne op til 40 år har Velfærdskommissionens fremskrivning stort set samme årlige reduktion som referencefremskrivningen. Der er imidlertid en betydelig forskel i startniveauet for dødelighederne, fordi Velfærdskommissionens fremskrivning ikke er korrigeret, så den tager udgangspunkt i den seneste observation. Forskellen mellem de to fremskrivninger viser derfor den potentielle effekt af at anvende den af Bell (1997) foreslåede korrektion ved fremskrivninger baseret på den lange historiske dataserie. Med anvendelsen af de kortere data serier er behovet for korrektionen betydeligt mindre på danske data.

Den historiske tendens til, at de senere års reduktion i dødeligheden er vokset relativt, fremgår ligeledes af Figur 13, hvor det såvel for mænd som kvinder gælder, at dødeligheden for 60-årige (og i mindre grad 80-årige) reduceres betydeligt mere i referencefremskrivningen end i Velfærdskommissionens fremskrivning.

For kvinder gælder generelt, at forskellen mellem de tre fremskrivninger er væsentligt mindre end for mænd.

**Figur 13. Dødshyppighed for udvalgte aldersgrupper (20, 40, 60 og 80 år), mænd (venstre figur) og kvinder (højre figur), logaritmisk skala**



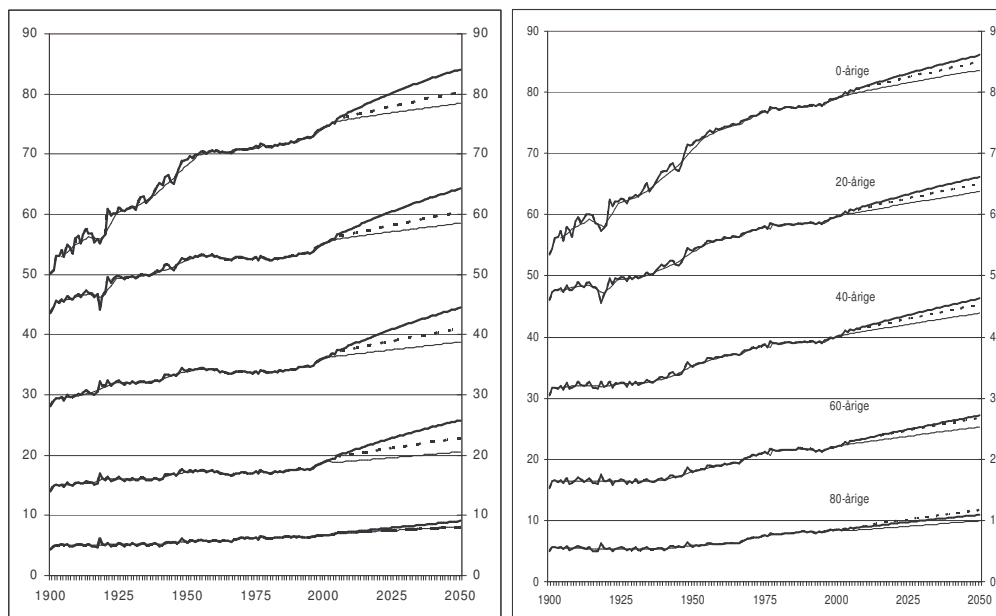
Note: Da dødeligheden er voksende med alderen er 80-årige øverst, efterfulgt af 60-årige osv.

### Udviklingen i den aldersbetingede restlevetid

Reduktionen i aldersbetingede dødelighed over tid giver anledning til en stigning i den aldersfordelte restlevetid. Udviklingen i denne i de tre ovenfor nævnte fremskrivninger er vist i Figur 14.

