

## Det standardberegneede lønindeks – supplerende metodebeskrivelse

Dette notat er et supplement til statistikdokumentationen. Notatet beskriver stikprøvedesign og opregning af stikprøven for den private sektor, der sammen med data for de offentlige sektorer danner grundlag for det standardberegneede lønindeks. Desuden uddybes hvordan del- og totalindeks konstrueres på baggrund af basisindeks.

### Stikprøve og opregning

For den offentlige sektor indsamles alle lønforhold, mens der for den private sektor indsamles lønforhold via en stikprøve af virksomheder.

Data til det standardberegneede lønindeks består af individuelle lønforhold fra kommunale, statslige og private løndataleverandører. Bemærk at lønforhold, der kommer fra henholdsvis private og offentlige løndataleverandører, kan godt indgå i andre sektorer.

### Stikprøvedesign

Stikprøven til den private sektor er en stratificeret stikprøve med simpel tilfældig udvælgelse. Stikprøven udtrækkes på virksomhedsniveau (defineret ved CVR-nummer), hvorefter de udtrukne virksomheder indberetter lønoplysninger på samtlige lønforhold i virksomheden (eventuelt via en løndataleverandør). Der stratificeres efter virksomhedens størrelse (baseret på antal ansatte) og branche. Alle virksomheder med mindst 100 beskæftigede samt medlemmer af Finanssektorens Arbejdsgiverforening (FA) indgår i stikprøven (*take all-strata*). Virksomheder med færre end 10 ansatte udtages ikke til stikprøven (*take none-strata*). For øvrige virksomheder med mellem 10 og 100 ansatte udtages en sandsynlighedsbaseret stikprøve (*take some-strata*).

Populationen, hvor stikprøven trækkes fra, er baseret på et udtræk fra Danmarks Statistiks erhvervsstatistiske register (ESR), og er det, der ligeledes danner grundlag for årets lønstrukturstatistik. Populationen fastlægges hvert år i 3. kvartal året før tællingsåret.

Der benyttes et stikprøvedesign med roterende paneler, hvor den samlede stikprøve er sammensat af flere paneler. Hvert panel er en repræsentativ stikprøve, og panelerne udskiftes (roteres ud) over tid. Det betyder, at virksomheder indgår i flere indsamlingsrunder til den pågældende statistik, men udskiftes i takt med at panelerne udskiftes. Fordelen ved et sådant design er primært øget præcision ved estimation af ændringer over tid, men også øget datakvalitet. Endvidere anvendes et SAMU-system<sup>1</sup> til koordinering mellem stikprøver for beslægtede statistikker for at mindske indberetningsbyrden for virksomhederne. Der er således (så vidt muligt)

---

<sup>1</sup> SAMU er en svensk forkortelse for ”samordnad urval”, altså koordineret stikprøveudtagelse.

negativ koordinering mellem stikprøven til *Standardberegnet lønindeks* og stikprøverne til *Øvrige arbejdsomkostninger* henholdsvis *Fravær*.

Stikprøvestørrelsen er årligt på ca. 5.000 virksomheder (inkl. virksomheder i take-all strata).

### Stikprøveopregning til målpopulation

Hvert kvartal er der virksomheder, der ikke er en del af stikprøven, der indberetter frivilligt. Ligeledes er der virksomheder i stikprøven, som ikke indberetter (non-response). Derfor skal designvægtene fra stikprøvedesignet korrigeres, hvilket sker i flere trin.

Først findes designvægten for de virksomheder, der er udtrukket til stikprøven. Disse korrigeres for det antal virksomheder, som ikke har indberettet, samt frivillige indberettere. De korrigerede designvægte beregnes på virksomhedsniveau (CVR-nummer) som:

$$\gamma_{cvr(i)}^k = \frac{N_{stratum(cvr)} - N_{isf, stratum(cvr)}}{n_{stratum(cvr)} - n_{nonresp, stratum(cvr)}}$$

hvor

- $n_{stratum(cvr)}$  er antal virksomheder i  $stratum(cvr)$ , der er udtrukket til stikprøven
- $n_{nonresp, stratum(cvr)}$  er antal ikke-responderende virksomheder i  $stratum(cvr)$
- $N_{stratum(cvr)}$  er det samlede antal virksomheder i  $stratum(cvr)$
- $N_{isf, stratum(cvr)}$  er antallet af virksomheder i  $stratum(cvr)$ , der er ikke er udtrukket til stikprøven, men indberetter frivilligt.

