



Udarbejdelse af et indeks af 25 indikatorer

Metodenotat

Januar 2019

Indhold

Introduktion	3
Pærer og bananer – sammenkog af 25 indikatorer til ét indekstal	3
Databearbejdning	4
Indsamling af data fra forskellige kilder og af forskellig størrelsesorden	4
Indikatorer, der berører få, er ofte de mest alvorlige problemer	4
Det simple alternativ – plus eller minus	5
Fokus på problemerne	5
Afgrænsning af problemerne	5
Indeksring	6
Absolut eller relativ score	6
Normalisering og standardisering	6
Diskussion af vægtning gennem Principal Component Analysis (PCA)	7
Ensretning af udviklingen for negativ udvikling	7
Beregning af det samlede indeks	7
Statistisk usikkerhed	7
Substitution	9
Metoden i Det sociale velfærdsindeks – og et supplement	11
En simpel indekseberegning – den historiske udvikling siden 2010	11
Plus eller minus - den aktuelle udvikling siden sidste år	12
Robusthedstjek – normalisering	12
Bilag	14
Uddybende om min-max normalisering	14
Indekseberegning med geometrisk gennemsnit	15
Indekseberegning ved forskellig baselines	16

Introduktion

Dette notat beskriver de metodiske overvejelser og valg, vi har gjort os omkring konstruktionen af Det sociale velfærdsindeks, SV25. Overordnet kan overvejelserne inddeles i følgende trin:

- Databearbejdning, herunder kategorisering og datakilder
- Indeksering, herunder normalisering, standardisering og relativt versus absolut indeks
- Indeksberegning og fortolkning, herunder statistisk usikkerhed, substitution og aritmetisk versus geometrisk gennemsnitsberegninger

Hvert af disse punkter er diskuteret i nærværende notat, og på denne baggrund argumenterer vi for, at den mest frugtbare beregning og præsentation af SV25 – med de præmisser og datakrav der er for indekset – er gennem en gennemsnitsberegning af indekserede indikatorer med en baseline som referencepunkt. Indekseringen af indikatorerne foregår ved at sætte dem til værdien 100 i år 2010, der fungerer som baseline, og hvert af de følgende år beregnes som den procentvise ændring fra den baseline. Metoden er mere gennemsnitsrig end metoder, hvor der indirekte vægtes på baggrund af eksempelvis standardafvigelser for en given årrække. Der er til gengæld også fordele ved andre metoder, hvilket diskuteres i notatet her.

Indeksværdien beregnes som gennemsnit af syv områdespecifikke indekser. Valget af disse områder beskrives i notatet "Indikatorer til et socialt velfærdsindeks – teoretisk baggrund og overvejelser om udvælgelse" (CASA og Analyse & Tal, 2018).

Derudover foreslår vi som supplement en opgørelse af udviklingen i de 25 indikatorer, der blot er summen af antallet af signifikant positive og negative udviklinger, der er fra et år til det næste. Fordelen ved denne metode er, at der ikke kan herske usikkerhed om vægtning af indikatorerne, da de alle tæller samme værdi, ét, og det er tolkningsmæssigt meningsfuldt at skelne mellem signifikante og insignifikante udviklinger. Til gengæld kan man udelukkende udtale sig om år til år og ikke se på udvikling over flere år.

Til sidst i notatet er en diskussion af fordele og ulemper ved de to metoder samt en yderligere forklaring af fremgangsmåden. Først her nogle overordnede betragtninger omkring dét at lave ét samlet mål for 25 mere eller mindre forskellige fænomener.

Pærer og bananer – sammenkog af 25 indikatorer til ét indekstal

Det sociale velfærdsindeks er en samling af 25 indikatorer, der spænder bredt over socialområdet. Disse er regnet sammen, først til 7 områder og dernæst til ét samlet indeks. Det betyder, at det samlede indeks giver en slags gennemsnitsbetragtning af en række mere eller mindre forskelligartede sociale udfordringer. Man kan med rette indvende, at vi således blander pærer og bananer og forsøger at sige noget fælles for begge dele. Det giver nogle statistiske og fortolkningsmæssige udfordringer, som i øvrigt også gør sig gældende for lignende indekser, såsom Human Development Index (HDI) og Social Justice Index, som dette indeks er inspireret af.

Man kan samtidigt argumentere for, at selvom indekset er en blandet frugtskål, så er det alt sammen er frugt. Fænomener som selvmord, arbejdsløshed og dårligt helbred er forskellige, men de afspejler hver deres hjørne af den sociale velfærd – eller snarere mangel på samme – på hver deres måde. Et indeks som dette skal således tages for det, det er: Et sammenkog af en række forskellige, komplekse problematikker, der tilsammen giver et forsimplet men informativt billede af de overordnede sociale forhold.

Selve konstruktionen af et indeks handler om at sikre en høj kvalitet og ensartethed af datamaterialet, så de samlede indekstal også fortæller en samlet historie. I den forbindelse er det afgørende, at der er fuld gennemsnitsrigdom omkring den måde, indikatorerne og indekseringen er foretaget, så det er muligt at regne baglæns fra det samlede indekstal til de 7 områder og de 25 indikatorer.

Helt overordnet er valget af indikatorer og kriterier for datakvaliteten afgørende for den metode, der er meningsfuld i konstruktionen af SV25. Valg af data går således forud for valg af metode. En anden tilgang ville have været at tage udgangspunkt i en metode for indeksberegningerne og udvælge data, der passer til forudsætningerne for disse beregninger efterfølgende. Det betyder for velfærdsindekset, at det bygger på et

omfattende datamateriale (25 indikatorer mod 4 i HDI), men at det til gengæld ikke er meningsfuldt at anvende de samme metodiske greb ift. rangering og gennemsnitsberegninger, som andre indekser benytter sig af.

I den følgende gennemgang beskriver vi processen fra indikator til indeks samt de faldgruber, der er ved de enkelte valg. Det indebærer en diskussion af datakvaliteten og usikkerheden, gennemsnitsberegninger, min-max normalisering og indeksering. En del er inspireret af metodenotater og artikler om Human Development Index og – i mindre omfang – af Social Justice Index. Det gælder både de tekniske greb og de mere pragmatiske valg, man må gøre sig på baggrund af det udvalgte datamateriale.