Stigningen i den aldersbetingede restlevetid er i hovedfremskrivningen en forlængelse af de seneste års udvikling kombineret med en tendens til gradvist aftagende vækst. Referencefremskrivningen betyder for mænd, at stigningen i fremskrivningens første år er lavere end i den foregående periode for alle aldersgrupper. For kvinder er forskellen mellem hoved- og referencefremskrivningen mindre markant, hvilket også gør sig gældende for de første års fremskrivning. Velfærdskommissionens fremskrivning har et lavere niveau fra starten (hvilket hænger sammen med at de seneste tre data-års betydelige stigninger i den aldersbetingede restlevetid ikke indgår), og stigningen er generelt endnu lavere end i referenceforløbet, hvilket først og fremmest hænger sammen med, at reduktionen i den aldersbetingede dødelighed for de ældre aldersgrupper undervurderes, fordi der anvendes så lang en dataserie.

**Figur 14. Restlevetid for udvalgte aldersgrupper (0, 20, 40, 60 og 80 år), mænd (venstre figur) og kvinder (højre figur)**



I det følgende fokuseres på udviklingen i restlevetiden for to udvalgte aldersgrupper: Middellevetiden (= restlevetiden for en 0-årig) og restlevetiden for en 60-årig.

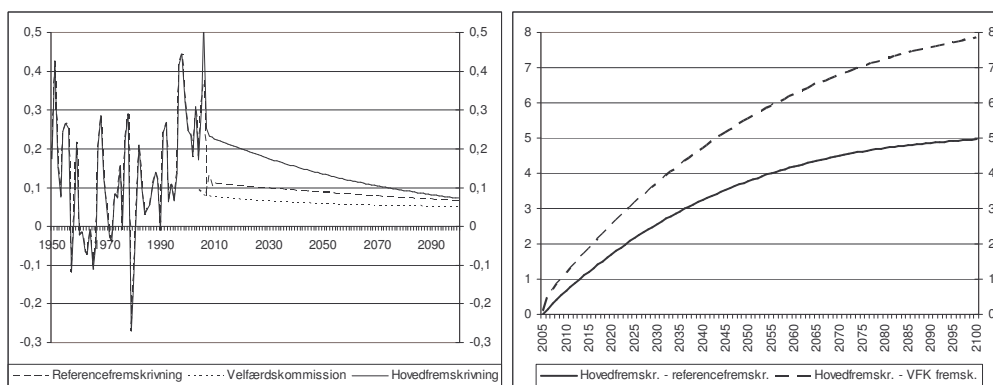
## Udviklingen i middellevetiden for mænd

I perioden fra 1995 til 2005 er middellevetiden for mænd vokset fra 72,7 år til 75,5 år, dvs. med 2,8 år. Det svarer til en gennemsnitlig årlig vækst på 3,4 måneder om året. Den samlede vækst over 10-året fra 1995-2005 er af samme størrelsesorden som væksten i hele perioden fra 1951-1995.<sup>11</sup> Udgangspunktet for væksten i den mandlige middellevetid er derfor højt.

Fremskrivningen indebærer en gradvis opbremsning i denne vækst, som betyder, at det forventes, at middellevetiden vokser med yderligere 1,5 år frem til 2010, hvor den vil være nået op på 77,0 år. I perioden fra 2010 til 2020 vil middellevetiden ifølge fremskrivningen vokse med yderligere 2,1 år og således nå op på 79,1 år i 2020. Det svarer til en gennemsnitlig vækst i middellevetiden på 2,5 måned om året. Middellevetiden for mænd når op på 80 år i 2025. I den efterfølgende periode vokser med knap 2 måneder om året og når således 81 år i 2031 og 84 år i 2050.

Udviklingen i den årlige tilvækst i middellevetiden i den historiske periode fra 1950 til 1990 udviser betydelig årlig variation omkring en vækst på ca. 0,1 år pr. år. Efter 1990 er væksten i middellevetiden steget og har som nævnt ligget på 0,3 år pr. år i gennemsnit. I fremskrivningen antages, at væksten i middellevetiden gradvist reduceres gennem det 21. århundrede fra et niveau på omkring 0,25 år i starten til et niveau på 0,07 år imod slutningen af århundredet, jf. Figur 15.