De korrigerede designvægte svarer dermed til de designvægte, man ville få, hvis de virksomheder, der havde indberettet, var de, der var udtrukket til stikprøven. De frivillige indberettere får vægten 1, da de kun repræsenterer sig selv.

De korrigerede designvægte justeres derefter til erhvervsbeskæftigelsen med en korrektionsfaktor,  $K_{(bra, stgrup)(i)}^k$ . Dette gøres ved et ESR-udtræk for det relevante kvartal, hvor korrektionsfaktor beregnes ved at holde beskæftigelsen beregnet ud fra stikprøven op mod erhvervsregisterbeskæftigelsen inden for krydset af branche og størrelsesgruppe (dvs. ESR-udtræk for 1. kvartal til kalibrering af vægte for 1. kvartal samme år). Hvis erhvervsregisterbeskæftigelsen er højere end den beregnede beskæftigelse er  $K_{(bra, stgrup)(i)}^k < 1$ , mens hvis erhvervsregisterbeskæftigelsen er lavere end den beregnede beskæftigelse er  $K_{(bra, stgrup)(i)}^k > 1$ . Korrektionen skal sikre, at virksomheder der ligger i et stratum, hvor den faktiske indberettede beskæftigelse er højere end erhvervsregisterbeskæftigelsen, vægtes ned, og omvendt.

## Konstruktion af indeks

Det standardberegnete lønindeks er konstrueret med inspiration fra prisindeksteori, hvor lønmodtagere, som udfører sammenligneligt arbejde inden for samme branche og sektor, grupperes i såkaldte homogene grupper. Disse grupper danner grundlaget for beregning af basisindeks, som herefter sammenvejes til del- og totalindeks, der er indeks opgjort for sektorer, brancher og hovedarbejdsfunktioner. Nedenfor følger en mere detaljeret beskrivelse af indekseberegningen.

### Basisindeks

Lønmodtagere, som udfører sammenligneligt arbejde, grupperes i homogene grupper defineret ved sektor, branche (36-gruppering), arbejdsfunktioner (53 grupper baseret på over- og mellemgrupper i DISCO-08) samt aflønningsform (time- eller fastlønnet). Disse grupperinger danner grundlaget for beregning af basisindeksene, som er det mest detaljerede niveau i det standardberegnete lønindeks.

Basisindeksene viser stigningen i gennemsnitstimelønnen (vægtet) for hver homogen gruppe siden referencekvartalet (sidste kvartal i forrige vægtår<sup>2</sup>). Beregningen af basisindeksene foregår lidt forskelligt for lønmodtagere i den offentlige sektor og lønmodtagere i den private sektor, da indberetninger fra den private sektor er baseret på en stikprøve, som beskrevet ovenfor. Derfor skal der tages højde herfor ved beregning af basisindeksene, således at stikprøven opregnes til målpopulationen.

Timelønsbegrebet, der anvendes i det standardberegnete lønindeks, er den samlede timefortjeneste (eksklusiv uregelmæssige betalinger) delt med præsterede timer (betalte arbejdstimer forudsat fast fravær). Hvis der sker ændringer i antallet af arbejdsdage eller arbejdstimer vil der korrigeres herfor i timelønsbegrebet. Det vil derfor ikke påvirke indekseberegningen.

I den kommunale, regionale og statslige sektor beregnes basisindeksene som:

$$B_{k_{ref}:k}^H = \frac{\frac{\sum_{i \in H_k} w_i^k t_i^k}{\sum_{i \in H_k} w_i^k}}{\frac{\sum_{j \in H_{k_{ref}}} w_j^{k_{ref}} t_j^{k_{ref}}}{\sum_{j \in H_{k_{ref}}} w_j^{k_{ref}}}}$$

hvor

- $w_i^k$  er den individuelle vægt for lønforhold (person)  $i$ , der angiver hvor stor en andel lønforholdet udgør i forhold til en fuldtidsstilling,
- $t_i^k$  er timelønnen til tid  $k$  for lønforhold  $i$  og
- $H_k$  er individerne i homogene gruppe  $H$  til tid  $k$ .