Databearbejdning

Indsamling af data fra forskellige kilder og af forskellig størrelsesorden

De 25 indikatorer bygger på to forskellige typer af data: Registerdata og surveydata. Nogle af indikatorerne kan direkte aflæses fra deres kilder, andre er lavet gennem beregninger på datamaterialet.

Det er et krav til datakilden, at indikatorerne kan ajourføres årligt og er blevet det over en vis periode. Da mange større surveys kun udføres hvert andet eller tredje år, såsom Tryghedsmålingerne og Den Nationale Sundhedsprofil, er størstedelen af data indsamlet fra registre. Eurostat står dog bag nogle årlige surveys af høj kvalitet, som er inddraget i registret, og også Justitsministeriets Offerundersøgelser indgår.

- Registerdata – i alt 15 indikatorer. Kilderne er
 - Statistikbanken i Danmarks Statistik (Indkomstulighed, Bruttoledighed, Langtidsledighed, Ingen kompetencegivende uddannelse, økonomisk fattige og børnefattigdom)
 - Domstolsstyrelsen (Udsættelse af lejebolig)
 - LO (Kompensationsgrader for dagpengemodtagere og kontanthjælpsfamilier)
 - AE (Økonomisk fattige, Børnefattigdom)
 - Egne beregninger på Danmarks Statistik (Langvarig passiv forsøgelse, Udsatte unge, Ingen g. klasseseksamen, Social arv i uddannelse)
 - Sundhedsdatastyrelsen (Selvmord)
 - RKI (Dårlige betalere)
- Surveydata – i alt 10 indikatorer. Surveymaterialet kommer fra kilder:
 - Eurostats SILC-program, indsamlet af Danmarks Statistik (Dårligt helbred, Aktivitetsbegrænsninger, Langvarig sygdom, Ufrivillig deltidsbeskæftigelse, Bagud med regningerne, Svage sociale relationer, Børnefamilier uden forældre i beskæftigelse)
 - Danmarks Statistiks velfærdsindikatorer (Svært at få pengene til at slå til)
 - Justitsministeriets Offerundersøgelse – (Ofre for kriminalitet, Frygt for kriminalitet)

De 25 indikatorer har ikke alene forskellige datakilder, de har også forskellig størrelsesorden – dvs. det antal personer, der er berørt af det sociale problem. Mange af indikatorer opgøres som procentdele af befolkningen. Størrelsen på indikatorerne varierer fra under 1 procent til over 30 procent alt efter, hvor stor en del af populationen, problemet berører. Det betyder, at der for nogle indikatorer skal ske en relativt stor ændring i antallet af berørte personer, for at det viser sig som et mærkbart procentvist udsving i indikatoren, som er det, vi måler i indekset. Dette er ikke nødvendigvis et problem, men det kræver at man overvejer konsekvenserne og eventuelle måder at ensarte indikatorerne, hvilket uddybes i de følgende afsnit.

Indikatorer, der berører få, er ofte de mest alvorlige problemer

Enkelte indikatorer er antalsmæssigt meget små. Det gælder især selvmord, treårsfattige og udsættelser fra lejeboliger. Udsving i disse indikatorer kan hurtigt få en relativt stor betydning for indekset. Fx steg andelen af treårsfattige fra 0,6 procent til 0,8 procent i perioden 2010-2016. Antalsmæssigt er dette ikke mange personer – ca. 13.000 – men procentvis er det 33 procent, hvilket er mest markante udviklinger blandt indikatorerne. Til sammenligning steg bruttoledigheden fra 4,2 til 4,3 procent i 2016-2017. Det er en procentvis stigning på ca. 2,5

procent, men antalmæssigt berører det ligeså mange mennesker, som stigningen i treårsfattige på 33 procent. Udgangspunktet har således en stor betydning for, hvordan en ændring giver udslag i den procentvise udvikling.

Argumentet for at acceptere dette forhold er, at de problematikker, der er medtaget i indekset og som berører relativt få mennesker, også er tilsvarende mere alvorlige end dem, der berører flere mennesker. Treårsfattigdom – såvel som selvmord og udsættelser – er så alvorligt et socialt problem, at en stigning i antallet fra ca. 42000 til 55000 personer skal kunne måles på indekset.

Et alternativ ville være at beregne udviklingen i absolutte personer frem for den relative, eller at vægte de indikatorer der berører mange personer højere end andre. Førstnævnte vil dog gøre de mere 'sjældne' indikatorer stort set ubetydelige for indekset. Sidstnævnte ville være problematisk ved likert-skala-indikatorer, hvor en større andel vil være berørt af et problem, når man medtager de andre end yderkategorierne.

Det simple alternativ – plus eller minus

Et andet alternativ er, at indekset kun angiver, om der er sket en positiv eller negativ udvikling for den enkelte indikator, og at størrelsesforholdet således ignoreres. Fordelen er, at man ikke skal tage stilling til, hvilken værdi udviklingen i en enkelt indikator har for indekset – den har altid den samme værdi. Det giver et enkelt og overskueligt billede af, om det på de fleste områder går fremad eller tilbage. Ulempen er, at fremgang ét år for en indikator ikke nødvendigvis er lige så stor som en efterfølgende tilbagegang, og således vil indekset med stor sandsynlighed være misvisende for udviklingen for en længere periode.

Således giver den simple tilgang nogle fordele, når vi skal identificere udviklingen fra et år til et andet, men også nogle problemer når vi ser på udviklingen over tid. I opsamlingen vender vi tilbage til denne problematik.

Fokus på problemerne

Et af kriterierne for udvælgelse af indikatorerne er, at de skal vise udviklingen i sociale problemer frem for velstanden. For så vidt er navnet "Det sociale velfærdsindeks" en smule misvisende. De fleste indikatorer måler netop andelen, der er udsat for et problem frem for andelen, der ikke er det. Fx anvender vi ledigheden frem for beskæftigelsen og andelen med dårligt helbred frem for andelen med godt helbred. Det giver en ensartethed i forhold til fortolkningen. Det har samtidig den virkning, at alle procentsatserne er i den lave ende, som oftest mellem 5 og 15 procent, da andelen der er berørt af hver enkelt problem ligger på det niveau. Indikatorerne får derved relativt sammenlignelige udgangspunkter for procentvise udsving, frem for hvis vi havde valgt at se på den modsvarende positive ende af fænomenet. Fx vil et fald i ledigheden fra 5 til 4 procent (dvs. med 20 procent) nærmest være usynligt, hvis vi i stedet havde målt beskæftigelsesstigningen fra 95 til 96 procent (1 procent).