**Figur 15. Væksten i middellevetiden for mænd (venstre figur) og differensen i middellevetid for mænd mellem hovedfremskrivningen og de øvrige fremskrivninger (højre figur)**



I såvel referencfremskrivningen som i Velfærdskommissionens fremskrivning indebærer metoden, at det lave langsigtede niveau stort set er gældende fra starten af fremskrivningsperioden. I begge disse fremskrivninger betragtes den seneste ti-års periode derfor som "unormal", og middellevetidsudviklingen vender tilbage til niveau-

<sup>11</sup> I denne periode var den gennemsnitlige årlige vækst på knap 0,9 måned pr. år.

et før 1990 fra første år i fremskrivningen. Referencefremskrivningen indebærer lidt højere vækst end Velfærdskommissionens fremskrivning, idet væksten falder fra et niveau på omkring 0,13 år pr. år i starten til et niveau på 0,06 år pr. år i slutningen af århundredet. Den tilsvarende udvikling i Velfærdskommissionens fremskrivning indebærer et fald i væksten fra 0,8 år pr. år til 0,5 år pr. år på langt sigt, jf. Figur 15.

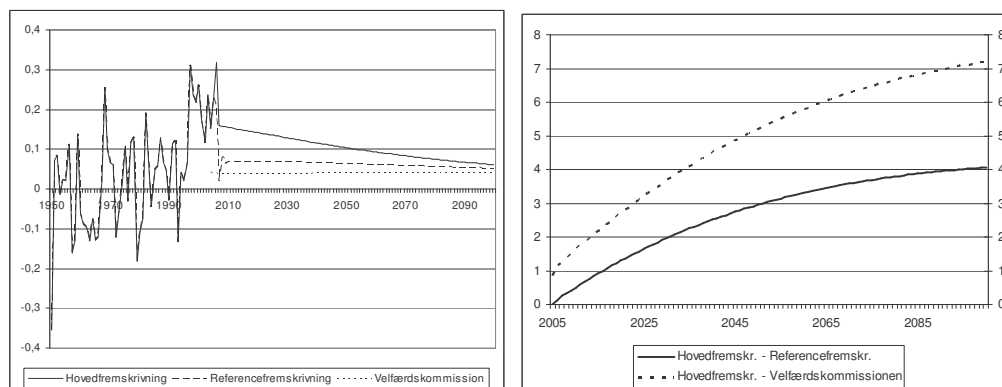
Forskellene i de årlige væksthiveauer indebærer, at fremskrivningerne af middellevetiden over tid kommer til at afvige betragteligt. I 2025 er middellevetiden godt 2 år højere i hovedfremskrivningen end i alternativet, mens forskellen til Velfærdskommissionens fremskrivning er på godt 3 år i 2025. I 2050 er de tilsvarende forskelle på henholdsvis 3,8 og 5,6 år, jf. Figur 15.

### Udviklingen i restlevetiden for 60-årige mænd

Restlevetiden for 60-årige mænd er vokset fra 17,7 år i 1995 til 19,7 år i 2005. Hvilket svarer til en gennemsnitlig årlig vækst i perioden på vækst på 2,4 måneder pr. år. Det er en bemærkelsesværdig høj vækst over denne periode. For at illustrere størrelsesordenen kan det noteres, at restlevetiden for en 60-årig mand også steg med 2,0 år i perioden fra 1923 til 1995. Væksten i de seneste ti år svarer således til væksten i de foregående 72 år. I perioden fra 1923 til 1995 var den gennemsnitlige vækst på 0,33 måneder pr. år.

Resultatet af fremskrivningen er, at væksten i restlevetiden for 60-årige 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 1,8 måneder om året – hvilket er noget lavere end de seneste 10 års gennemsnitlige vækst. Fra dette niveau falder den årlige vækst gradvist mod et langsigtet niveau på godt 0,7 måneder om året.