Basisindekset er altså forholdet mellem den gennemsnitlige timeløn i periode  $k$  og den gennemsnitlige timeløn i referenceperioden  $k_{ref}$  idet gennemsnitstimelønnen for hver periode er beregnet som et gennemsnit af de observerede timelønninger vægtet med deres andel af en fuldtidsstilling.

---

<sup>2</sup> Vægtåret er de fire sammenhængende kvartaler, hvor samme lønsumsvægte anvendes.

I den private sektor beregnes basisindeksene som:

$$B_{k_{ref}:k}^H = \frac{\frac{\sum_{cvr} \sum_{i \in H_k^{cvr}} \gamma_{cvr(i)}^k K_{(bra,stgrup)(i)}^k w_i^k t_i^k}{\sum_{cvr} \sum_{i \in H_k^{cvr}} \gamma_{cvr(i)}^k K_{(bra,stgrup)(i)}^k w_i^k}}{\frac{\sum_{cvr} \sum_{j \in H_{k_{ref}}^{cvr}} \gamma_{cvr(j)}^k K_{(bra,stgrup)(j)}^{k_{ref}} w_j^{k_{ref}} t_j^{k_{ref}}}{\sum_{cvr} \sum_{j \in H_{k_{ref}}^{cvr}} \gamma_{cvr(j)}^k K_{(bra,stgrup)(j)}^{k_{ref}} w_j^{k_{ref}}}}$$

Igen angiver  $w_i^k$  den individuelle vægt,  $t_i^k$  er timelønnen til tid  $k$  for et individ  $i$  og  $H_k$  er individerne i homogene gruppe  $H$  til tid  $k$ .

Nu findes imidlertid yderligere et led for de korrigerede designvægte  $\gamma_{cvr(i)}^k$  samt en korrektionsfaktor  $K_{(bra,stgrup)(i)}^k$  således at stikprøven opregnes til den fulde målpopulation jf. afsnittet om stikprøveopregning til målpopulation ovenfor.

I udtrykket ovenfor betragtes de individuelle timelønninger  $t_i^k$ , men udtrykket kan omskrives til at være et udtryk i gennemsnitlige timelønninger for de forskellige virksomheder, som de individuelle lønforhold er observeret indenfor. Da vil de forskellige gennemsnitlige timelønninger for forskellige virksomheder blive vægtet med mængden af udført arbejde, så en enkelt, mindre virksomhed med afvigende timelønninger ikke kan komme til at dominere den gennemsnitlige timeløn for den homogene gruppe.

Forsimplet kan formlen for basisindeks opskrives som et prisrelativ:

$$B_{k_{ref}:k}^H = \frac{p_H^k}{p_H^{k_{ref}}}$$

hvor  $p_H^k$  er den vægtede gennemsnitstimeløn for homogen gruppe  $H$  i kvartal  $k$ , og  $p_H^{k_{ref}}$  er den vægtede gennemsnitstimeløn for homogen gruppe  $H$  i referencekvartalet  $k_{ref}$ .

Dette er generelt ikke det samme som kvartalsstigningen,  $p_H^k/p_H^{k-1}$ , hvor gennemsnitstimelønnen for den homogene gruppe i et kvartal sættes i forhold til gennemsnitstimelønnen for det umiddelbart foregående kvartal.

Tabel 1 viser et simpelt eksempel på, hvordan basisindeksene beregnes for to homogene grupper.

**Tabel 1** Eksempel: Beregning af basisindeks

	Referencekvartal	1. kvartal	2. kvartal	3. kvartal	4. kvartal
<b>Gennemsnitstimeløn</b>			kr.		
Gns. timeløn, Homogen gruppe 1	204,6	207,1	243,1	230,1	218,9
Gns. timeløn, Homogen gruppe 2	207,8	216,9	222,0	224,6	225,1
<b>Basisindeks</b>			Indeks		
Homogen gruppe 1		101,3	118,8	112,5	107,0
Homogen gruppe 2		104,4	106,8	108,1	108,3
Eksempel på beregning:			$B_{ref:2}^1 = 100 \times \frac{243,1}{204,6} = 118,8$		

### Del- og totalindeks

Når basisindeksene er beregnet sammenvejes de til del- og totalindeks (aggregerede indeks). Dette gøres ved brug af *lønsumsvægte* (sum af løn uden uregelmæssige betalinger). Vægtene udtrykker lønsumsandelene for de homogene grupper, der indgår i det standardberegnete lønindeks, og beregnes på baggrund af de data, der bruges til Lønstrukturstatistikken. Hver homogen gruppe tildeles en vægt som fastholdes i fire kvartaler. Denne vægt sikrer, at den estimerede lønudvikling i mindre grad påvirkes af ændringer i antallet af lønmodtagere i de homogene grupper.