Afgrænsning af problemerne

På trods af et fast udgangspunkt for indikatorerne – den negativt berørte gruppe – er der forskel på, hvor negativt berørt, gruppen skal være. Skal vi måle på den gruppe med enten 'meget dårligt' eller 'dårligt' helbred, eller skal vi udelukkende måle andelen med 'meget dårligt' helbred? Det kan betyde en forskel på, om niveauet for indikatoren er 30 pct. eller 5 pct., og således hvor meget en forskydning i procentfordelingen betyder for indekset. I sådanne tilfælde af en likert-skala, har vi konsekvent valgt at sammenlægge de to eller tre negative kategorier. Det skyldes flere ting. For det første kan man argumentere for, at eksempelvis dårligt helbred også er et problem, selvom det ikke er *meget* dårligt helbred. For det andet skyldes det et hensyn til målesikkerheden i spørgeskemaet – *reliabiliteten*. Ved at lægge kategorierne i den negative ende sammen, mindsker vi kravet til besvarelsernes præcision og den øgede usikkerhed, der følger.

Indeksering

Absolut eller relativ score

Der findes overordnet to forskellige måder at konstruere et indeks for et land, sådan som det gøres her: 1) relativt til andre lande eller til en gennemsnitsværdi, eller 2) relativt til én fast værdi. Det sociale velfærdsindeks er har udelukkende til formål at vise udviklingen i Danmark. Derfor er det mest hensigtsmæssigt at sætte sig fast på et udgangspunkt – en *baseline* – som hver ny værdi i indekset kan referere til over tid.

Da de alle indikatorer på nær én går tilbage til 2010, er dette år baselineåret. Indikatorerne i dette år sættes til værdien 100. De følgende værdier beregnes som den procentvis udvikling fra 2010:

$$\text{Indeks}_t = \frac{\text{værdi}_t}{\text{værdi}_{2010}} \cdot 100$$

Årstallet 2010 er således valgt som baseline for udviklingen af pragmatiske grunde. Det ville være interessant at kunne følge udviklingen endnu længere tilbage, men flere af indikatorerne mangler observationer før 2010, og det vil derfor give større usikkerhed, hvis vi medtager flere år før 2010. Valget af baseline har naturligvis betydning for, hvor udviklingen man sammenligner over starter, men det har kun en marginal betydning for udviklingen fra år til år (se bilag).

Normalisering og standardisering

Der er enkelte indikatorer, der afviger fra systematikken i definitionerne for andele og procentdele. Det gælder indkomstuligheden målt ved Gini-koefficienten og kompensationsgraderne for hhv. dagpenge og kontanthjælp. De ligger på omkring hhv. 27 procent og 50-60 procent. De kræver samtidigt relativt store ændringer over tid for at det vil komme til udtryk i mærkbare procentvise udsving. Særligt risikerer Gini-koefficienten at drukne, da den over de sidste mange har bevæget sig med under 0,5 procentpoint fra år til år (ca. to procent), hvorimod kompensationsgraden er mere følsom overfor politiske ændringer fra år til år.

En måde at justere dette på, som blandt andet anvendes i HDI, er den såkaldte *min-max normalisering*. Ved den sættes et minimum og maximum for, hvor store og små værdier, indikatoren kan antage. Disse grænser sættes typisk med udgangspunkt i normative mål eller overvejelser af, hvad værdierne rent praktisk kan antage. Grænserne får betydning for, hvor meget udviklingen i en indikator kommer til at veje, hvorfor det kræver en særlig grund eller et eksplicit politisk mål at flytte det fra 0.

Andre hyppigt anvendte metoder til at skalere datamaterialet, så indikatorerne kan sammenlignes, er standardisering – eksempelvis z-score standardisering. Standardisering er brugbart ved indekser, hvor man rangerer observationer i forhold til hinanden (fx lande eller kommuner) eller hvis man forholder hvert nye år til et gennemsnit, der flytter sig med udviklingen.¹ Det sociale velfærdsindeks er dog ingen af delene, da det udelukkende er et mål for Danmark med udgangspunkt i et givent år (2010).

Et mere relevant alternativ til den valgte metode kunne være at anvende min-max normalisering, hvor et givent antal standardvariationer fra gennemsnittet for perioden 2010-2017 udgjorde minimum og maximum. På den måde kunne man sikre procentvist mere ensartede udsving i indikatorerne uden at skulle tage normativt stilling til, hvor minimum og maximum skulle placeres. Der er dog det potentielt store problem, at de beregnede standardafvigelser og gennemsnit i perioden 2010-2017 ikke er repræsentative for den efterfølgende udvikling. Det kan medføre, at indikatorerne bliver vægtet u hensigtsmæssigt og u gennemsigtigt op eller ned. For

¹ For en yderligere forklaring af dette, se fx Mazziatto og Pareto (2013): *Methods for constructing composite indices: one for all or all for one?* i *Rivista Italiana di Economia Demografia e Statistica* Volume LXVII n. 2 Aprile-Giugno 2013

indeksudviklingen 2010-2017 er resultatet stort set identisk med den valgte, mere simple fremgangsmåde. I bilaget findes yderligere redegørelse for metoden og eksempler på resultater.

Diskussion af vægtning gennem Principal Component Analysis (PCA)

En relativt almindelig metode til at lave et indeks for eksempelvis rigdom eller socio-økonomisk status er ved hjælp af en såkaldt PCA. En PCA er en metode, der indfanger mest mulig variation i én dimension, og mest muligt af den resterende variation i en anden dimension – og så fremdeles. Med det kan man simplificere og identificere en eventuel hovedtendens i sit datasæt. Hver variabel får en score for, hvor meget af dens variation, der opfanges af den nye dimension, og disse scorer kan anvendes til at konstruere et vægtet gennemsnit til indekseberegningen. Således er PCA en metode til at vægte det væsentlige, baseret på korrelation i data.

