

Ingvild Johansen og Ragnhild Nygaard

**Skjevheter i prisindeksen for matvarer og
alkoholfrie drikkevarer?**

	Standardtegn i tabeller	Symbol
© Statistisk sentralbyrå, august 2010 Ved bruk av materiale fra denne publikasjonen skal Statistisk sentralbyrå oppgis som kilde.	Tall kan ikke forekomme Oppgave mangler Oppgave mangler foreløpig Tall kan ikke offentliggjøres Null Mindre enn 0,5 av den brukte enheten Mindre enn 0,05 av den brukte enheten Foreløpig tall Brudd i den loddrette serien Brudd i den vannrette serien Desimaltegn	.
ISBN 978-82-537-7911-9 Trykt versjon ISBN 978-82-537-7912-6 Elektronisk versjon ISSN 1891-5906 Emne: 08.02.10		..
Trykk: Statistisk sentralbyrå		...
		;
		0
		0,0
		*
		—
		,

Forord

I de senere årene har det både i Norge og internasjonalt vært en økende interesse for å utnytte strekkodedata ved beregning av prisindeksene generelt og i konsumprisindeksen (KPI) spesielt. Statistisk sentralbyrå (SSB) er ett av svært få statistikkbyråer i verden som utnytter strekkodedata direkte i KPI. Prisindeksen for matvarer og alkoholfrie drikkevarer, som inngår i KPI, er i sin helhet basert på strekkodedata fra det norske dagligvaremarkedet. Dette notatet redegjør for prisindeksen for matvarer og alkoholfrie drikkevarer generelt og for et evalueringssarbeid som er gjennomført av denne prisindeksen.

Det er seksjon for prisstatistikk som er faglig ansvarlig for KPI. Ingvild Johansen og Ragnhild Nygaard har gjennomført evalueringssarbeidet og utarbeidet denne publikasjonen.

Sammendrag

Siden august 2005 er prisindeksen for matvarer og alkoholfrie drikkevarer, som inngår i KPI, i sin helhet basert på strekkodedata fra dagligvarekjedene. Strekkodedata utgjør enorme elektroniske datamengder som gjør det mulig å lage prisindeks på detaljert nivå basert på alle matvarer og alkoholfrie drikkevarer som omsettes på det norske markedet. Med bakgrunn i den store informasjonsmengden som blir tilgjengelig fra strekkodedata er en superlativ prisindeks, som nyttiggjør seg både pris- og omsetningsinformasjon, med månedlig kjeding tatt i bruk på elementært nivå.

Ettersom stadig flere land ønsker å utnytte strekkodedata i KPI, er det nå et økende internasjonalt fokus på å etablere gode internasjonale rutiner og praksis. Hovedinnvendingen mot månedlig kjeding og bruk av både pris- og omsetningsinformasjon for beregning av prisindeks er at såkalt drifting kan oppstå når prisindeks kjeses. Drifting er en skjevhetsform som oppstår i kjedede prisindeks når priser vender tilbake til sitt utgangspunkt mens prisindeksen ikke gjør det. I dette notatet ser vi etter mulig drifting i prisindeksen basert på våre norske data.

For å avdekke mulige skjevheter i prisindeksen er det utarbeidet alternative prisindeks - (uveid) Jevons prisindeks og en helt nyutviklet prisindeks - RYGEKS prisindeks. RYGEKS prisindeks er per definisjon driftfri, men er ennå ikke tatt i bruk av noen statistikkbyråer og må ses på som en eksperimentell prisindeks.

De alternative prisindeksene gir klare indikasjoner på nedadgående skjevheter i prisindeksen for matvarer og alkoholfrie drikkevarer. På bakgrunn av konklusjone fra evalueringssarbeidet vil SSB gjøre endringer i prisindeksen for matvarer og alkoholfrie drikkevarer for å redusere dagens skjevheter.

Innhold

Forord.....	3
Sammendrag.....	4
1. Innledning	6
2. Generelt om matvareindeksen	7
2.1. Utvalg	7
2.2. Beregning	8
2.3. Revisjon.....	10
3. Skjevheter i matvareindeksen	11
3.1. Superlativ prisindeks og drifting.....	11
3.2. Sesongvarer	15
3.3. Manglende prisobserasjoner.....	18
4. Alternative prisindekser.....	21
4.1. Jevons prisindeks	21
4.2. RYGEKS prisindeks.....	22
5. Sammenligninger mellom Törnqvist, Jevons and RYGEKS	24
5.1. Sammenligninger på totalnivå	24
5.2. Sammenligninger på detaljert nivå	25
6. Oppsummering og konklusjon.....	30
Referanser.....	32
Vedlegg A: Klassifisering av matvarer og alkoholfrie drikkevarer, etter COICOP	33
Vedlegg B: Andel som fjernes fra COICOP6-nivå i Jevons prisindeks med "cut-off" metode	39
Figurregister	42
Tabellregister.....	42

1. Innledning

I de senere årene har det både i Norge og internasjonalt vært et økende fokus på å utnytte strekkodedata ved beregning av prisindeks generelt og i KPI spesielt. SSB er ett av svært få statistikkbyråer i verden som utnytter strekkodedata direkte i KPI, og det er ulike løsninger som ligger til grunn både for innhenting og implementering av dataene i de ulike statistikkbyråene. Etter vår kjennskap er det kun Central Bureau of Statistics in Netherlands (CBS) i tillegg til SSB som utnytter både pris- og omsetningsinformasjon i sine KPI beregninger. Swiss Federal Statistical Office (FSO) utnytter strekkodedata som en kilde for prisinnsamling der prisinformasjonen fra strekkodedata erstatter mer tradisjonell datainnsamling. I SSB har vi brukt strekkodedata i KPI helt siden slutten av 90-tallet ved å utnytte prisinformasjon på noen utvalgte representantvarer. I august 2005 utvidet vi bruken av strekkodedata betydelig ved å nyttiggjøre oss pris- og omsetningsinformasjon på alle varer¹ som omsettes i dagligvarebutikkene som inngår i vårt utvalg.

Strekkodedata er elektroniske data som genereres på utsalgsstedet og ved salgstidspunktet. Dataene innhentes fra kjedenes hovedkontorer og inneholder blant annet informasjon om omsatt mengde, gjennomsnittspris i en gitt periode samt varebeskrivelse.

Tradisjonelt ved indeksberegninger har ikke informasjon om omsatt mengde helt ned på produktnivå (mikronivå) vært tilgjengelig. Strekkodedata gir statistikkbyråer større valgmuligheter i beregning av prisindeks gjennom å kunne ta hensyn til både pris og volum. Samtidig gir det oss nye utfordringer, blant annet hvordan man skal aggregere disse enorme datamengdene på best mulig måte. Ved valg av metode spiller også det overordnede formålet med KPI en viktig rolle.

Flere prinsipielle og metodiske valg ble tatt i prosessen med å innføre strekkodedata i full skala i den norske delindeksen for matvarer og alkoholfrie drikkevarer i 2005. Noen av disse var nokså opplagte, mens andre igjen var mer kontroversielle. Kanskje det mest kontroversielle valget vi gjorde var å innføre månedlig kjedede superrelative prisindeks på det mest detaljerte beregningsnivå for alle varegrupper inklusive volatile varer. Den store fordelen med månedlig kjeding er at det muliggjør rask oppdatering av alle tilganger og avganger av varer. Ettersom utskiftshastigheten på varer på strekkodenivå (EAN-koder) er høy, er det viktig å raskt fange opp den løpende varesammensetningen. Internasjonalt har det riktignok vært stor skepsis til å bruke månedlig kjeding for varer som over tid svinger mye i pris og volum. Det vises til at det ved bruk av superrelative prisindeks, som nyttiggjør seg både pris- og mengdeinformasjon, kan oppstå såkalt drifting² når prisindeksene kjedes. Drifting er en skjevhetsform som oppstår i kjedede prisindeks når priser vender tilbake til sitt utgangspunkt mens prisindeksen ikke gjør det. Desto kortere kjedeperioden er, desto større blir skjevheten.

Høsten 2009 ble det startet et evalueringssarbeid av prisindeksen for matvarer og alkoholfrie drikkevarer. Det var flere grunner til at vi igangsatte et slik prosjekt. Det er en pågående internasjonal debatt rundt drifting og bruk av superrelative prisindeks. Det er nå en økende interesse blant statistikkbyråene for strekkodedata i KPI, men det er samtidig ikke mye erfaring å støtte seg til. Matvarer og alkoholfrie drikkevarer er også en relativ viktig konsumgruppe som vektmessig utgjør om lag 11 prosent av den norske KPI. Et av formålene i dette evalueringssarbeidet har vært å kontrollere beregningsmetoden i prisindeksen for matvarer og alkoholfrie drikkevarer³ for drifting på bakgrunn av den internasjonale debatten som pågår.

¹ Innen matvarer og alkoholfrie drikkevarer.

² Den engelske betegnelsen er "chain drift" eller "chain link bias".

³ Herunder også kalt matvareindeksen.

Rapporten dokumenterer det evaluatingsarbeidet vi har gjort av beregningsmetoden i matvareindeksen. I kapittel 2 redegjøres det for dagens beregnings- og revisjonsmetode. Kapittel 3 ser nærmere på drifting, sesongvarer og manglende prisobservasjoner. I kapittel 4 redegjør vi for alternative prisindeksler mens vi i kapittel 5 undersøker for mulige skjevheter i den norske matvareindeksen ved hjelp av disse alternative prisindeksene. Vi oppsummerer i kapittel 6.

2. Generelt om matvareindeksen

2.1. Utvalg

Tilgangen på strekkodedata gir oss muligheter til å beregne prisindeksen på matvarer og alkoholfrie drikkevarer basert på tilnærmet alle varer som omsettes på det norske dagligvaremarkedet. Hver måned mottas om lag 300 000 prisobservasjoner fordelt på om lag 14 000 forskjellige varer innenfor matvarer og alkoholfrie drikkevarer.

Tabell 1. Antall prisobservasjoner⁴ per måned i matvareindeksen etter COICOP⁵

COICOP	Antall priser
0111: Brød og kornprodukter	50 031
0112: Kjøtt	29 464
0113: Fisk	14 861
0114: Melk, ost og egg	34 785
0115: Oljer og fett	5 566
0116: Frukt	11 321
0117: Grønnsaker	27 670
0118: Sukker, syltetøy mv.	43 440
0119: Andre matvarer	59 144
0121: Kaffe, te og kakao	9 137
0122: Mineralvann mv.	25 142
Totalt	310 561

Det norske dagligvaremarkedet kan karakteriseres med høy kjedekonsentrasjon og SSB mottar strekkodedata fra alle aktørene på det norske dagligvaremarkedet. I følge Dagligvarerapporten 2010 fordeler dagligvaremarkedet seg på følgende måte (holder spesialforretninger og torghandel utenfor); Norgesgruppen 40,0 prosent, Coop Norge 24,0 prosent, Rema 1000 20,3 prosent og ICA Norge 15,7 prosent. Til sammen dekker disse fire store aktørene 100 prosent av det norske dagligvaremarkedet.