**Figur 16. Væksten i restlevetiden for 60-årige mænd (venstre figur) og differensen i restlevetid for 60-årige mænd mellem hovedfremskrivningen og de øvrige fremskrivninger (højre figur)**



I såvel referencefremskrivningen som Velfærdskommissionens fremskrivning er resultatet, at den årlige vækst også i fremskrivningens første år har et niveau, der svarer til det langsigtede og derfor er på henholdsvis 0,6 – 0,7 måneder pr. år og 0,5 – 0,6 måneder pr. år, jf. Figur 16.

Den højere vækst i de første år af hovedfremskrivningen betyder, at niveauet for restlevetiden for 60-årige mænd bliver højere end i de øvrige fremskrivninger. Frem til 2010 vokser restlevetiden for 60-årige mænd fra 19,7 år i 2005 til 20,6 år. I 2020 er restlevetiden for 60-årige mænd vokset til 22,1 år og i 2050 til 25,7 år. I referencefremskrivningen er restlevetiden i 2050 vokset til kun vokset til 22,8 år, mens Velfærdskommissionens fremskrivning indebærer en restlevetid for 60-årige mænd på 20,5 år i 2050.

Det er bemærkelsesværdigt, at forskellen mellem fremskrivningernes udvikling i restlevetiden for 60-årige er næsten lige stor som forskellen mellem fremskrivningernes middellevetidsudvikling. I løbet af det 21. århundrede bliver forskellen mellem hovedfremskrivningen og Velfærdskommissionens fremskrivning 8,0 år, mens forskellen i restlevetiden for 60-årige mænd er 7,2 år. Det gælder således, at stort set hele den ekstra stigning i middellevetiden skyldes en ekstra stigning i restlevetiden for 60-årige. Det samme gør sig gældende for forskellene mellem hovedfremskrivningen og referencefremskrivningen. Her vokser forskellen i middellevetiden til 5,0 år mens forskellen i restlevetiden for 60-årige vokser til 4,0 år.

### **Udviklingen i middellevetiden for kvinder**

I perioden fra 1995 til 2005 er middellevetiden for kvinder steget med 2,2 år eller gennemsnitligt med 2,6 måneder om året. Selvom det er væsentligt lavere end væksten i mændenes middellevetid i samme periode (3,4 måneder pr. år), er der også for kvinderne tale om en betydelig acceleration i væksten. Til sammenligning steg middellevetiden for kvinder også med 2,2 år over perioden fra 1970 til 1995. Den gennemsnitlige årlige vækst er kun på 1,1 måned i denne periode.

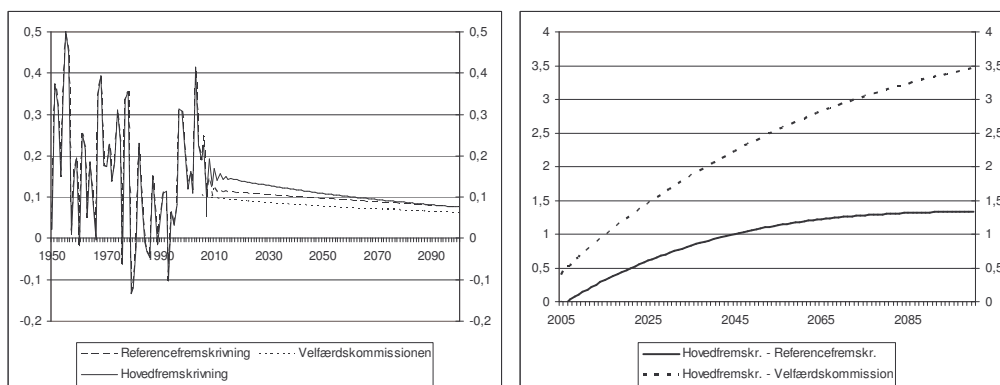
Hovedfremskrivningen indebærer, at væksten i middellevetiden i de første år er lavere end i den seneste historiske periode og ligger på ca. 1,8 måneder pr. år. Væksten reduceres gradvist gennem fremskrivningsperioden og ligger på 0,9 måned om året ved udgangen af århundredet, jf. Figur 17.