I SV25 er vi interesseret i at indfange 7 teoretisk funderede områder igennem 25 indikatorer. Således er alle indikatorer lige vigtige, og vi kan ikke retfærdiggøre at vægte én indikator højere end en anden. PCA kan reducere data – muligvis til et mere klart og forståeligt billede af eksempelvis materiel velfærd – men er mindre egnet til indeks, der skal indfange flere forskellige og uafhængige aspekter under overskriften: Social velfærd.

Ensretning af udviklingen for negativ udvikling

Langt de fleste indikatorer måler andelen af personer, der er berørt af et problem. En stigning i værdien betyder således en negativ udvikling. Den eneste undtagelse er kompensationsgraden, hvor en positiv ændring også fortolkes som en positiv udvikling. Derfor vender vi de to indikator for kompensationsgraden (af dagpenge og kontanthjælp), så en negativ udvikling medfører en stigning. Det gøres ved at trække kompensationsgraden, der opgøres som procent af et givet lønniveau, fra 100 procent. Indikatoren er eksempelvis 40, hvis kompensationsgraden er 60.

Beregning af det samlede indeks

Statistisk usikkerhed

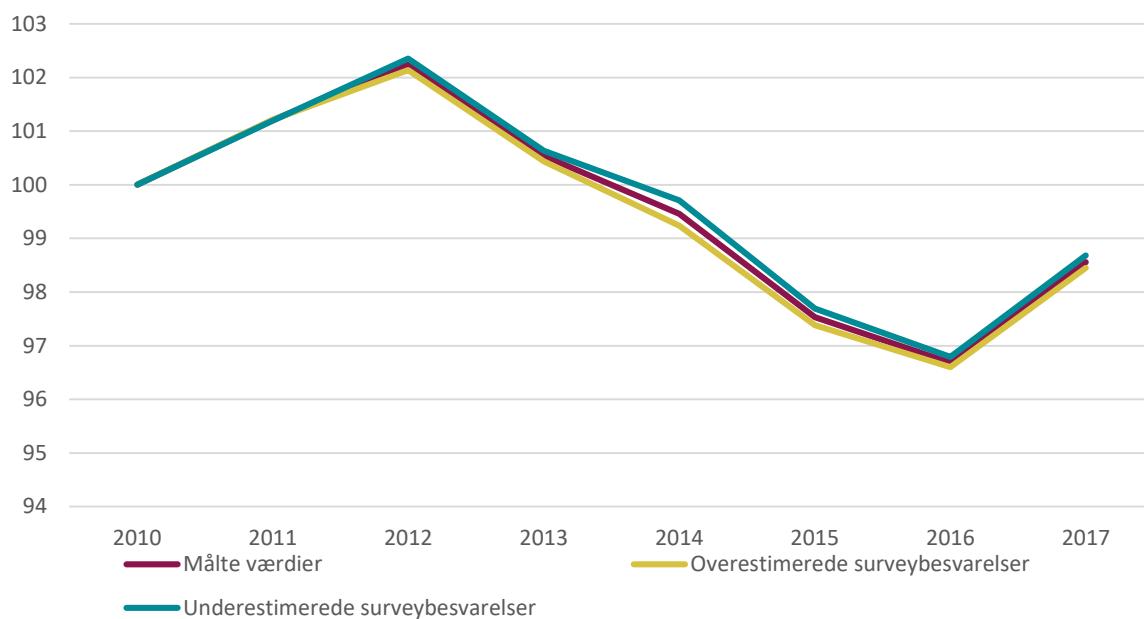
Godt halvdelen af indikatorerne baserer sig på spørgeskemaundersøgelser og har derfor en vis statistisk usikkerhed – stikprøveusikkerhed. Det medfører en risiko for, at en udvikling fra ét år til et andet skyldes målestøj og ikke en faktisk udvikling. For indikatorerne der stammer fra Eurostats Statistics on Income and Living Conditions (SILC)² er stikprøveusikkerheden størst – omkring 0,7 procent hver vej – mens stikprøveusikkerheden i Justitsministeriets offerundersøgelser og i Danmarks Statistiks Arbejdskraftsundersøgelse er noget mindre (stikprøvestørrelser på hhv. 12.000 og 84.000).

For den enkelte surveybaserede indikator kan usikkerheden gøre små udviklinger svære at tolke på. For det samlede billede af indikatorer har usikkerheden dog mindre betydning, da den bliver mindsket af, at indekset består af 25 indikatorer, hvoraf halvdelen ikke er påvirket af stikprøvestøj.

I figuren nedenfor ses beregningen af et indeks, hvor alle surveybaserede indikatorer er hhv. overestimerede og underestimerede på grund af stikprøveusikkerheden – det vil sige det værste tænkelige scenarie. Beregningen viser, at det vil medføre en fejlmargen på 0,3 procentpoint på hver side. Det samlede billede vil altså kun i meget lille omfang være styret af stikprøvestøj – selv i det tilfælde hvor støjen i de enkelte indikatorer trækker indekset i samme retning og således ikke udligner hinanden.

² <http://ec.europa.eu/eurostat/web/income-and-living-conditions/>

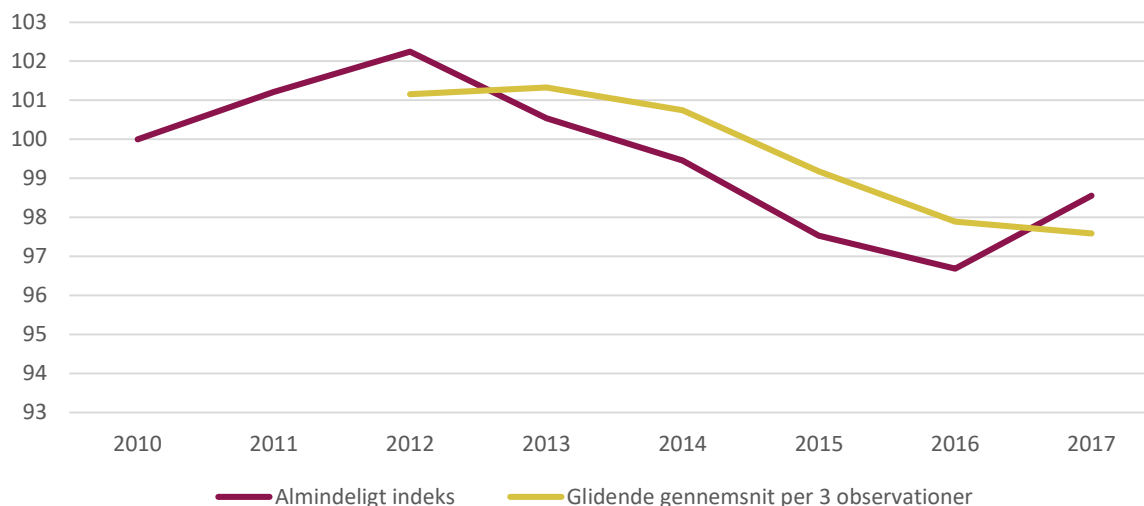
Figur 1: SV25 2010-2017 (2010=100) ved hhv. overestimerede og underestimerede surveybesvarelser