Generelt i KPI benyttes et utvalg av representative bedrifter for å estimere prisutviklingen på varer og tjenester. Bedriftsutvalget blir trukket fra en bedriftspopulasjon⁶ og trekkes etter *PPS-prinsippet (Probabilities proportional to size)*, dvs. trekking med hensyn på størrelse hvor omsetning er valgt som størrelsesmål. Populasjonen stratifiseres etter næring og geografisk område. Når det gjelder utvalgsplanen for trekking av bedrifter til matvareindeksen stratifiseres populasjonen etter butikkjede og kjedekonsept/profil. Analyser som ble gjort i forbindelse med omleggingen i 2005 tyder på at kjedeprofil har større betydning for prisutviklingen enn hvor bedriften befinner seg geografisk. Konseptene stratifiseres etter kiosk, lavpris, supermarked og nærbutikk. Lavprissegmentet framstår som et konsept for hverdagsinnkjøp med begrenset vareutvalg og relativt lave priser mens for eksempel supermarkedene profileres som et konsept med større mangfold av varer, høyere servicenivå samt høyere fokus på ferskvarer. Utvalgstørrelsen er på om lag 150 butikker og domineres av supermarked- og lavprissegmentet. Utvalget, som ble etablert i 2005, er basert på såkalt Neymans allokeringsmetode og er trukket ”en-

⁴ Antall prisobservasjoner er et gjennomsnitt av alle prisobservasjonene i perioden oktober 2008 til september 2009.

⁵ COICOP-Classification of individual consumption by purpose: konsumklassifisering utarbeidet av FN. Klassifiseringskriteriet er sluttformålet med konsumet.

⁶ SSBs Bedrifts- og foretaksregister (BOF).

kelt tilfeldig uten tilbakelegging”, se Henriksen (2006). Per juni 2010 er en ny utvalgsplan for matvareindeksen utarbeidet som innebærer mer detaljert stratifisering, økt utvalgstørrelse samt trekking av bedrifter etter proporsjonal allokering etter omsetning.

For matvareindeksen innhentes strekkodedata via hovedkontorene til aktørene innen det norske dagligvaremarkedet. Trekkpopulasjonen utgjør de butikker inn under disse kjedene som har elektronisk rapportering til hovedkontorene. Andre forretninger som spesialforretninger og torghandel holdes utenfor trekkpopulasjonen. Disse forretningene utgjør en liten del av den totale omsetningen innen nærings- og nytelsesmidler. SSBs Varehandelsstatistikk for 2007 viser at spesialforretninger og torghandel står for om lag 10 prosent av denne omsetningen. Vi har heller ingen indikasjon på at prisutviklingen i disse spesialforretningene vil være vesentlig forskjellig fra kjedebutikkene.

Bedriftene i KPI utgjør normalt et panelutvalg der hver bedrift som blir trukket ut skal være med i undersøkelsen i en 6 års periode. Hvert år blir utvalget ajourført og 1/6 av utvalget utgår. For bedriftsutvalget til matvareindeksen er det etablert bedriftsrulleringsrutiner for å fange opp eventuelle vridninger i konsumentenes forbruksmønster/skift mellom kjeder og konsepter på linje med den generelle utvalgsplanen for KPI.

2.2. Beregning

Den publiserte KPI har basis i 1998 (1998=100). Den består av en årlig kjedet indeks basert på såkalte korttidsindeksene. Korttidsindeksene representerer prisutviklingen for en periode på 12 måneder. For korttidsindeksene benyttes juli hvert år som basis (juli=100), hvilket innebærer at gjennomsnittlig pris for hver vare i juli benyttes som prisreferanse fra og med indeksen for august år t til og med indeksen for juli år $t+1$. Utviklingen i korttidsindeksene er dermed kun sammenlignbar innenfor samme korttidslink.

Korttidsindeksene på det mest detaljerte nivået, mikronivå, fremkommer ved å beregne et geometrisk gjennomsnitt av prisobservasjonene for en gitt representantvare innen hvert av de åtte markedene som inngår i KPI i måned m , i forhold til tilsvarende gjennomsnitt i prisreferanseperioden θ . Ved å benytte et geometrisk gjennomsnitt tar man hensyn til en antatt substitusjonseffekt mellom de ulike utsalgsstedene. Generelt i KPI er ikke informasjon om omsetning for hver representantvare i det enkelte utsalgssted tilgjengelig. Derfor beregnes et uveid geometrisk gjennomsnitt, en såkalt Jevons prisindeks, av prisobservasjonene fra de ulike utsalgsstedene innenfor hvert marked. Markedet veies deretter sammen til en indeks for representantvaren ved hjelp av områdevekter. Aggregering til gruppe-, konsumgruppe- og totalindeksnivå er basert på Laspeyres formel, der utgiftsandelene som benyttes i beregningen er basert på Forbrukssundersøkelsen (FU)⁷.

Strekkodedataene for matvarer og alkoholfrie drikkevarer gjør det mulig å beregne indeksene ved å også ta hensyn til omsetningsdata. Strekkodedataene inneholder informasjon om gjennomsnittspris for månedenes midtuke definert som total omsetning delt på total mengde, mengde, utsalgssted, tidspunkt samt en kort beskrivelse av produktene per EAN-/PLU-kode⁸ og per utsalgssted.

I beregningen av mikroindeksene for matvarer og alkoholfrie drikkevarer basert på strekkodedata benyttes en vektet Jevons prisindeks. Utgangspunktet er at beregningsmetoden skal være så lik som mulig den metoden som ble benyttet tidligere, men tilpasset de mulighetene som strekkodedata gir. EAN-/PLU-kodene er koplet

⁷ Fra januar 2011 vil utgiftsandelene baseres på tall fra nasjonalregnskapet.

⁸ EAN (European Article Number) er en standardisert identifiseringskode som gjelder for et produkt, og som er felles uavhengig av utsalgssted. PLU, som står for ”Product Look-Up” eller ”Price Look-Up”, er også en identifiseringskode for et produkt, men som kun er felles innenfor en kjede.

til en COICOP6-gruppe, som representerer en gruppe av homogene varer som har til hensikt å tjene samme formål. Beregningene på mikronivå basert på en vektet Jevons prisindeks er gitt ved:

$$(1) \quad \ln P_K(p^0, p^1, s^i) \equiv \sum_{m=1}^M s_m^i \ln \frac{p_m^1}{p_m^0}$$

Dvs. logaritmen av indeks P for varene som inngår i COICOP6-gruppe K er en funksjon av prisene i basisperiode p^0 , tellingsperiode p^1 , og omsetningsandel s i periode i ; $i=0, 1$ (dvs en kombinasjon av begge periodene) per EAN-/PLU-kode og utsalgssted.

Merk at (1) kommer frem ved å ta logaritmen av:

$$(2) \quad P_K(p^0, p^1, s^i) \equiv \prod_{m=1}^M \left(\frac{p_m^1}{p_m^0} \right)^{s_m}$$

Under forutsetning av

$$(3) \quad \sum_{m=1}^M s_m^i = 1$$

er (2) den sanne elementære prisindeksen sett fra en økonomisk tilnærming, se Pollak (1989). Derfor er valget av (1) for å beregne mikroindeks støttet av både indeksteorien og dagens rutiner for beregning av KPI. Merk ellers at (3) stiller krav om å bruke prisobservasjoner for alle varene som inngår i aggregat K, se Rodriguez og Haraldsen (2005). I beregningen benyttes en kombinasjon av omsetningsandelen fra basis- og tellingsperioden der begge perioder blir behandlet symmetrisk noe som genererer en Törnqvist superlativ prisindeks.

Formålet til KPI er å måle utviklingen i levekostnadene til private husholdninger. En levekostnadsindeks skal gi svar på spørsmålet om hvor mye utgiftssummen må endres når prisene endrer seg fra ett tidspunkt til et annet gitt samme levestandard (nyttenivå). I motsetning til en indeks basert på en fast varekurv tar den ideelle levekostnadsindeksen hensyn til at den konsumerte mengden er avhengig av prisene. Rasjonelle konsumenter antas å endre den relative mengden de konsumerer i forhold til endringen i de relative prisene. En levekostnadsindeks legger til grunn at konsumenten minimerer kostnadene for å opprettholde et gitt nyttenivå. Det beste målet på en levekostnadsindeks kan derfor sies å være en superlativ indeks, se kapittel 17 i ILO manualen (2004). Men i praksis er informasjonen for å beregne en superlativ indeks sjeldent tilgjengelig på det tidspunktet, kun retrospektivt. Derfor beregnes det vanligvis et uveid gjennomsnitt av prisobservasjonene på mikronivå. Strekkodedata gjør det derimot mulig å beregne indeks der man også tar hensyn til endringen i mengden. En karakteristisk egenskap ved en superlativ indeks er at den behandler priser og mengder i begge perioder symmetrisk.

Korttidsindeksene beregnes med utgangspunkt i et vektet geometrisk gjennomsnitt i løpende periode i forhold til forrige måned. Forrige måned er i matvareindeksen benyttet som prisreferanse måned i motsetning til juli som benyttes ellers i KPI. Årsaken er at utskiftningshastigheten på varer på EAN-nivå er relativt høy. Ved å benytte løpende vekter kan endringer i forbruksmønsteret bli tatt hensyn til fortløpende.

I tabell 2 vises en oversikt over antall prisobservasjoner, i prosent, som er felles med basisperioden etter 1, 6 og 12 måneder. Tabellen illustrerer utskiftningshastigheten på varer i 2003 og er hentet fra Rodriguez (2004). Ved å benytte forrige måned som prisreferanse måned vil om lag 73 prosent av prisobservasjonene ha sam-

me EAN-kode i de to periodene. Settes derimot basisperioden til ett år tidligere er det kun 42 prosent av observasjonene med samme EAN-kode.

Tabell 2. Antall prisobservasjoner i prosent som er felles med basisperiode etter 1, 6 og 12 måneder. 2003

COICOP	Felles etter 1 mnd.		Felles etter 6 mnd.		Felles etter 12 mnd.	
	Omsetning i %	Priser i %	Omsetning i %	Priser i %	Omsetning i %	Priser i %
01: Matvarer og alkoholfrie drikkevarer	83,6	73,1	65,6	54,6	53,5	42,0
011: Matvarer	82,5	71,5	65,9	53,3	54,4	42,6
012: Alkoholfrie drikkevarer	84,4	74,7	68,2	56,0	54,8	41,2
0111: Brød og kornprodukter	86,3	74,0	70,2	56,8	53,6	42,9
0112: Kjøtt	80,2	71,1	60,5	51,5	46,6	37,0
0113: Fisk	79,5	69,5	62,4	51,5	46,7	36,6
0114: Melk, ost og egg	84,9	76,5	75,1	60,3	62,7	50,4
0115: Oljer og fett	89,7	77,8	75,1	60,3	65,3	53,5
0116: Frukt	71,2	59,0	56,5	39,9	56,4	41,8
0117: Grønnsaker	84,2	70,4	68,8	53,9	53,7	42,8
0118: Sukker, syltetøy mv.	86,6	75,7	63,2	54,5	57,0	42,9
0119: Andre matvarer	81,8	67,1	67,3	53,6	55,9	42,6
0121: Kaffe, te og kakao	77,9	80,0	53,8	58,5	48,9	44,8
0122: Mineralvann mv.	83,2	78,2	64,7	57,9	49,1	40,6

Fra COICOP6-nivå aggregeres indeksene videre opp til 5-, 4-, 3-, og 2-sifret COICOP-nivå samt totalindeks sammen med representantvarer for varer og tjenester som representerer resten av KPI basert på Laspeyres formel med årlige vektskifter. Vektandelene fra 5-sifret COICOP er basert på FU, mens vektandelene mellom COICOP6-gruppene er basert på omsetningsandeler for de ulike varegruppene hentet fra strekkodedataene.