Middellevetiden vokser således fra 80,1 år i 2005 til 80,9 år i 2010 og 82,3 i 2020. I 2050 er middellevetiden for kvinder vokset til 86,0 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.

Den gennemsnitlige vækst i middellevetiden i hovedfremskrivningen er samtidig højere end væksten i både referencefremskrivningen og Velfærdskommissionens fremskrivning. Forskellen skyldes ligesom for mændene først og fremmest, at hovedfremskrivningen indebærer en højere vækst i de førstkommande år, fordi de to øvrige fremskrivninger indebærer, at langsigtsvækstraten stort set indtræder fra første år i

fremskrivningen, således at væksten bliver mindre end det halve af væksten i de seneste år i disse fremskrivninger. I 2050 betyder disse antagelser, at middellevetiden for kvinder 85,0 år i referencefremskrivningen og 83,6 år ifølge Velfærdskommissionens fremskrivning. Det svarer til en forskel på henholdsvis 1,0 år og 2,4 år til hovedfremskrivningen, jf. Figur 17.

**Figur 17. Væksten i middellevetiden for kvinder (venstre figur) og differensen i middellevetid for kvinder mellem hovedfremskrivningen og de øvrige fremskrivninger (højre figur)**



### Udviklingen i restlevetiden for 60-årige kvinder

Ligesom for mændene er der sket en mærkbar stigning i væksten i restlevetiden for 60-årige kvinder i de seneste ti år. I denne periode er restlevetiden vokset med 1,4 år, hvilket svarer til en gennemsnitlig stigning på 1,7 måneder om året. Til sammenligning steg restlevetiden for 60-årige kvinder også med 1,4 i perioden fra 1968 til 1995. I de seneste 10 år er restlevetiden derfor vokset lige så meget som i de foregående 27 år.

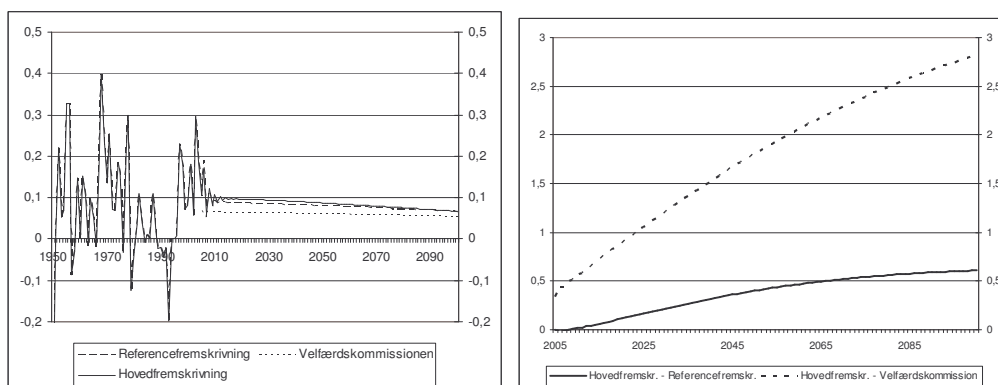
Hovedfremskrivningen indebærer, at væksten på kort sigt bliver noget lavere end i den seneste periode, men fortsat vil ligge på omkring 1,2 måneder om året, hvilket er næsten dobbelt så højt som i perioden forud for 1995. Væksten er svagt faldende over tid og er på langt sigt 0,9 måneder om året, jf. Figur 18. Det bemærkes, at fremskrivningen af væksten i restlevetiden for 60-årige kvinder i referencefremskrivningen er stort set sammenfaldende med hovedfremskrivningen. Den tilsvarende vækst i Velfærdskommissionens fremskrivning ligger til sammenligning noget under de to øvrige med et niveau på mellem 0,6 og 0,7 måneder pr. år.