Kilde: Egne beregninger

En anden metode, der kan udjævne udsving fra stikprøvestøj eller målefejl i et enkelt år, er at beregne indikatorerne som glidende gennemsnit af de tre seneste observationer. Dette udjævner pludselige udsving op og ned, men det forsinker samtidigt udviklingen i indekset. Derfor bliver udviklingen i indekset mere stabil, men det bliver svært at lave en meningsfuld fortolkning af resultaterne.

Figur 2: Hhv. almindeligt indeks og indeks baseret på glidende gennemsnit for de seneste 3 observationer (2010=100)



Kilde: Egne beregninger

Figuren illustrerer, at udviklingen forekommer mere jævn, når man beregner et glidende gennemsnit for det pågældende år og de to forrige. Men den viser også, at udviklingen først slår igennem på indekset efter nogle

år. Det giver nogle fortolkningsmæssige udfordringer og aktualitetsproblemer, som vi vurderer er større end den gevinst, der er ved at foretage det glidende gennemsnit.

Substitution

Når de enkelte indikatorer er fastlagt, beregnes først områdespecifikke indekser og dernæst det samlede velfærdsindeks. Et centralt aspekt ved sammenlægning af flere indikatorer til ét indeks er spørgsmålet om substitution og compensation. Det handler kort fortalt om, at en positiv udvikling i én indikator kan kompensere for en negativ udvikling i en anden indikator.

Vi kan forestille os, at arbejdsløsheden falder mens erhvervs sygdomme stiger. Indekset giver da ikke noget incitament til komme erhvervs sygdommene til livs, hvis det er lettere at mindske arbejdsløsheden yderligere frem for at mindske antallet af erhvervs sygdomme. Substitutionen af en negativ udvikling i en indikator med en positiv udvikling i en anden gør således, at problemer kan gå usete hen, så længe der er fremgang på andre områder.

Denne problemstilling er diskuteret indgående i litteraturen omkring HDI og er også relevant for det sociale velfærdsindeks.³ I HDI valgte man i 2010 at mindske problemet ved at erstatte det almindelige aritmetiske gennemsnit med det geometriske gennemsnit. Det geometriske gennemsnit beregnes således:

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$$

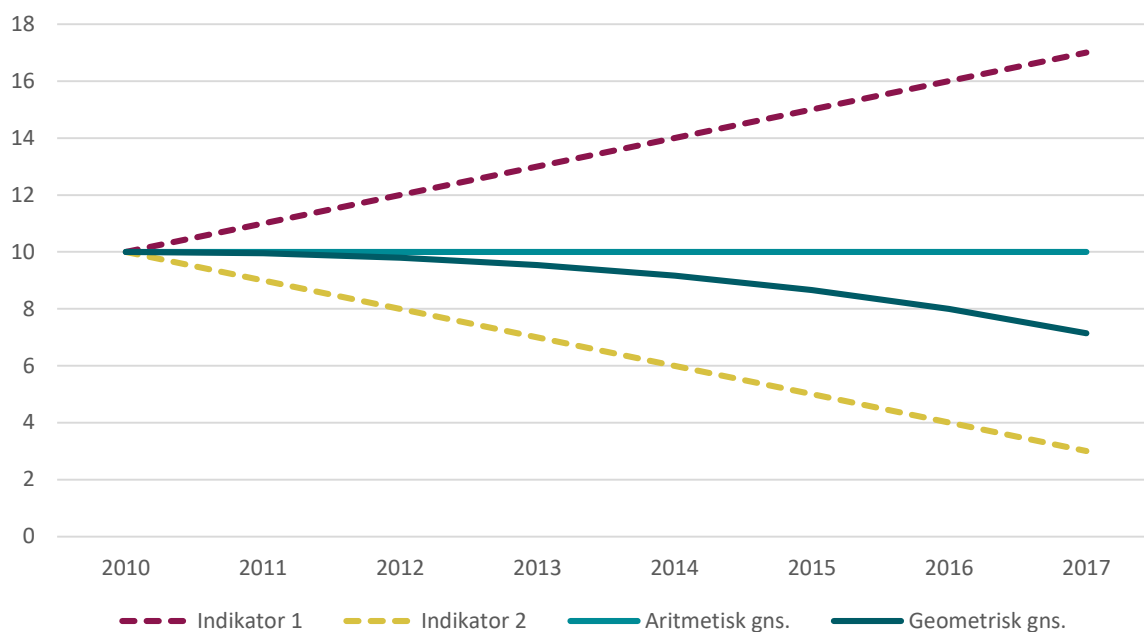
, mens det mere almindelige aritmetiske gennemsnit beregnes sådan:

$$A = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

Fordelen ved det geometriske gennemsnit er, at ulighed mellem indikatorerne har en negativ påvirkning på gennemsnittet i modsætning til det aritmetiske gennemsnit. Et fald i en indikator kan ikke kompenseres for 1:1 med en tilsvarende stigning i en anden indikator. Dette er illustreret i figuren nedenfor, hvor to indikatorer (de stiplede) hhv. stiger og falder med ét procentpoint hvert år. Med det aritmetiske gennemsnit er det ensbetydende med status quo, mens det geometriske gennemsnit bliver mindre i takt med, at afstanden mellem indikatorerne øges.

³ Klugman, Rodríguez og Choi (2011): *The HDI 2010: New controversies, old critiques*. UNDP.

Figur 3: Illustration af substitution ved geometrisk og aritmetisk gennemsnit



Kilde: Egne beregninger

Denne egenskab ved det geometriske gennemsnit medfører, at det er en fordel at have en god balance indenfor og imellem de forskellige områder, hvis indekset skal stige. Med andre ord fjerner man den fuldkomne substitution, hvor modsatrettede udviklinger i samme størrelsesorden udligner hinanden.