Tolkningen av prisindeksen for matvarer og alkoholfrie drikkevarer er mer komplisert enn de fleste andre indeksene som inngår i KPI grunnet valg av beregningsmetode. Ren prisindeks som KPI generelt er bygget på, er enklere å tolke ettersom de kun tar hensyn til utviklingen i de relative prisene. Ettersom prisindeksen for matvarer og alkoholfrie drikkevarer også tar hensyn til omsetning må endringene i indeksen tolkes som en kombinasjon av endringer i priser og omsetning.

2.3. Revisjon

Prisene som inngår i beregningsgrunnlaget til KPI gjennomgår flere kontroller. Først kontrolleres dataene for *åpenbare feil* som blant annet tierfeil og punsje feil. Deretter kontrolleres de for *mulige feil*, ved at observasjonene utifra gitte kriterier blir definert som ekstremverdier. Metoden som benyttes for å definere ekstremverdier i KPI er en kombinasjon av HB (Hidiroglo-Berthelot) og Chebyshevs metode. Kontrollene som benyttes i revisjonsprosessen i KPI foretar ingen automatiske korrigeringer av datamaterialet. Alle korrekjoner foretas/godkjennes av revisor.

Revisjonen av matvareindeksen er derimot basert på både automatiske og manuelle kontroller. Først fjernes observasjonene som oppfattes som åpenbare feil⁹ (F-verdier) fra beregningsgrunnlaget. Grunnet den enorme mengden prisobservasjonene fjernes observasjonene i stedet for å kontrollere dem mot kilden. Alle måned-til-måned prisrelativer (P_t/P_{t-1}) mindre eller lik 0,33 og større eller lik 3 oppfattes som åpenbare feil og slettes fra beregningsgrunnlaget. Dette gjelder imidlertid svært få observasjoner, kun om lag 0,3 promille av prisobservasjonene.

Typiske sesongvarer fjernes også fra beregningsgrunnlaget. Dette er varer som på forhånd er definert med sterkt sesongmønster ved at de er på markedet noen få måneder og deretter forsvinner.

⁹ Observasjonene som klassifiseres som åpenbare feil kan vise seg å ikke være det. Eksempelvis kan reelle prisendringer på frukt og grønnsaker overstige disse revisjonsgrensene.

Deretter defineres såkalte kritiske observasjoner. Dette er observasjoner som påvirker resultatet på elementær-nivå mest. Bidraget fra hver prisobservasjon beregnes basert på både prisendring og omsetningsandeler i begge perioder. Deretter beregnes gjennomsnitt og standardavvik for alle bidragene innenfor en COICOP6-gruppe. Observasjoner som ligger utenfor intervallet gjennomsnitt +/- 5 ganger standardavviket defineres som kritiske verdier. Disse verdiene fjernes ikke automatisk fra datamaterialet, men vurderes av revisor. Om lag 1 600 prisobservasjoner, det vil si om lag 0,5 prosent av beregningsgrunnlaget, blir kodet som kritiske verdier hver måned. Etter en gjennomgang av de kritiske verdiene for COICOP6-gruppene med størst endring er det kun et fåtall av observasjonene som fjernes fra beregningsgrunnlaget.

3. Skjevheter i matvareindeksen

3.1. Superlativ prisindeks og drifting

De store datamengdene med både pris- og omsetningsdata gjør at strekkodedata gir oss muligheter for å fange opp konsumentenes løpende forbruksmønster. Samtidig gir mengden av strekkodedata nye utfordringer i hvordan man skal aggregere. Det pågår nå en internasjonal debatt om hvordan man best mulig skal gjøre dette. Hypsig kjeding i kombinasjon med superlativ prisindekser kan gi opphav til skjevheter i tilfeller hvor både priser og volum varierer sterkt over tid. Årsaken til dette er drifting som kan oppstå når indekser kjedes. Drifting er skjevheter som oppstår i kjedede prisindekser når priser er tilbake til sitt utgangspunkt mens prisindeksen ikke er det, se kapittel 20 i ILO manualen (2004). Dersom kjedeperioden er kort, altså at det er kort tid mellom beregningsmåneden og en løpende prisreferanse-måned, kan ulike indekser i to etterfølgende måneder forekomme til tross for at prisen er helt identiske. Jo kortere måleperioden er, jo større drifting kan oppstå i motsetning til en prisindeks med fast basis der man ikke har slike skjevheter.

Kjedede prisindekser kan tendere til drifting på grunn av store svingninger i pris og volum blant annet når varer er på tilbud. Norske strekkodedata viser tydelig at om varer faller sterkt i pris fører dette ofte til en sterk økning i omsetning. Ivancic, Fox og Diewert (2009) påpeker at sammenlignet med prisindekser med fast basis kan kjedede superlativ indekser tendere til å *drifte nedover* når varer settes på tilbud. Når prisen på en vare er tilbake på sitt ”normale” nivå forventer vi at den kjedede prisindeksen også skal være det. Drifting kan derimot oppstå om det tar en viss tid før omsetningen normaliserer seg. Når en vare er tilbake på prisnivået før salg kan det være at konsumenten har hamstret inn så mye av varen at omsetningen en periode er lavere enn den var før varen var på salg og det tar tid før man igjen handler normalt. Triplett (2003) mener at disse omsetningsskiftene (på grunn av tilbudsaktivitet) forårsakes av to typer av konsumenter; de som bare handler når varene er på tilbud og de som hamstrar opp og fyller lagrene sine når varene er på salg. Denne ”hamstringseffekten” der konsumenter hamstrar og fyller sine lagre bryter riktig nok med forutsetningen i KPI om at varer antas konsumert umiddelbart ved kjøp og at kjøpstidspunkt og tidspunkt for konsum er sammenfallende.

De Haan og van der Grient (2009) forklarer driftingproblematikken på følgende måte;

La p_i^0 og s_i^0 stå for pris og omsetningsandel for vare i i basisperioden 0 mens p_i^t og s_i^t står for tilsvarende verdier i periode t ($t > 0$). For en fast kurv av varer U er Fisher og Törnqvist prisindeks definert som

$$(4) \quad P_F^{0t} = \left[\frac{\sum_{i \in U} s_i^0 (p_i^t / p_i^0)}{\sum_{i \in U} s_i^t (p_i^t / p_i^0)^{-1}} \right]^{1/2}$$

$$(5) \quad P_T^{0t} = \prod_{i \in U} (p_i^t / p_i^0)^{(s_i^0 + S_i^t)/2}$$

Vi innfører tre perioder; 0, 1 og 2. Kjedet Fisher og Törnqvist prisindeks fra periode 0 til periode 2 kan skrives som

$$(6) \quad P_{F,kjedeindeks}^{02} = \left[\frac{\sum_{i \in U} s_i^0 (p_i^1 / p_i^0)^1}{\sum_{i \in U} s_i^1 (p_i^1 / p_i^0)^{-1}} \right]^{1/2} \left[\frac{\sum_{i \in U} s_i^1 (p_i^2 / p_i^1)^1}{\sum_{i \in U} s_i^2 (p_i^2 / p_i^1)^{-1}} \right]^{1/2}$$

$$(7) \quad P_{T,kjedeindeks}^{02} = \prod_{i \in U} (p_i^1 / p_i^0)^{(s_i^0 + S_i^1)/2} \prod_{i \in U} (p_i^2 / p_i^1)^{(s_i^1 + S_i^2)/2}$$

Vi antar at vare 1 er på salg i periode 1 og prisen har falt ($p_1^1 / p_1^0 < 1$), i periode 2 er prisen tilbake til sin "normale" verdi ($p_1^2 = p_1^0$ eller $p_1^2 / p_1^1 = p_1^0 / p_1^1$). Vi antar at prisene på alle andre varer ligger fast. Uttrykkene (6) og (7) kan dermed uttrykkes som

$$(8) \quad P_{F,kjedeindeks}^{02} = \left[\frac{s_1^0 \{ (p_1^1 / p_1^0) - 1 \} + 1}{s_1^2 \{ (p_1^1 / p_1^0) - 1 \} + 1} \right]^{1/2}$$

$$(9) \quad P_{T,kjedeindeks}^{02} = (p_1^1 / p_1^0)^{(s_1^0 - S_1^2)/2}$$

Om prisene faller for så å gå tilbake til utgangspunktet forventer vi at mengde solgt og derfor omsetningsandelene vil returnere til utgangsverdiene ($s_1^2 = s_1^0$) slik at

$$P_{F,kjedeindeks}^{02} = P_{T,kjedeindeks}^{02} = 1.$$

Vi ser imidlertid at ulike "forvridninger" kan skape forskjeller mellom s_1^0 og s_1^2 . I dette eksemplet har vi at

$$P_{F,kjedeindeks}^{02} < 1 \text{ og } P_{T,kjedeindeks}^{02} < 1 \text{ for } s_1^2 < s_1^0$$

og

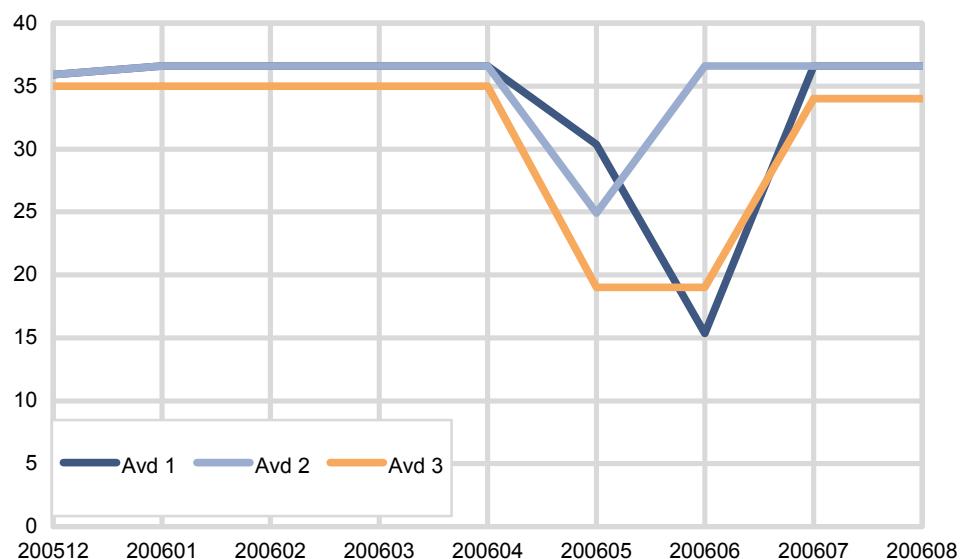
$$P_{F,kjedeindeks}^{02} > 1 \text{ og } P_{T,kjedeindeks}^{02} > 1 \text{ for } s_1^2 > s_1^0.$$

Eksemplet ovenfor er ikke et helt typisk forløp i norske strekkodedata. Vi finner ingen klar systematikk i at omsetningen er lavere i måneden etter tilbud enn måneden før tilbud noe som gjør at denne "hamstringseffekten" ikke er særlig synlig i norske data. Det er heller ikke alltid slik at priser henter seg inn igjen til utgangspunktet i etterfølgende periode. Norske data viser at lengden på salgsperioden kan ha stor betydning. Et forløp vi finner igjen er at prisen holdes lav over flere måltidspunkter (midtuken i hver måned); I periode 0 selges vare 1 til normalpris og