Den højere vækst i de første år af hovedfremskrivningen og i referencefremskrivningen betyder, at niveauet for restlevetiden for 60-årige kvinder bliver højere i disse fremskrivninger end i Velfærdskommissionens fremskrivning. Frem til 2050 vokser restlevetiden for 60-årige kvinder til 27,1 år, hvilket er 4,3 år mere end i 2005. De tilsvarende tal for referencefremskrivningen og Velfærdskommissionens fremskrivning er 3,9 år og 2,5 år.



Som hos mændene er forskellen mellem fremskrivningernes udvikling i restlevetiden for 60-årige næsten lige stor som forskellen mellem fremskrivningernes middellevetidsudvikling. I løbet af det 21. århundrede bliver forskellen mellem hovedfremskrivningen og Velfærdskommissionens fremskrivning 4,0 år, mens forskellen i restlevetiden for 60-årige mænd er 3,5 år. Det gælder således, at stort set hele ændringen i middellevetidsudviklingen mellem fremskrivningerne skyldes en ekstra stigning i restlevetiden for 60-årige.

**Figur 18. Væksten i restlevetiden for 60-årige kvinder (venstre figur) og differensen i restlevetid for 60-årige kvinder mellem hovedfremskrivningen og de øvrige fremskrivninger (højre figur)**



### Sammenfatning af fremskrivningsresultater

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 10 år end i de foregående 40 til 50 år. Udviklingen er særligt markant 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.

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

Denne fremskrivning betyder en opjustering i den forventede vækst i den aldersbetingede restlevetid i forhold til tidligere fremskrivninger, som var baseret på en data-serie fra 1900, hvor perioden med lav vækst indgik med betydelig vægt. Opjusteringen af væksten i de aldersbetingede restlevetid skyldes i meget høj grad en opjustering af restlevetiden for personer på 60 år og derover.

## Litteratur

Andersen, T. M. & L. H. Pedersen (2005): Demografi, velstanddilemma og makro-økonomiske strategier, *Nationaløkonomisk Tidsskrift* vol. 143, no. 2, 189-229

Andreev, K. F. & J. W. Vaupel (2006): Forecasts of Cohort Mortality after age 50, Working paper 2006-012, Max-Planck-Institut für demografische Forschung

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

Booth, H., J. Maindonald, & L. Smith (2002): Applying Lee-Carter under conditions of variable mortality decline, *Population Studies* vol. 56, 325-336

Booth, H, L. Tickle, & L. Smith (2005): Evaluation of the Variants of the Lee-Carter Method of Forecasting Mortality: A Multi-Country Comparison, *New Zealand Population Review*, vol. 3 no.1, 13-34

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

Carter, L. & A. Prskawetz (2001): Examining Structural Shifts in Mortality Using the Lee-Carter Method, Working paper 2001-007, Max-Planck-Institut für demografische Forschung

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

Giroso, F. & G. King (2005): A Reassessment of the Lee-Carter Mortality Forecasting Method, Working paper Harvard University

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

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

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

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

Lundström, H. & Qvist, J. (2004): Mortality forecasting and Trend Shifts: an Application of the Lee-Carter Model to Swedish Mortality Data, *International Statistical Review* vol. 72 no.1, 37-50

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 (2004a): Fremtidens velfærd kommer ikke af sig selv. Analyserapport. Velfærdskommissionen, København.

Wilmoth, J. R. (1995): Are Mortality Projections Always More Pessimistic When Disaggregated by Cause of Death?, *Mathematical Population Studies* vol. 5, no.4 293-319

Wilmoth, J. R. (1996): "Mortality Projections for Japan: A Comparison of four Methods" in Casselli, G. & A. Lopez: Health and Mortality among Elderly Populations