Teoretisk ville der være en pointe i at anvende det geometriske gennemsnit i det sociale velfærdsindeks. Det ville betyde, at de områder, hvor der sker en negativ udvikling, ville særlig stor betydning for indekset. I praksis er der dog det problem, at indekstallene bliver større, når der sker en negativ udvikling (omvendt i HDI). Det betyder, at vi med det geometriske gennemsnit ville pynte på situationen, når der er stor forskel på indikatorerne, modsat intentionen. Det vil således fortsat ændre på forholdet med fuldkommen substitution, men med nogle uheldige konsekvenser.

Af den grund anvender vi ikke det geometriske gennemsnit til at beregne det samlede indeks eller de enkelte områder, men det mere almindelige aritmetiske indeks. Samtidig kommer vi ikke tættere på at løse substitutionsproblematikken. Det er det endnu mere påkrævet, at aflæsning af det samlede indeks ikke kan stå alene for analyse af den sociale situation, men må suppleres ved at se nærmere på udviklingen i de enkelte områder og indikatorer.

Metoden i Det sociale velfærdsindeks – og et supplement

De foregående afsnit viser, at det ikke er simpelt at konstruere et fornuftigt indeks, der ikke gør vold på det datamateriale, der ligger til grund for det. Her følger en opsamling og diskussion af den metode, vi mener er den mest fordelagtige – eller mindst problematiske – til beregningen af det sociale velfærdsindeks. Dernæst uddyber vi en metode, der kan fungere som et supplement, da den har nogle andre kvaliteter, men også ulemper.

En simpel indekseberegning – den historiske udvikling siden 2010

Til at vise udviklingen i sociale problemer over en længere periode er der behov for en metode, der refererer tilbage til et udgangspunkt – en baseline. Vi anbefaler en almindelig indekseberegning, hvor man tager udgangspunkt i et år og sætter værdien for indikatorerne for det år til eksempelvis 100, sådan som man blandt andet gør i aktieindekset c25 og Danmarks Statistiks forbrugerprisindeks. Indikatorerne for de følgende år beregnes så som den procentvise udvikling fra dette år. I SV25 går indikatorerne tilbage til 2010, så for dette år sættes de til 100.

Den almindelige indekseberegning tager ikke forbehold for, at indikatorerne kan variere i volatilitet og skala, sådan som standardiseringer og normaliseringer gør det. Det er imidlertid en ulempe, man kan være nødt til at acceptere for at undgå andre ulemper. Alternativer er nemlig at konstruere en skala på baggrund af gennemsnit og standardafvigelser af eksisterende data (z-score standardisering) frem for at tage ét år som baseline, eller at konstruere en ny skala for alle ved hjælp af min-max normalisering.

Problemet ved førstnævnte, z-score standardisering, er, at den konstruerer en skala på baggrund af observationer fra en relativt kort tidsperiode (2010-2017). Såfremt indikatorer i denne periode har opført sig unormalt volatibelt eller stabilt, så vil udviklingen for de følgende år blive hhv. dæmpet eller forstærket, hvis den adskiller sig fra den foregående periode. Det er svært at sige, hvor sandsynligt dette er, men konsekvenserne med en uigennemsigtig vægtning af enkelte indikatorer vil være u hensigtsmæssig.

Min-max normaliseringen har, som diskuteret tidligere i notatet, den ulempe, at de to yderpunkter fastsættes gennem subjektive vurderinger, fx ved at se på forhold i andre lande, politiske mål eller ud fra en vurdering af eksistensminimum⁴. For nogle indikatorer vil det være muligt at argumentere for sådan nogle yderpunkter i en skalering, men for de fleste vil det være vanskeligt og i høj grad præget af personlige vurderinger.

Således er en simpel, og derved gennemsigtig, indeksering den mest fordelagtige i denne sammenhæng. Lidt groft kan man også sige, at det er 'den mindst ringe'. Som notatet indledte med, så er det at lade ét tal repræsentere 25 sociale og velfærdsmæssige fænomener ikke uden problemer. Ved at lade indikatorerne, områderne og det samlede indekstal repræsentere de faktiske data i videst mulig udstrækning, bliver aflæsning og fortolkning enklere.

Konkret beregnes indekset ved at tage gennemsnittet af de indekserede indikatorer for hvert af de 7 områder, og dernæst tage gennemsnittet af disse 7 områder. Der er således tale om 7 områdespecifikke indekser, der tilsammen udgør det samlede indeks. Det er en selvstændig pointe, at der først beregnes gennemsnit af de enkelte områder, så hvert område bidrager lige meget til det samlede indeks, på trods af at der kan være forskel på antallet af indikatorer i hvert område. Denne fremgangsmåde adskiller sig i øvrigt fra "plus eller minus"-metoden, der har til formål at give et simpelt tal for, hvor mange af de 25 indikatorer der udvikler sig positivt, hvor mange der udvikler sig negativt, og hvor mange der er stabile fra det ene år til det andet.

⁴ I Human Development Index er minimum for midlertidigt fx sat til 20, da man vurderer, at et samfund ikke kan reproducere sig selv hvis midlertidigheden er lavere. (Klugman, Rodríguez og Choi (2011): *The HDI 2010: New controversies, old critiques*. UNDP.)

Plus eller minus - den aktuelle udvikling siden sidste år

Den mest simple måde at opgøre et indeks på er ved udelukkende at skelne mellem positiv, negativ og ingen udvikling. Metoden er, at hver indikator sættes til 1, hvis den er signifikant højere end sidste år, og -1 hvis den er signifikant mindre (95 pct. signifikansniveau). Det er således et almindelig plus-minus beregning af antallet af indikatorer, der går i hhv. den ene eller den anden retning. Indekset vil ligge mellem -25 og 25, hvor et tal tæt på 25 betyder, at de fleste sociale og velfærdsmæssige problemer er steget siden året før, og et højt negativt tal betyder, at de er faldet. 0 indikerer status quo siden året før for de samlede billede af indikatorerne, men der kan naturligvis være bevægelser imellem dem.