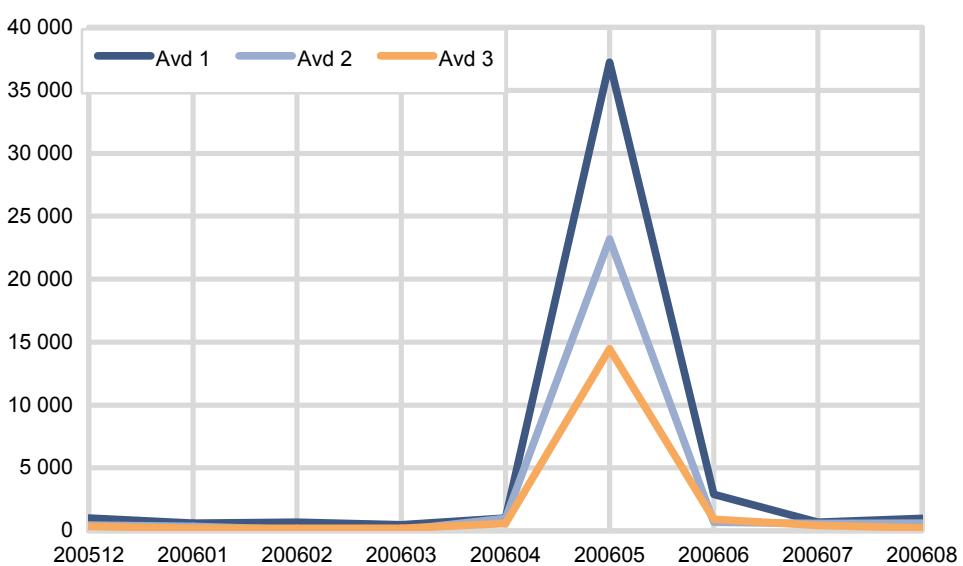
mengde solgt er lav eller nesten lik null. I periode 1 selges varen til en lav tilbudspris og mengde solgt er svært høy. I periode 2 er det fortsatt tilbudsaktivitet, men nå er mengdetallene lave, men ikke så lave som i periode 0. I periode 3 er både pris og mengde solgt tilbake til normalnivåer. Prisnedgangen i periode 1 får høy vekt mens prisoppgangen i periode 3 kun får liten vekt noe som vil føre til at

$P_{T,kjedeindeks}^{03}$ vil være lavere enn utgangsverdien før salg og vi får nedadgående drifting¹⁰. I figur 1 og 2 vises pris- og omsetningsmønsteret for rømmegrøt i 3 ulike avdelinger i perioden desember 2005 til desember 2006. Alle de tre avdelingene setter varen på tilbud i mai 2006 med sterk vekst i omsetningen.

Figur 1. Rømmegrøt 900g. pris i kroner. Desember 2005-desember 2006



Figur 2. Rømmegrøt 900g. Omsetning i kroner. Desember 2005-desember 2006

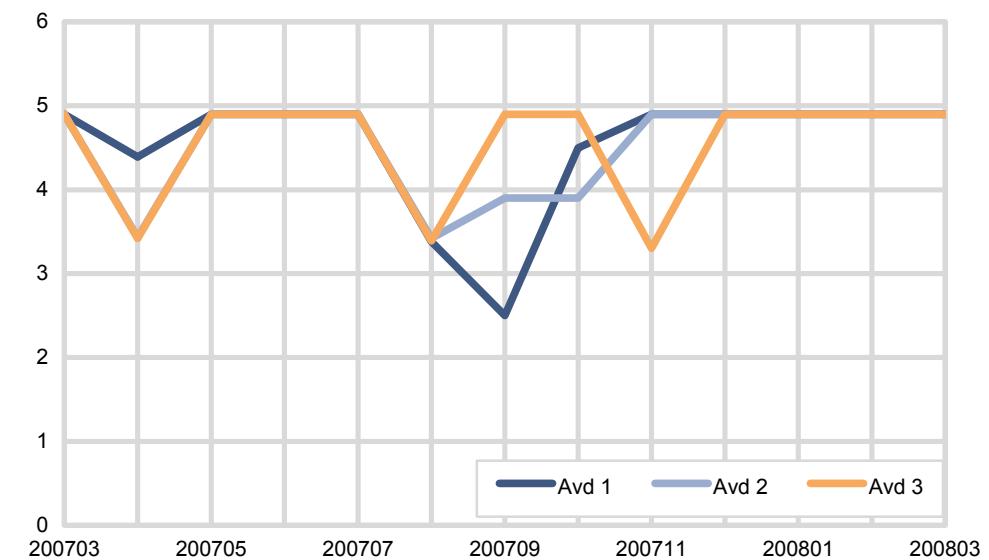


I alle avdelingene faller omsetningen sterkt i måneden etter, også for avdeling 1 der prisen faller ytterligere og i avdeling 3 som viderefører tilbudsprisen. Med et slikt pris- og omsetningsmønster får vi asymmetri i omsetningsvektene der prisnedgangen veies sterkere enn prisoppgangen i prisindeksen.

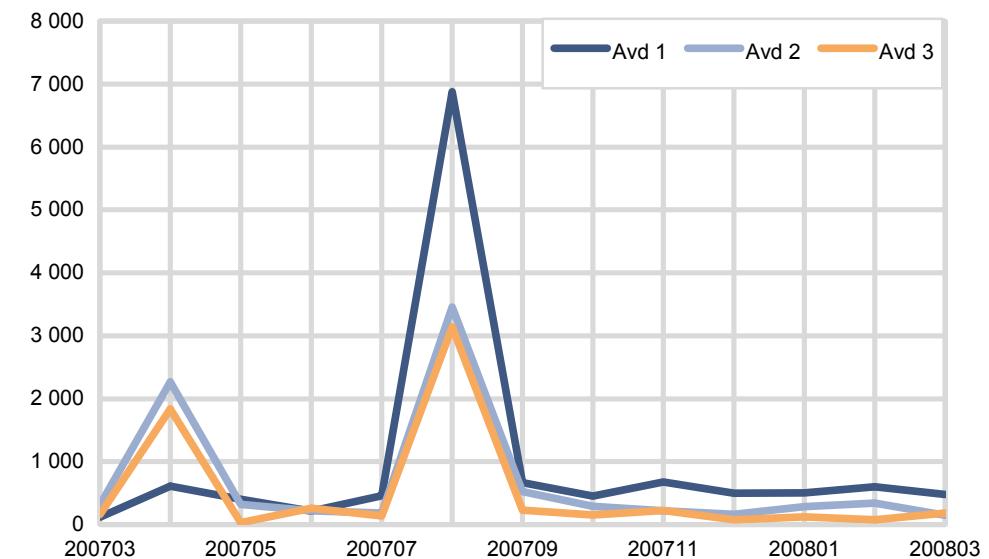
¹⁰ I den internasjonale litteraturen omtales dette som ”downward drift”.

I figur 3 og 4 er pris- og omsetningsmønsteret på skivet sjampinjong i boks i 3 ulike avdelinger illustrert. I dette eksemplet faller prisen i august 2007 i alle de tre avdelingene og konsumentene responderer raskt med høye omsetningstall. I avdeling 3 henter prisen seg inn igjen måneden etter og i avdeling 2 stiger prisen kun delvis tilbake til ”normal” nivå. I avdeling 1 faller prisen ytterligere i måneden etter, men som i de to andre avdelingene faller omsetningstallene kraftig. Dermed vil vi også i dette eksemplet få nedadgående drifting der prisnedgang vektes kraftigere enn prisoppgang.

Figur 3. Skivet sjampinjong i boks. pris i kroner. Mars 2007-mars 2008



Figur 4. Skivet sjampinjong i boks. Omsetning i kroner. Mars 2007-mars 2008



Det er ikke alle varer som er potensielle ”hamstringsvarer”, frukt er et eksempel på det. I internasjonale studier er toalettpapir og bleier eksempler på varer som konsumentene kan tendere til å hamstre. Den norske KPI inneholder også en del strekkodedata på andre produkter enn matvarer og alkoholfrie drikkevarer, men for disse utvalgte produktene anvendes kun prisinformasjonen i tillegg til at disse, i likhet med ellers i KPI, beregnes med årlig kjeding.

Eksemplene ovenfor illustrerer godt hvor komplekse strekkodedataene er med svært mange ulike prisforløp og med ulike varer på tilbud til ulike tidspunkter. Strekkodedataene består av ekstremt mange varer og denne kompleksiteten gjør det

vansklig å skulle ”forstå” hvilken retning prisindeksene tar. Det er heller ikke slik at kjedeindeksene kun drifter nedover, vi ser av våre beregninger at de også kan drifte oppover selv om dette ser ut til å være mindre vanlig i norske data¹¹. Ivancic, Fox og Diewert (2009) viser til at ukentlig kjedede superlative prisindeks kan føre til svært kraftige drifting og at aggregering over måned og kvartal i stedet for uke antas å redusere denne skjevheten klart. Denne ”hamstringseffekten” hadde trolig vært mye tydeligere i de norske strekkodedataene med en ukentlig kjedet prisindeks. De Haan og van der Grient (2009) har analysert ukentlige nederlandske strekkodedata og viser til at konsumentene reagerer umiddelbart på sterke prisnedganger og at omsetningen faller dramatisk nesten til null i uker etter salg. I SSB innhentes pris- og omsetningsdata over midtuken i måneden mens i Nederland innhentes pris- og omsetningsdata fra hele måneden¹². De Haan og Opperdoes (1997) kom fram til at å bruke enhetspriser over en uke hver måned i stedet for å bruke enhetspriser over hele måneden gir opphav til langt større avvik enn selve formelvalget. Dette indikerer at aggregering over tid er svært viktig. Det arbeides nå, per juni 2010, for å utvide datainnsamlingsperioden i den norske matvareindeksen fra en til to uker. Hvilken betydning dette vil ha er foreløpig ikke analysert.

Aggregering over butikker kan også føre til drifting. I henhold til Ivancic, Fox og Diewert (2009) vil man kunne få større skjevhetsgrad i prisindeksen om man aggregerer over butikker enn om man ikke gjør det. I den norske matvareindeksen aggregerer vi verken over kjeder eller butikker. I henhold til Ivancic, Fox og Diewert (2009) er riktignok aggregering over tid av mye større betydning når det gjelder mulig drifting enn aggregering over butikker.

Så vidt vi vet er det ikke gjort mange analyser på drifting ved bruk av strekkodedata på aggregert nivå internasjonalt, mange internasjonale studier fram til nå har basert sine analyser på enkeltprodukt/produktgruppe noe som gjør det vanskelig å skulle konkludere eller generalisere om mulige effekter. Diskusjonen om hvordan man skal aggregere strekkodedata er særsviktig ettersom stadig flere land nå er interessert i å ta i bruk strekkodedata i sine prisindeks.

3.2. Sesongvarer

Enkelte varegrupper som blant annet sesongvarer må behandles med stor forsiktighet for å unngå skjevheter i kjedede prisindeks. Sesongvarer er produkter som enten ikke er tilgjengelig i visse perioder av året eller produkter som er tilgjengelig hele året, men med et pris- og omsetningsmønster som svinger stort i forhold til visse perioder av året. I den norske matvareindeksen har vi definert ulike påske- og juleprodukter som sesongvarer og tatt disse ut av beregningene. Andre mulige sesongvarer som fisk, frukt og grønnsaker har blitt behandlet likt som alle andre varer og er beregnet på samme måte dette til tross for at ILO manualen (2004) påpeker ”*When monthly prices are subject to regular and substantial seasonal fluctuations, monthly chaining cannot be recommended*”.