Den faste vægt, der anvendes ved beregning af del- og totalindeks, revideres én gang om året i forbindelse med beregningen af 3. kvartal, hvor data fra sidste års lønstrukturstatistik er tilgængelig.

De aggregerede indeks er vægtede aritmetiske gennemsnit af basisindeksene:

$$I_k = \sum_H \frac{v_H}{v_D} \times B_{k_{ref}:k}^H$$

hvor  $\frac{v_H}{v_D}$  er lønsumsvægten, som angiver lønsumsandelene for hver homogen gruppe  $H$ , i forhold til den samlede lønsum for lønmodtagere i del- eller totalindekset.

*Delindeks* er aggregerede indeks på branche- og arbejdsfunktionsniveau. Delindekset afspejler således udviklingen i den gennemsnitlige timeløn for en given branche eller hovedarbejdsfunktion.

I 2. kvartal ganges delindeksene for den private sektor med en korrektionsfaktor for at tage højde for den lønstigning, der er en effekt af aftaler omkring *fritvalg* og lignende ordninger i de private overenskomster, dvs. delindeks i den private sektor beregnes som

$$I_{k_{ref}:k}^{bra} = \begin{cases} I_{k_{ref}:k}^{bra} F_{\text{år}(k)}^{bra(H)}, & kv(k) = 2 \\ I_{k_{ref}:k}^{bra}, & kv(k) \neq 2 \end{cases}$$

hvor  $F_{\text{år}(k)}^{bra(H)}$  er den korrektionsfaktor der benyttes til at tage højde for *fritvalgs-effekten* indenfor branchen i året  $\text{år}(k)$ .

*Totalindeks* er det mest aggregerede niveau standardberegnet lønindekset opgøres (og offentliggøres) på. Der estimeres et indeks på sektorniveau (privat, stat,

kommuner og region), et indeks for offentlig samlet, kommuner og regioner samlet samt et samlet indeks for alle lønmodtagere i Danmark uanset sektor. Totalindeksene viser således udviklingen i den gennemsnitlige timeløn for en given sektor, offentlig/private virksomheder og den samlede lønudvikling på et overordnet niveau.

Tabel 2 viser et simplificeret eksempel på hvordan (ikke-kædet) del- og totalindeks beregnes.

**Tabel 2** Eksempel: Beregning af del- og totalindeks, ikke-kædet værdier

	År 1				År 2			
	1. kvartal	2. kvartal	3. kvartal	4. kvartal	1. kvartal	2. kvartal	3. kvartal	4. kvartal
<b>Basisindeks</b>					Indeks -----			
Basisindeks 1	101,2	100,8	100,5	101,5	102,4	103,0	103,2	104,1
Basisindeks 2	102,3	101,9	102,2	103,4	104,5	105,0	106,0	106,7
Basisindeks 3	100,1	100,0	100,6	102,4	101,2	101,1	103,2	103,4
Basisindeks 4	99,6	98,2	99,1	100,2	100,5	102,1	104,0	103,9
Basisindeks 5	98,3	99,2	97,8	96,9	98,2	99,3	99,7	101,3
<b>Lønsumsvægte år 1</b>					<b>Lønsumsvægte år 2</b>			
Basisindeks 1	10				Basisindeks 1		10	
Basisindeks 2	5				Basisindeks 2		5	
Basisindeks 3	25				Basisindeks 3		25	
Basisindeks 4	50				Basisindeks 4		50	
Basisindeks 5	10				Basisindeks 5		10	
<b>Delindeks</b>					Indeks -----			
Delindeks 1 (basisindeks 1, 2 og 4)	100,1	98,9	99,6	100,6	101,1	102,5	104,0	104,1
Delindeks 2 (basisindeks 3 og 5)	99,6	99,8	99,8	100,8	100,3	100,6	102,2	102,8
Eksempel på beregning:	$D_{\text{år } 1, k=4}^2 = 102,4 \times \frac{25}{25 + 10} + 96,9 \times \frac{10}{25 + 10} = 100,8$							
<b>Totalindeks</b>	<b>99,9</b>	<b>99,2</b>	<b>99,6</b>	<b>100,7</b>	<b>100,8</b>	<b>101,8</b>	<b>103,4</b>	<b>103,7</b>
Eksempel på beregning:	$T_{\text{år } 1, k=4} = 100,6 \times \frac{10 + 5 + 50}{10 + 5 + 25 + 50 + 10} + 100,8 \times \frac{25 + 10}{10 + 5 + 25 + 50 + 10} = 100,7$							