Fordelen ved metoden er først og fremmest, at den sikrer gennemsigthed og ensartethed for aflæsning af indekset. Hvor mange andre former for indeksering risikerer at vægte indikatorer forskelligt pga. dataseriens størrelsesorden og volatilitet (tendens til at svinge), så bliver diskussionen af vægt nærmest meningsløs med en enkel kategorisering af indikatorerne til enten plus eller minus. To udviklinger inden for et område vil altid veje lige meget – nemlig 1 – uanset størrelsen.

En indvending mod den simple kategorisering er, at det forstærker substitution imellem indikatorerne – en negativ udvikling vil altid opvejes af en positiv. Omvendt er substitution en uundgåelig faktor i denne type indeks, men hvornår én udvikling kan substituere en anden kan være relativt vilkårlig, da de typisk måler forskellige fænomener. Ved at kategorisere enhver udvikling til hhv. 1 og -1 undgår vi, at et stort udsving i én – måske mere volatil – indikator, kan opveje udviklingen i flere andre indikatorer.

Den simple 'plus-minus' opgørelse af udviklingen i indikatorerne fra ét år til det næste er på flere måder den mest robuste i forhold til at give et ensartet billede af de mange forskellige dataserier. Et problem ved metoden er dog, at vi ikke kan bruge den til at sige noget om udviklingen over længere tid. Eksempelvis vil tre års negativ udvikling først blive opvejet efter tre års positiv udvikling, til trods for at indikatorerne værdimæssigt steg lige så meget på et enkelt år som de faldt på tre år. Metoden kan således kun bruges som mål for, i hvilken retning indikatorerne er gået siden det foregående år. For et billede af de historiske tendenser, hvilket er et af målene med Det sociale velfærdsindeks, må vi således anvende en anden metode.

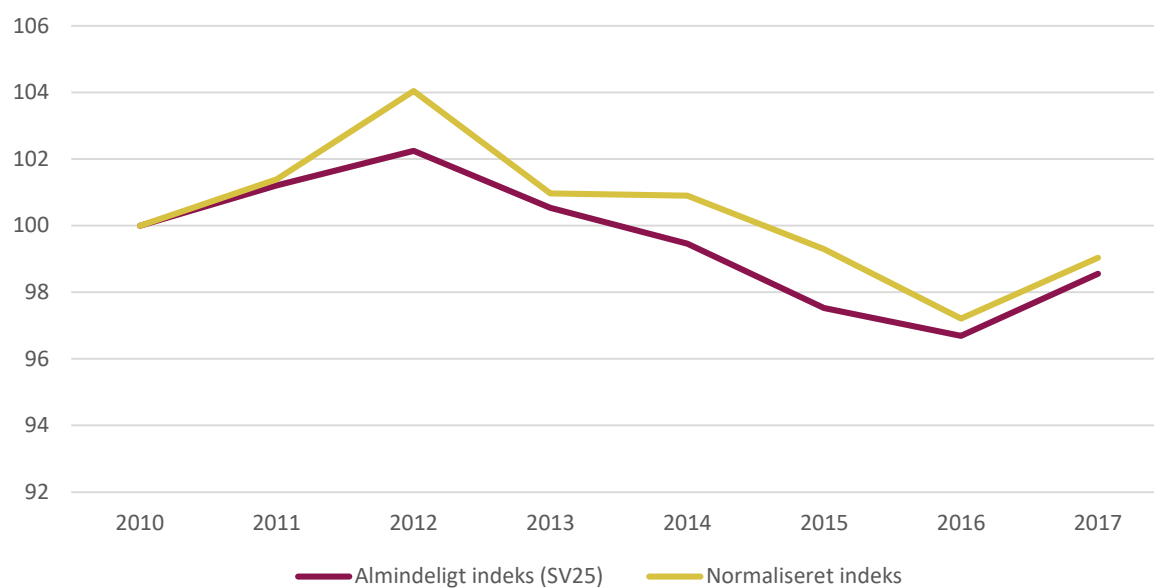
Robusthedstjek – normalisering

Som et robusthedstjek af, hvad den forslåede beregningsmetode betyder for indekset, frem for et valg af en mere kompliceret metode, sammenligner vi her med et normaliseret indeks. Indikatorerne er normaliseret ved at sætte minimum og maximum til hhv. 10 standardafvigelser over og under gennemsnittet for perioden 2010-2017. Det er altså på sin vis både en standardisering og normalisering. Indikatorerne bliver omregnet til alle at ligge mellem 0 og 1, og følsomheden for udsving bliver standardiseret på baggrund af udsvingene i det eksisterende data. Indikatorer med store udsving får et minimum længere fra gennemsnittet end indikatorer med små udsving, hvilket gør at udsving vejer en smule mindre for disse indikatorer, og vice versa.

Med denne metode kan udviklingen fortsat beregnes som en procentvis udvikling fra baseline-året, men hvor indekstallene er standardiseret på baggrund af variationen i 2010-2017.

Den overordnede trend i resultatet er ikke meget forskellig fra den almindelige indekseberegning. De er dog heller ikke ens. Det normaliserede indeks viser sig at være mere følsomt over for udsving i data end det almindelige. Det skyldes sandsynligvis, at udsvingene i nogle indikatorer er blevet forstærket gennem den skalering, som sker ved normaliseringen.

Figur 4: Det almindelige indeks og et normaliseret indeks (2010=100)



Kilde: Egne beregninger

Se bilag for det mere tekniske om normaliseringsmetoden.

Bilag

Uddybende om min-max normalisering

Min-max normalisering er en slags skalering, hvor hver enkelt indikator får en værdi mellem 0 og 1. Metoden anvendes blandt andet i Human Development Index (HDI), dels for at konstruere en fælles skala (mellem 0 og 1), dels fordi skalaen kan konstrueres efter det, der er praktisk muligt eller på baggrund af et subjektivt mål. Det får betydning for, hvor stor procentvis betydning det skal have, når en indikator går et procentpoint op eller ned. Den normaliserede indikator beregnes således:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

For indikatorer der baserer sig på procentdele, er det naturligt at sætte minimum til 0 procent og maximum til 100. I det tilfælde er den eneste effekt af normaliseringen, at indikatoren divideres med 100 – altså flyttes to decimaler. For andre indikatorer kan det give mening at sætte grænserne anderledes. I HDI anvendes minimum som en slags eksistensminimum, hvor middellevetiden må være mindst 20 år for at et samfund skal kunne reproducere sig selv. Inspireret af det, kan man overveje at sætte minimum for kompensationsgraden til det, der svarer til en SU (ca. 30 pct. hvis det er efter skat). Det vil betyde, at en ændring fra 60 til 50 vil være et fald på 33 procent frem for 16 procent, hvis minimum var 0.