Vi har beregnet matvareindeksen på nytt der vi i tillegg til påske- og juleprodukter har definert flere sesongvarer særlig innen gruppene fisk, kjøtt og frukt. Dette gjelder spesielt varer som kun omsettes visse deler av året. På samme måte som i publisert prisindeks blir sesongvarene fjernet fra beregningene. Antallet sesongvarer fra måned til måned i prisindeksen med bredere definisjon av sesongvarer varierer med langt flere sesongvarer rundt november og desember enn ellers i året. I de-

¹¹ Feenstra og Shapiro (2003) som analyserte strekkodedata for tunfisk på boks fant ut at ukentlig kjedede Törnqvist prisindeks viste oppadgående drifting ”i perioder med lave priser men ingen annonsering, er ikke omsatt mengde høy[...]. Ettersom annonsering skjer i sluttfasen av salget, får prisoppganger som etterfølger salgsperioden mye høyere vekt enn prisnedganger ved salgsstart. Dette fører til en dramatisk drifting *oppover*”. Disse funnene synes ikke å være representative for norske strekkodedata.

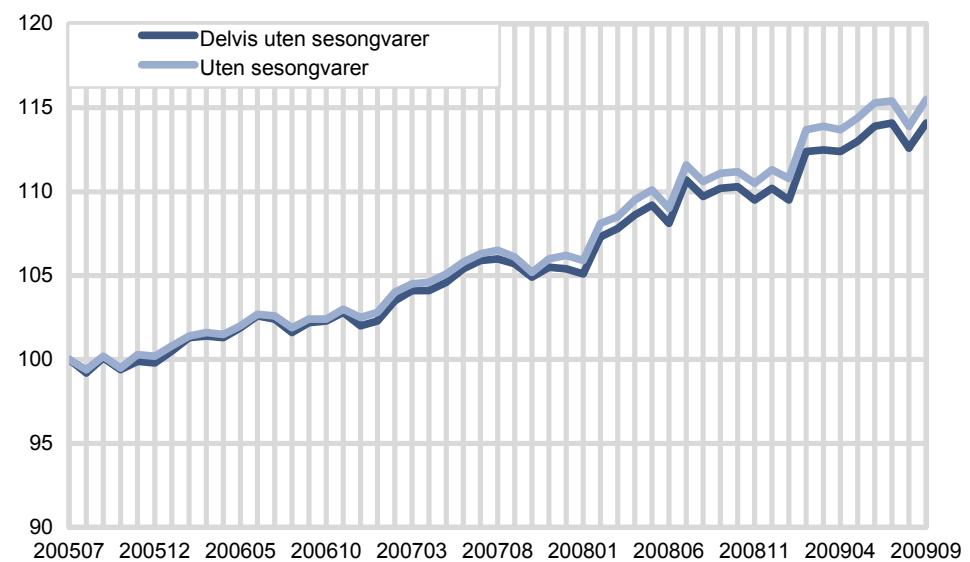
¹² Strekkodedata fra de tre første ukene inngår i KPI beregningene i Nederland.

september 2008 ble eksempelvis over 8 000 prisobservasjoner utelatt fra beregningene mot i underkant av 1 500 prisobservasjoner per måned ellers i året.

Våre beregninger viser at COICOP6-grupper med sterkt innslag av sesongvarer viser klare nedadgående skjevheter. Det typiske med sesongvarer som kun er tilgjengelig i visse perioder av året, er at med månedlig kjeding og uten imputering av manglende prisobservasjoner inkluderes disse først i beregningene i måned t+1. Med sesongvarer, og for så vidt med alt temporært frafall, mister vi prisendringer mellom siste måned de er i utvalget og første måned de kommer inn i utvalget på nytt. Sesongvarer kommer ofte inn på markedet med en høyere pris sammenlignet med forrige pris sist den var i sesong og denne effekten fanges ikke opp. I etterfølgende måneder, og når sesongvarene nærmer seg slutten av sin sesong, kan prisene falle markant (muligens kombinert med høye omsetningstall) for deretter å falle ut av utvalget. Dermed kan disse varene gi opphav til nedadgående skjevhets. Ettersom vi i matvareindeksen har definert få sesongvarer og dermed latt sesongvarene i stor grad inngå i beregningene er denne skjevheten synlig i resultatene, særlig på COICOP6-nivå. For å redusere denne skjevheten har vi fra og med 2009 utvidet koding av sesongvarer i publisert prisindeks og tar dermed disse ut av beregningene.

Det er opplagt noen klare svakheter ved å fjerne sesongvarer fullt og helt fra prisindeksen, særlig om sesongvarene utgjør en betydelig andel av konsumentenes totale forbruk. På sikt bør sesongvarer med størst omsetning inkluderes i prisindeksen. En mulig løsning er å inkludere transaksjonspriser i beregningene i sesongperioden og estimere priser utenfor sesongperioden. I dagens matvareindeks er det per i dag ingen løsning for imputering av priser, verken for sesongvarer eller andre varer. Fra og med januar 2010 implementerte CBS en løsning med imputering av temporært frafall der manglende prisobservasjoner blir imputert lik prisutviklingen for tilsvarende varer. Figurene 5, 6 og 7 viser at innslag av sesongvarer i prisindeksen har bidratt til nedadgående skjevheter;

Figur 5. Prisindeks for matvarer og alkoholfrie drikkevarer, Törnqvist prisindeks delvis¹³ uten sesongvarer og Törnqvist prisindeks uten¹⁴ sesongvarer. Juli 2005-september 2009. Indeks juli 2005=100



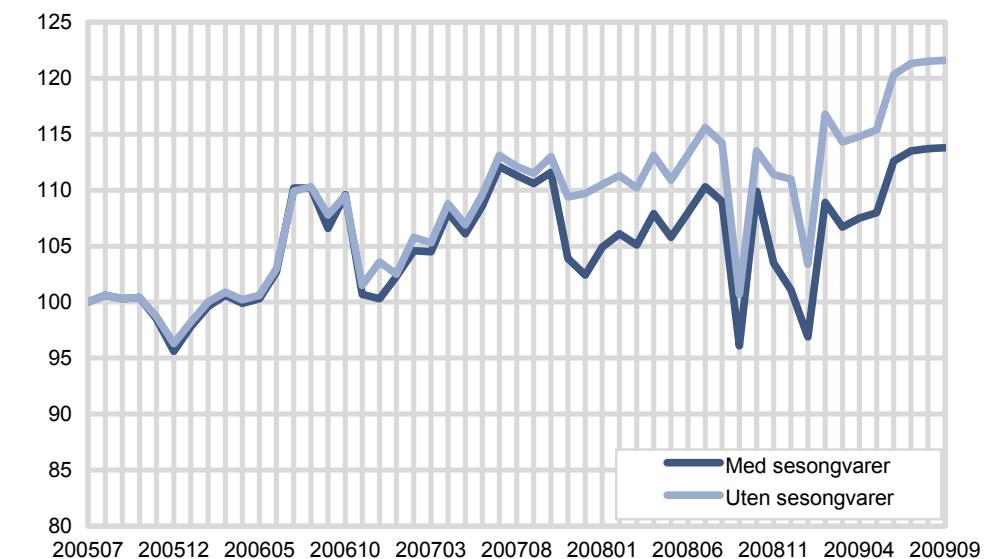
Törnqvist prisindeks delvis uten sesongvarer og Törnqvist prisindeks uten sesongvarer, viser at innslaget av sesongvarer har bidratt til å trekke ned prisveksten for delindeksen for matvarer og alkoholfrie drikkevarer med om lag 1,5 prosentpoeng hele perioden fra juli 2005 til september 2009 sett under ett, som betyr om lag 0,4

¹³ Törnqvist prisindeks delvis uten sesongvarer dvs. uten jule- og påskeprodukter.

¹⁴ Törnqvist prisindeks uten sesongvarer dvs. uten en bredere definisjon av sesongvarer.

prosentpoeng i gjennomsnitt per år. For KPI betyr ”feilen” om lag 0,04 prosent i de årlige vekstratene. Derimot er avvikene til dels betydelige innen enkelte COICOP6-grupper;

Figur 6. Prisindeks for saltet, tørket og røkt fisk, Törnqvist prisindeks med og uten sesongvarer. Juli 2005-september 2009. Indeks juli 2005=100

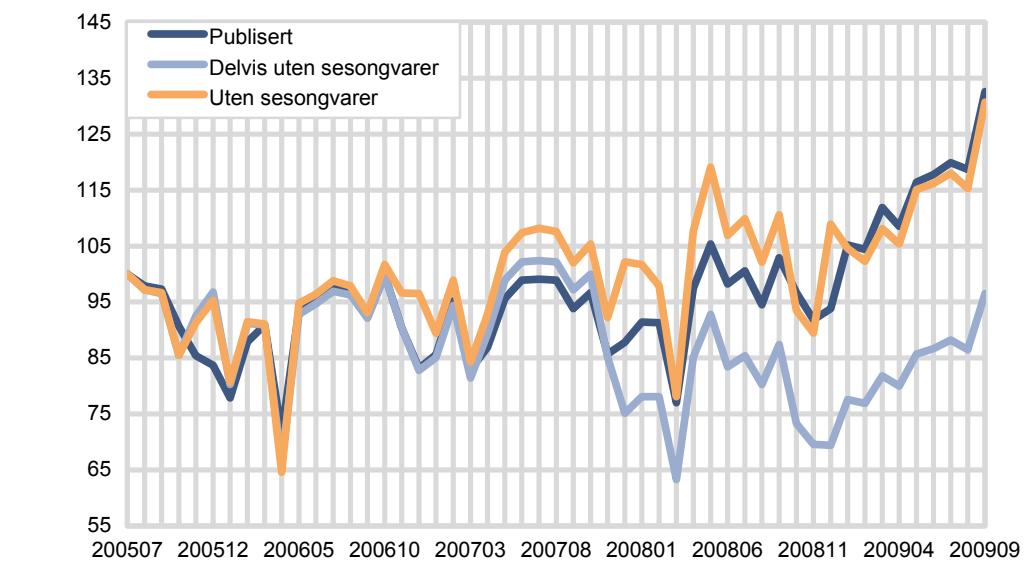


Sesongvarer trekker ned prisveksten for COICOP6-gruppen ”Saltet, tørket og røkt fisk” med om lag 2 prosentpoeng i gjennomsnitt per år. Fom november 2007 er avviket mellom prisindeksene betydelig. Lutefisk er tidligere ikke kodet som sesongvare og avviket mellom seriene viser at prisforløpet på lutefisk gir store effekter i denne prisindeksen. Lutefisk står for en stor del av prisnedgangen i gruppen fra november og desember 2007 for deretter å helt falle ut av prisindeksen. Dermed henter ikke prisindeksen seg inn igjen og prisindeksen viser en nedadgående skjevhets-

Også COICOP6-gruppen ”Situsfrukter” har et klart innslag av sesongvarer i gruppen, se figur 7. Om vi ikke hadde revidert¹⁵ i denne gruppen ville sesongvarer gitt opphav til omfattende skjevhet i denne gruppa. Det er store avvik mellom seriene med og uten sesongvarer. Sammenligner vi Törnqvist prisindeks uten sesongvarer med *publisert prisindeks* er det riktig nok mindre forskjeller. I den publiserte prisindeksen har vi fjernet en del av disse skjehetene manuelt. Dette forteller oss at vi til dels klarer å fange opp mulige skjevheter forårsaket av sesongvarene. Med revisjon kan vi dempe utslag fra observasjoner som avviker mye fra hovedtrenden i de enkelte gruppene. Om vi overlater sesongvarebehandlingen til manuell revisjon kan dette riktig nok føre til en del tilfeldige utslag som våre analyser viser.