Hvert år indsamles lønoplysninger fra den offentlige og private sektor for virksomheder med mindst 10 fuldtidsbeskæftigede. Det er disse oplysninger, som de faste lønsumsvægte beregnes på baggrund af. Det er derfor også kun basisindeks, hvortil der kan beregnes en lønsumsvægt (dvs. hvor der er lønmodtagere i den homogene gruppe i den årlige tælling), som indgår i del- og totalindeksene. Samtidig medtages kun basisindeks, som indeholder observationer både i referencekvartalet og i ultimo kvartalet (indeværende kvartal). Dermed indgår kun homogene grupper, hvor der er indberetninger i både referencekvartalet og indeværende kvartal, i de relevante del- og totalindeks.

Nystartede virksomheder indgår først i indeksberegningen når de indgår i vægtgrundlaget. På samme vis vil virksomheder, der er gået konkurs, først udgå, når

de ikke indgår i vægtgrundlaget. I den mellemliggende periode håndteres konkursramte virksomheder på samme vis som virksomheder, der ikke har indberettet, nemlig ved at lignende virksomheder, der har indberettet, får korrigeret deres vægt i stikprøveopregningen.

### Kædning af indeks og beregning af årsstigninger

Det standardberegne lønindeks beregnes som et kædet fastvægtsindeks, hvor vægtgrundlaget opdateres en gang om året. Ved at kæde de enkelte indeksserier sikres, at de viser den korrekte udvikling over tid, selvom der er skiftet vægtgrundlag.

Når hele indeksserien skal beregnes, så beregnes først indeks 100 som gennemsnittet af de første fire kvartaler i 2016. Derefter beregnes den samlede indeksserie ved at kæde alle delindeks for vægtåret på ultimo kvartalet for forrige vægtår. I det standardberegne lønindeks foretages kædningen kun på de aggregerede indeksserier, dvs. del- og totalindeks.

$$I_k = \frac{I_{k, ukædet} \cdot I_{kædet, vægtår-1, ultimo}}{100}$$

På den kædede indeksserie kan årsstigningerne  $stg_{\text{år}}$  beregnes som stigningen i indeksværdi i forhold til samme kvartal året før (altså fire kvartaler før):

$$stg_{\text{år}} = \frac{I_{k, kædet} - I_{k-4, kædet}}{I_{k-4, kædet}}$$

Med udgangspunkt i samme eksempel som vist i tabel 2 viser tabel 3, hvordan en kædet indeksserie beregnes.

**Tabel 3** Eksempel: Beregning af del- og totalindeks (kædet værdier) samt årsstigninger

	År 1				År 2			
	1. kvartal	2. kvartal	3. kvartal	4. kvartal	1. kvartal	2. kvartal	3. kvartal	4. kvartal
<b>Totalindeks</b>	Indeks -----							
Ikke-kædet	99,9	99,1	99,6	100,7	100,8	101,8	103,4	103,7
Kædet	99,9	99,1	99,6	100,7	101,6	102,5	104,1	104,4

Eksempel på beregning:

$$T_{\text{år } 2,4 \text{ (kædet)}} = \frac{103,7 \cdot 100,7}{100} = 104,4$$

**Årsstigninger**

	Pct. -----							
	-	-	-	-	0,9	2,6	3,8	2,9

Eksempel på beregning:

$$stg_{\text{år } 1,4:\text{år } 2,4} = \frac{104,4 - 100,7}{100,7} \times 100 = 2,9 \text{ pct.}$$

Anm.: Referencekvartalet er sat til 4. kvartal.