Således kan et teoretisk eller politisk funderet minimum få betydning for, hvor meget udviklingen i en indikator skal veje, hvorfor det kræver en særlig grund eller et eksplicit politisk mål at flytte det fra 0. Af netop denne grund har vi som udgangspunkt valgt ikke at anvende min-max normaliseringen i indekskonstruktionen.

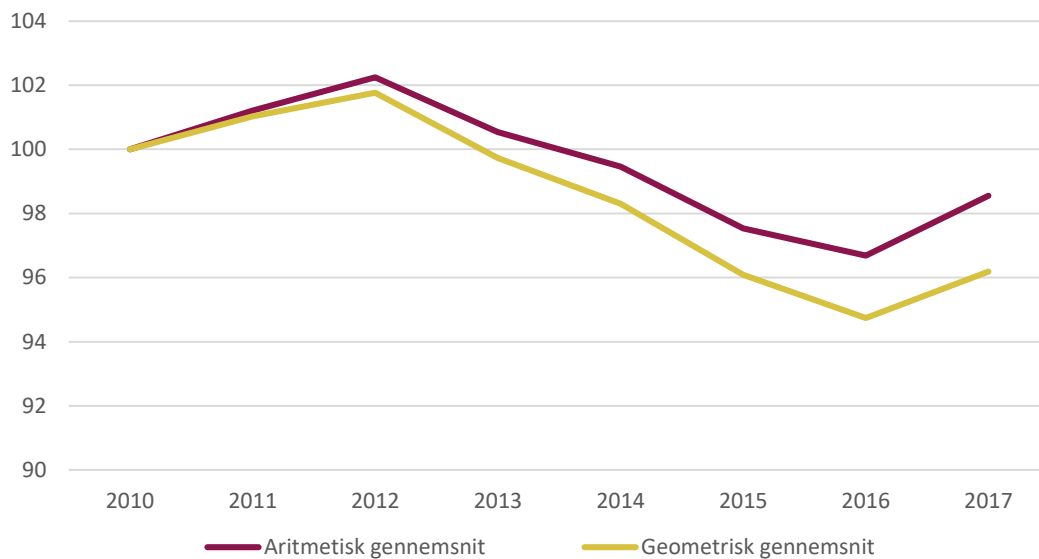
Ændring af maximum svarer til at skalere indikatoren, men det har ingen betydning for den procentvise udvikling. Maximumværdien har ingen betydning for indekset, da det ikke ændrer ved den procentvise udvikling.

Indeksberregning med geometrisk gennemsnit

Anvendelsen af det aritmetiske gennemsnit i det sociale velfærdsindeks betyder, at indekset ligger en smule over det niveau, det geometriske gennemsnit havde resulteret i. Det geometriske gennemsnit trækker værdien ned, hvis der er stor forskel mellem de enkelte indikatorer. Jo større ulighed i indikatorerne og områder, des lavere bliver gennemsnittet. Da vi med SV25 ikke er interesseret i at honorere, at det går fremad op visse områder, men tilsvarende tilbage for andre, er det teoretisk fornuftige valg at anvende det aritmetiske gennemsnit.

Figuren her viser forskellen ved de to typer af gennemsnit.

Figur 5: SV25 2010-2017 (2010=100), beregnet med hhv. aritmetisk og geometrisk gennemsnit

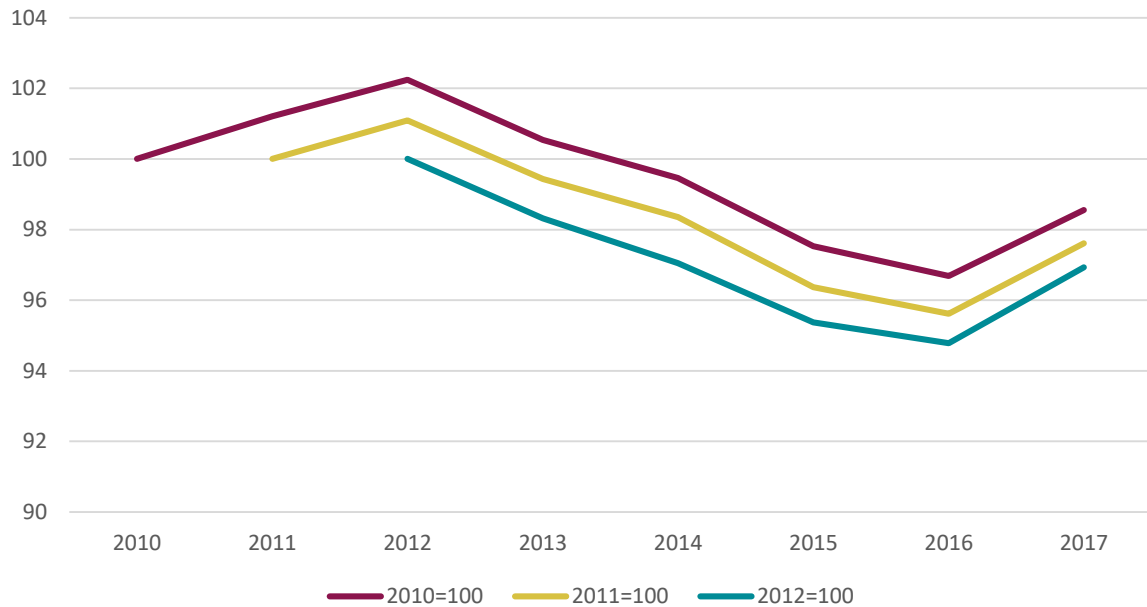


Kilde: Egne beregninger

Indeksberedning ved forskellig baselines

Figur 10 viser, at valget af baseline ikke spiller en væsentlig rolle for resultatet af indekset. Der er dog en lille forskel mellem de forskellige baselines, der tydeligst ses i udviklingen 2016-2017, hvor der er et meget lille fald ved baseline 2010, mens der ingen udvikling er i baseline 2012.

Figur 6: SV25 ved tre forskellige baselines: 2010, 2011 og 2012.



Kilde: Egne beregninger