¹⁵ Revisjon i matvareindeksen betyr det samme som å ta observasjoner ut av beregningene. Revisor tar avgjørelse på hvilke prisobservasjoner som skal fjernes.

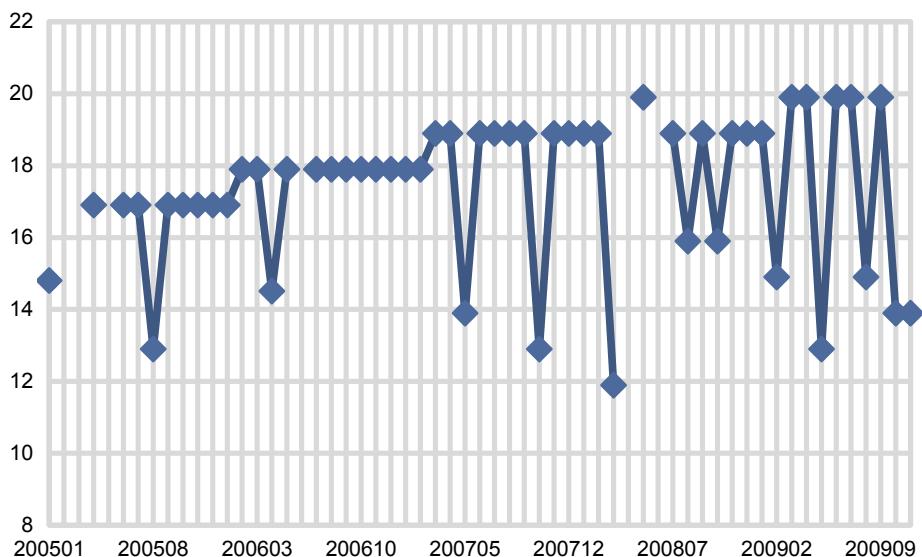
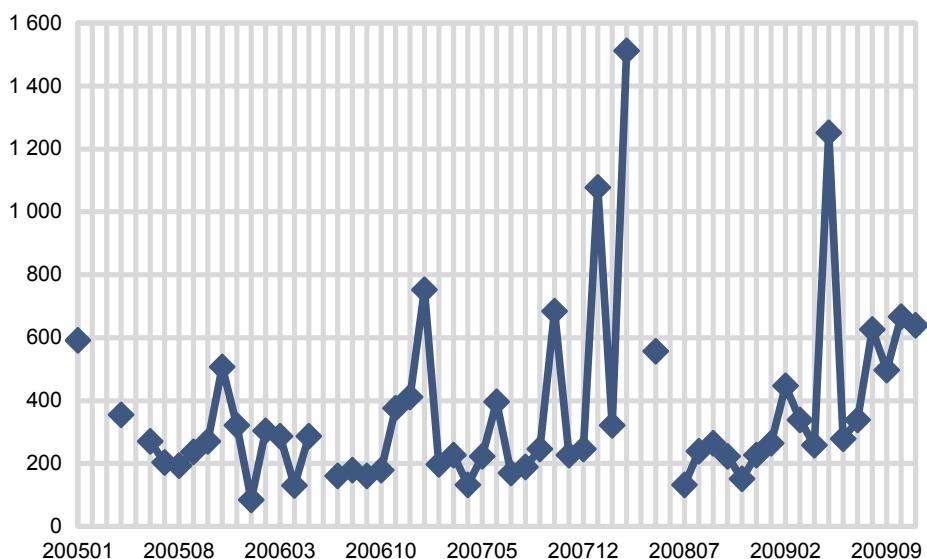
Figur 7. Prisindeks for sitrusfrukter. Publisert prisindeks, Törnqvist prisindeks delvis uten sesongvarer og Törnqvist prisindeks uten sesongvarer. Juli 2005-september 2009. Indeks juli 2005=100



Så langt har vi kun vurdert sesongvarer som er i sesong i visse deler av året. Mange varer har et pris- og omsetningsmønster som gjør at også de kan bli klassifisert som sesongvarer selv om varene omsettes året rundt. Sjokoladen Kvikk Lunch er en slik vare. Kvikk Lunch er en typisk påskesjokolade som ofte blir satt på tilbud rundt påsketider med tilhørende sterk økning i omsetningen. Når påskan er over går normalt konsumet av sjokoladen sterkt ned, trolig til et nivå som er klart lavere enn før salget eller kanskje omsetningen faller helt ned i null. Dette vil også kunne føre til asymmetriske omsetningsandeler og til at prisindeksen drifter nedover.

3.3. Manglende prisobservasjoner

Vi har ovenfor sett at en vare kan falle ut av utvalget på grunn av at den ikke lengre er i sesong. En vare kan også falle ut av beregningene av andre årsaker enten på grunn av temporært eller permanent frafall. Temporært frafall kan skyldes, foruten sesongvarer utenfor sesong, varer som er midlertidig utsolgte eller varer som det ikke selges noe av. Ettersom det er transaksjonspriser vi mottar fra de store dagligvarekjedene er det kun pris og omsetning på omsatte varer i midtuken i måneden vi mottar. Etter en tilbudsperiode med høye omsetningstall kan det være slik at omsetningen faller til null når prisen er tilbake til sitt ”normal” nivå. I disse tilfeller skjer ingen transaksjon og prisoppgangen vil ikke bli registrert i prisindeksen. Dermed faller varen ut av prisindeksen på et lavt prisnivå som igjen kan føre til nedadgående skjevhets. Figurene 8 og 9 viser eksempler på temporært frafall for filtermalt kaffe i en enkeltavdeling. Fra februar til mars 2008 faller prisen fra 18,90 til 11,90 kroner med høy omsetning før deretter å falle ut av prisindeksen. I mai 2008 kommer den inn igjen for så å falle ut på nytt. Prisendringen på varen registreres derfor ikke før august 2008.

Figur 8. Kaffe filtermalt 250g. pris i kroner. Januar 2005-november 2009**Figur 9.** Kaffe filtermalt 250g. Omsetning i kroner. Januar 2005-november 2009

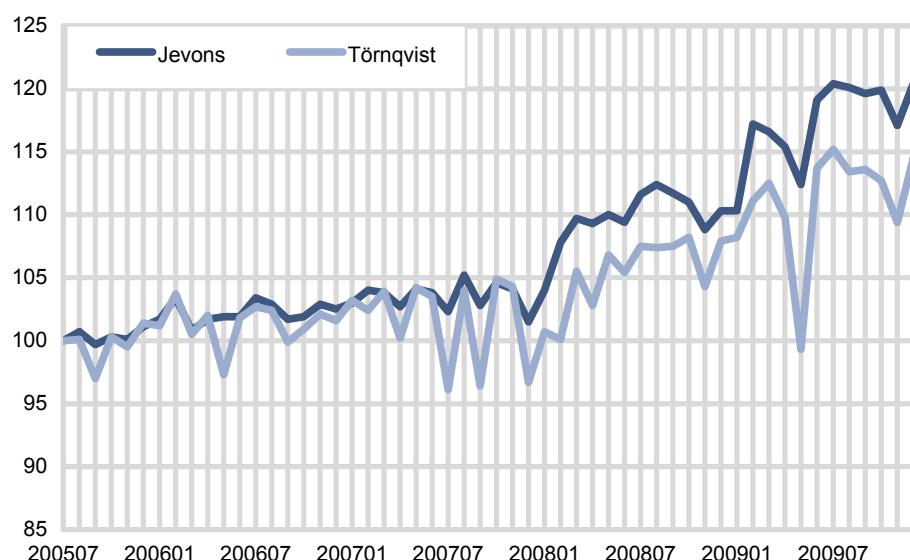
Permanent frafall kan skyldes varer som går helt ut av produksjon og som ikke kommer inn igjen i utvalget. For at butikker skal få tømt sine lagre, ser vi ofte at varer settes på tilbud kombinert med muligens høye omsetningstall før de går ut av produksjon. Dette bidrar til det samme forløpet med høy vekt knyttet til prisnedgangen, prisoppgangen uteblir og konsekvensen er nedadgående skjevhets-

I KPI løser vi vanligvis temporært og permanent frafall med å imputere pris for manglende prisobservasjoner. Manglende prisobservasjoner på grunn av temporært frafall, får imputert en prisutvikling lik tilsvarende produkter. Permanent utgårte produkter som ikke blir erstattet av nye produkter imputeres med en gjennomsnittspris av tilsvarende produkter ettersom konsumenten i framtiden ikke lengre står ovenfor dette produktet. I KPI ellers vet vi som oftest også noe om årsaken til en manglende prisobservasjon, enten fordi prisene innhentes gjennom skjema hvor butikken oppgir frafallsgrunn, eller fordi prisinnsamlingen skjer sentralt. Med strekkodedata vet vi ikke noe om årsaken til frafall av prisobservasjoner. I dagens matvareindeks er det som tidligere nevnt ingen løsning for imputering av priser, verken for temporært eller permanent frafall. Vi ser av våre beregninger at dette fører til en underestimering av prisveksten og nedadgående skjevheter særlig innen enkelte konsumgrupper.

Når en vare er permanent utgått i KPI skal våre oppgavegivere dvs. butikkene ersette den utgåtte varen med en vare av tilsvarende kvalitet. I matvareindeksen og med bruk av strekkodedata gjøres ingen slike erstatninger, varen går ganske enkelt ut av utvalget. Kvalitetsendringer får en annen dimensjon når vi snakker om strekkodedata. En stor fordel med strekkodedata og identifikasjon gjennom EAN-koder er at vi er sikre på at vi sammenligner identiske varer. Samtidig er det et problem at samme type varer kan få forskjellige EAN-koder dersom butikkene enten skifter produsent eller at produsenter skifter artikkelnummer, da går varen (EAN-koden) helt enkelt ut av utvalget. Man kan si at i enkelte tilfeller blir EAN-nivået for detaljert, varer som er identiske ut fra konsumentenes perspektiv kan ha ulike EAN-koder. Å koble på EAN-koder kan dermed underestimere antallet varer som kan kobles mot hverandre og overestimere varenes utskiftningshastighet.

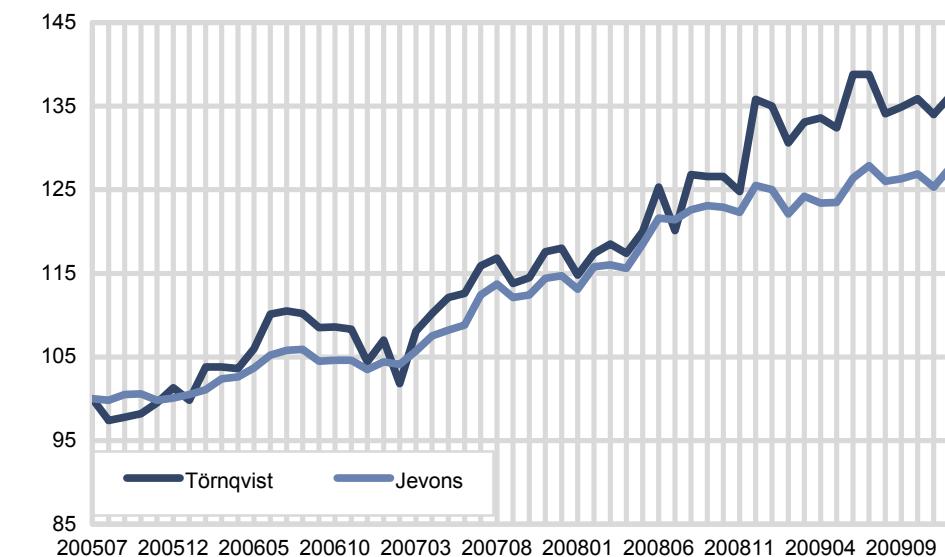
I figur 10 er prisutviklingen på pizza illustrert. Eksemplet viser hvilke effekter vi kan få av prisobservasjoner som faller ut av datagrunnlaget. I desember 2007 faller prisindeksene klart, i januar 2008 henter de aller fleste prisene seg inn igjen (noe vi ser av Jevons prisindeks som vil bli nærmere forklart i kapittel 4.1), men med unntak av noen priser (med høye omsetningstall i desember) som faller ut og vi får en situasjon med nedadgående skjevhet i Törnqvist prisindeks. De manglende prisobservasjonene påvirker også Jevons prisindeks, men uten eksplisitt vektning er betydningen mindre.

**Figur 10. Prisindeks for pizza med Törnqvist og Jevons formel. Juli 2005-desember 2009.
Indeks juli 2005=100**



Vi ser også av våre beregninger at prisindeksene kan vise skjevhet *oppover* som en følge av manglende prisobservasjoner. I visse tilfeller er det slik at vi ikke får registrert prisnedgangen sannsynligvis som en følge av at varen er midlertidig utsolgt, men i stedet registreres kun prisoppgangen. Dette finner vi eksempler på i COICOP6-gruppene "Storfe", "Svin" og "Laks/ørret". I figur 11 er prisutviklingen på storfe illustrert. I desember 2008 stiger den superlativ prisindeksen for storfe med nesten 10 prosent. En prisobservasjon på indrefilet styrer mye av denne prisoppgangen etter tilbud. Tilbudsperioden ble imidlertid ikke registrert da prisobservasjonen ikke var i utvalget i oktober. Men alt i alt tyder våre analyser likevel på at det er langt flere tilfeller av at prisindeksen viser skjevheter nedover enn oppover.

**Figur 11. Prisindeks for storfe med Törnqvist og Jevons formel. Juli 2005-desember 2009.
Indeks juli 2005=100**



4. Alternative prisindekser

4.1. Jevons prisindeks

For å undersøke for mulig drifting i vår prisindeks for matvarer og alkoholfrie drikkevarer har vi beregnet en månedlig kjedet prisindeks ved hjelp av (uveid) Jevons formel på mikronivå i perioden juli 2005 til og med desember 2009. Vi benytter Jevons prisindeks som et sammenligningsgrunnlag med Törnqvist prisindeks ettersom Jevons prisindeks viser hvorvidt prisen er tilbake til utgangspunktet eller ikke. I de tilfeller hvor Jevons prisindeks er tilbake til utgangspunktet etter tilbud mens den superlativ prisindeksen ikke er det, kan dette være en indikasjon på mulig drifting. I Jevons prisindeks vil også eventuelle skjevheter forårsaket av manglende prisobservasjoner og sesongvarer forekomme, men betydningen vil sannsynligvis være mindre. Jevons prisindeks kan defineres som uveid geometrisk gjennomsnitt av prisrelativer som er identisk med relativet av uveide geometriske gjennomsnittspriser;

$$(10) \quad P_J^{0t} = \prod \left(\frac{p_i^t}{p_i^0} \right)^{1/n} = \frac{\prod (p_i^t)^{1/n}}{\prod (p_i^0)^{1/n}}$$

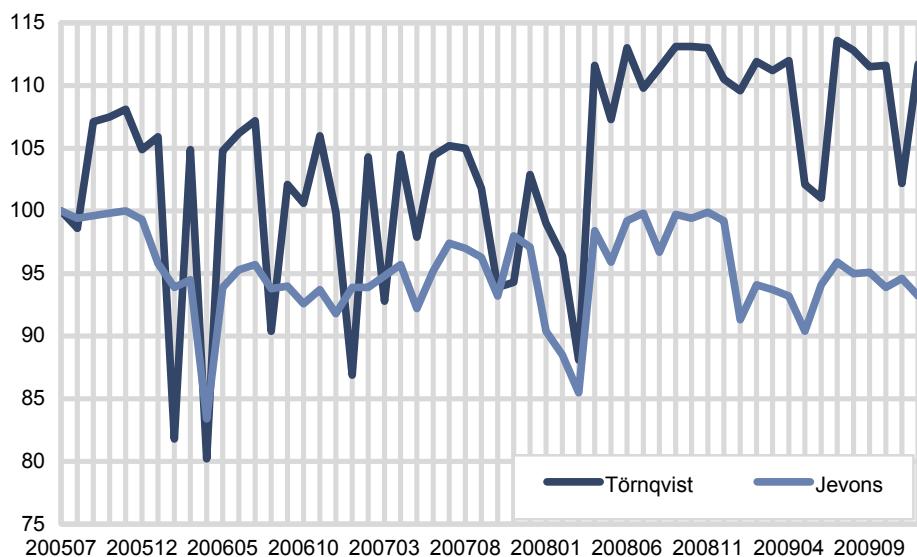
Fra COICOP6-nivå og videre opp til totalindeks for matvarer og alkoholfrie drikkevarer er i likhet med publisert indeks, Laspeyres formel med årlig kjeding benyttet.

I januar 2010 introduserte CBS en ny metode for å beregne prisindekser i KPI basert på strekkodedata. På grunn av risikoen for drifting ved bruk av månedlig kjeding og superlativ indeks har CBS valgt å beregne en månedlig kjedet Jevons prisindeks. Ettersom Jevons prisindeks er en uveid prisindeks hvor man kun tar hensyn til endringer i priser, ekskluderer CBS varer med lav omsetning. Vi har valgt å beregne Jevons prisindeks etter samme prinsipp. Ved hjelp av ”cut-off” metoden blir varer med høy omsetning inkludert i utvalget, mens varer med omsetning under en viss terskel vil bli ekskludert fra utvalget. Metoden kan forklares på følgende måte;

En vare i vil bli inkludert i utvalget for måned $t-1$ og måned t om den gjennomsnittlige omsetningsandelen i begge perioder, ($s_i^{t-1} + s_i^t$)/2 er over en viss terskel. ”Cut-off” terskelen er satt til $1/(N^{t-1,t} \chi^{t-1,t})$ der $N^{t-1,t}$ står for antallet observasjoner på COICOP6-nivå. Vi har valgt $\chi^{t-1,t}=2$. Dette betyr at om for eksempel $N^{t-1,t}=100$ vil alle varer med en gjennomsnittlig omsetningsandel på mer enn 0,5 % bli inkludert i beregningsgrunnlaget. Dermed er det ingen fast utvalgstørrelse, men antallet varer som ekskluderes fra utvalget kan variere over tid og fra gruppe til gruppe.

Våre beregninger viser at om lag 12 prosent av omsetningen totalt sett blir ekskludert mens i underkant av 50 prosent av prisobservasjonene fjernes. Andelen varer som fjernes varierer riktignok mye fra COICOP6-gruppe til COICOP6-gruppe, se vedlegg B. Eksempelvis fjernes om lag 11 prosent fra gruppen ”Sideflesk og spek” mens i gruppen ”Resten av sjokolade” fjernes så mye som i overkant av 80 prosent. Vi har imidlertid valgt å legge til en forutsetning om at dersom antall prisobservasjoner innen en COICOP6-gruppe er mindre enn 100, vil det ikke bli ekskludert noen prisobservasjoner. Vi ser at ved å sette verdien på $\chi^{t-1,t}=2$ faller til dels en veldig stor del av datagrunnlaget bort som eksempelvis i gruppen ”Resten av sjokolade”, se figur 12. Jevons og Törnqvist prisindeks viser svært avvikende forløp for denne COICOP6-gruppen. Det er en lite homogen gruppe og har dessuten et sterkt innslag av sesongvarer som også fjernes fra datagrunnlaget. I sin testfase valgte CBS å benytte $\chi^{t-1,t}=2$, men i offisiell indeks fra januar 2010 har de valgt $\chi^{t-1,t}=1,25$ noe som innebærer en reduksjon av ”cut-off” utvalget.

Figur 12. Prisindeks for Resten av sjokolade med Törnqvist og Jevons formel. Juli 2005-desember 2009. Indeks juli 2005=100



4.2. RYGEKS prisindeks

Ivancic, Fox og Diewert (2009) har nylig lansert en ny metode for å konstruere superlativ prisindekser basert på strekkodedata ved å bruke multilaterale indeksteknologi. Denne metoden er en versjon av multilaterale GEKS indekser (Gini, 1931; Eltentö og Köves, 1964; Szulc, 1964) som ofte blir brukt for prissammenligninger mellom land (kjøpekraftspariteter), men tilpasset for sammenligninger over tid der hver tidsperiode (i stedet for land) blir brukt som en enhet. Disse GEKS indeksene er transitive slik at de kjedede indeksene er identiske med de direkte indeksene og dermed driftfrie. Ettersom denne metoden er driftfri er den et godt verktøy for å avdekke mulig drifting. GEKS metoden utnytter all mulige ”match” som er i data-

ene og beregner direkte ”match-item” prisindeksene ved å bruke hver tidsperiode som en basisperiode, det betyr at bare prisrelativer av varer som er kjøpt i begge de to periodene som sammenlignes blir med i beregningene. Imputering for å hankses med manglende prisobservasjoner er derfor ikke nødvendig (jmf. kapittel 3.3).

GEKS prisindeks er et geometrisk gjennomsnitt av alle de direkte bilaterale prisindeksene. La P^{jl} og P^{kl} være bilaterale prisindeksene (som eksempelvis Törnqvist eller Fisher prisindeks) mellom enhetene j og l ($l=1\dots m$) og tilsvarende mellom k og l . GEKS prisindeks mellom j og k kan da skrives som;

$$(11) \quad P_{GEKS}^{jk} = \prod_{l=1}^M \left[P^{jl} / P^{kl} \right]^{1/M} = \prod_{l=1}^M \left[P^{jl} \times P^{lk} \right]^{1/M}$$

Ettersom GEKS prisindeks er transitiv av oppbygning gjelder også det andre uttrykket i (11).

Når nye data blir tilgjengelig krever metoden at prisindeksene revideres, en løsning som i KPI sammenheng vil være uakseptabel. For å unngå denne problemstillingen har Ivancic, Fox og Diewert (2009) lansert en ”Rolling year” GEKS prisindeks (RYGEKS)¹⁶. ”Rolling year” metoden bruker pris- og omsetningsinformasjon fra de siste 13 månedene for å beregne en RYGEKS prisindeks der hver måned i dette ”13-månedersvinduet” blir brukt som basisperiode. La oss anta at vi har periodene 1, ..., 13. Når en ny periode med data er tilgjengelig flytter ”vinduet” seg en periode framover og vil da bestå av data for perioden 2, ..., 13+1, det betyr at for hver nye tidsperiode som er tilgjengelig, fjernes første perioden (måned) fra ”vinduet” og den siste perioden (måned) legges til ”vinduet”.

Ettersom RYGEKS metoden er driftfri, anbefales det å beregne en superlativ prisindeks basert på både pris- og omsetningsinformasjon. Om vi velger å beregne prisendring fra periode (måned) 13 til 14 basert på RYGEKS prisindeks etter Törnqvist prisindeksformel betyr dette følgende; først beregner vi prisendringen fra periode 2 til periode 14 (kun basert på prisobservasjoner som fins i begge disse periodene) som ses opp mot med tilsvarende prisindeks for periode 14-1; Törnqvist prisindeks_{14/2}/Törnqvist prisindeks_{13/2}. Deretter beregner vi Törnqvist prisindeks_{14/3}/Törnqvist prisindeks_{13/3}, Törnqvist prisindeks_{14/4}/Törnqvist prisindeks_{13/4},Törnqvist prisindeks_{14/14}/Törnqvist prisindeks_{13/14}. For å få RYGEKS prisindeks tas det altså et geometrisk gjennomsnitt av alle disse prisrelativene.

Fordelen med RYGEKS prisindeksen er at den tillater sammenligninger av sesongvarer som kun er tilgjengelig i visse deler av året. Ettersom RYGEKS metoden beregner direkte indeks over en 13-måneders periode sammenlignes sesongvarer på vei ut av sesong med sesongvarer på vei inn i sesong.

For mer informasjon om RYGEKS prisindeksene og det teoretiske grunnlaget, se Ivancic, Fox og Diewert (2009) og de Haan og van der Grient (2009).

Ivancic, Fox og Diewert (2009) anbefaler statistikkbyråer å ta i bruk RYGEKS metoden for beregning av prisindeksene basert på strekkodedata. Enn så lenge er denne nyutviklede metoden ikke tatt i bruk av noen statistikkbyråer eller blitt implementert i noen offisielle KPI-tall. Metoden befinner seg fortsatt på et tidlig stadium og må inntil videre anses som en eksperimentell prisindeks. Resultater fra metoden er uten tvil svært lovende og metoden testes nå ut på ulike lands strekkodedata. De Haan og van der Grient (2009) har testet ut RYGEKS prisindeksene for en rekke varegrupper over en periode på flere år på nederlandske strekkodedata. Til tross for svært lovende empiriske resultater valgte CBS ikke å direkte innføre en RYGEKS prisindeks ved omleggingen i januar 2010, men derimot en modifisert

¹⁶ RYGEKS prisindeks ble i utgangspunktet utviklet spesielt for CBS som et alternativ til uveid gjennomsnitt (Jevons prisindeks) på elementært nivå.

Jevons prisindeks. CBS anser likevel RYGEKS metoden som en ”benchmark” prisindeks og har brukt sammenligninger med RYGEKS metoden i sin utforming av Jevons prisindeks. SSB har testet ut RYGEKS prisindekser for matvarer og alkoholfrie drikkevarer og gjort sammenligninger mellom en reberegning av offisiell prisindeksmetode og RYGEKS metoden, se kapittel 5 nedenfor.

5. Sammenligninger mellom Törnqvist, Jevons and RYGEKS

5.1. Sammenligninger på totalnivå

Nedenfor har vi sammenlignet prisindeks for matvarer og alkoholfrie drikkevarer beregnet med Jevons og RYGEKS prisindeksformel med vår offisielle beregningsmetode - en Törnqvist prisindeks. Törnqvist, Jevons og RYGEKS prisindeks avviker fra hverandre ved at RYGEKS prisindeks har alle sesongvarer *inkludert* i datagrunnlaget, mens i Törnqvist og Jevons prisindeks har vi, som i publisert prisindeks, fjernet de typiske sesongvarene. Vi har også beregnet RYGEKS prisindeks uten de viktigste sesongvarene for å gjøre indeksene mer sammenlignbare og teste betydningen av sesongvarene. Alle indeksene er revidert likt, det vil si ingen manuell revisjon¹⁷ er utført og kun åpenbare feil (F-verdier)¹⁸ er automatisk fjernet. Alle kritiske verdier er dermed inkludert noe som medfører at Törnqvist prisindeks vil avvike fra publisert prisindeks. I tillegg har vi ikke for RYGEKS prisindeksene kontrollert alle prisrelativene i 13-månedersperioden, noe som kan gi visse utslag for RYGEKS prisutviklingen. Det er viktig å understreke at fjerning av kritiske verdier i mange tilfeller er vårt verktøy for å redusere eventuell drifting. Sammenligninger med *publisert* prisindeks viser at fjerning av kritiske verdier har bidratt til redusert drifting for flere undergrupper. For andre undergrupper har revisjonen bidratt til mer drifting noe som viser at manuell revisjon kan føre til noen tilfeldige utslag.

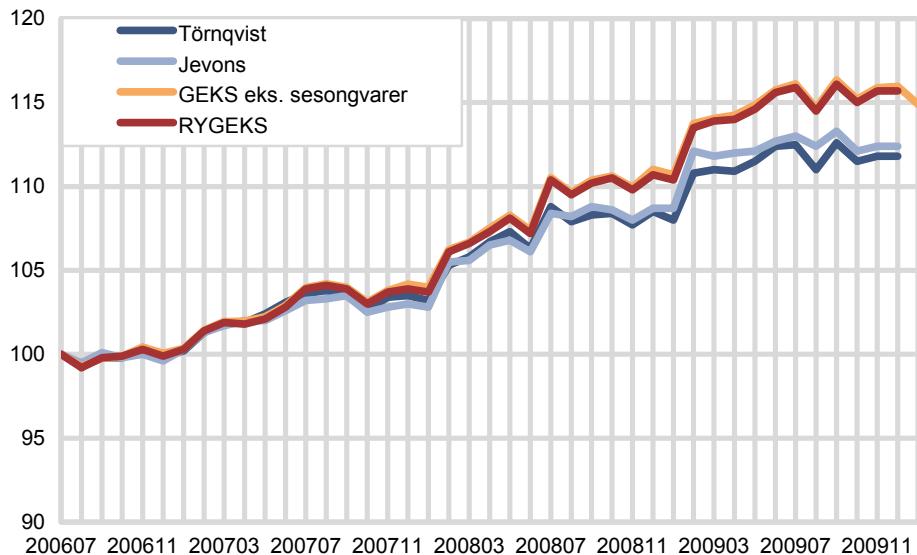
RYGEKS prisindeks er beregnet for perioden juli 2006 til og med desember 2009 (basert på data for perioden juli 2005-desember 2009), mens Törnqvist og Jevons prisindeks er beregnet for perioden juli 2005-desember 2009. Sammenligningene gjøres for perioden juli 2006 til desember 2009.

Figur 13 viser et klart avvik mellom Törnqvist og RYGEKS prisindeks. Differansen mellom de to prisindeksene for matvarer og alkoholfrie drikkevarer totalt utgjør 3,9 prosentpoeng for hele perioden juli 2006 til desember 2009, eller litt i overkant av 1 prosentpoeng i gjennomsnitt i året. Avviket stiger utover perioden med størst avvik i juli 2007 og juli 2008 der prisutviklingen på frukt og ulik behandling av sesongvarer er hovedårsaken. I februar 2010 er det økende avviket trolig forårsaket av en økende andel varer som går ut av utvalget på salg. Det er kun mindre avvik mellom Törnqvist og Jevons prisindeks. Fram til august 2008 er det ingen systematiske forskjeller i utviklingen mellom de to indeksseriene. Fra august 2008 blir Törnqvist prisindeks liggende systematisk under Jevons prisindeks. Avviket som oppstår i august 2008 skyldes blant annet bidraget fra COI-COP4-gruppene ”Fisk” og ”Kaffe, te og kakao”. Totalt sett i perioden juli 2006 til desember 2009 er differansen mellom de to prisindeksene for matvarer og alkoholfrie drikkevarer i underkant av 1 prosentpoeng. RYGEKS prisindeks ligger systematisk over begge de andre alternativene fra og med 2. halvår 2007. Anses RYGEKS prisindeks som en ideell indeks eller ”benchmark” prisindeks er det klar indikasjon på skjevheter på aggregert nivå i den offisielle prisindeksmetoden i perioden juli 2006 til og med desember 2009.

¹⁷ Dvs. ingen fjerning av bidrag som betyr mest for resultatet på elementært nivå.

¹⁸ Måned-til-måned prisrelativer ≥ 3 eller $\leq 0,33$ fjernes.

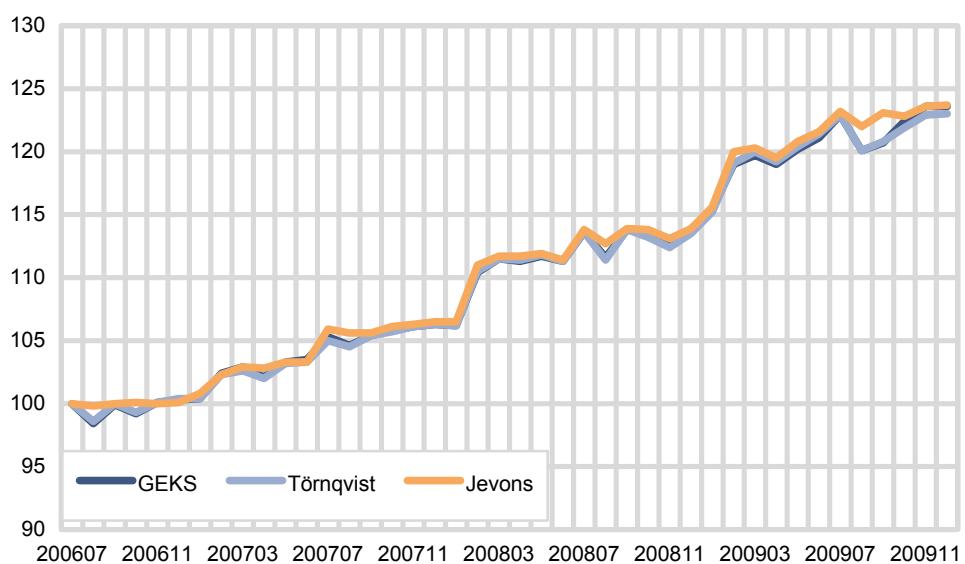
Figur 13. Prisindeks for matvarer og alkoholfrie drikkevarer med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100



5.2. Sammenligninger på detaljert nivå

Sammenligninger på mer detaljert nivå viser at forholdet mellom de ulike prisindeksene er noe ulikt mellom konsumgruppene. I de tilfeller med større avvik mellom prisindeksene ligger RYGEKS prisindeks i de fleste tilfellene systematisk over både Törnqvist og Jevons prisindeks, der Jevons prisindeks synes å ligge oftere over enn under Törnqvist prisindeks. For mange konsumgrupper er det kun mindre avvik og ingen systematiske forskjeller i prisutviklingen mellom de ulike prisindeksene. For gruppene ”Melk, ost og egg” og ”Oljer og fett” er det svært små avvik mellom de ulike prisindeksene i mesteparten av analyseperioden. Se figur 14 for COICOP4-gruppen ”Melk, ost og egg”.

Figur 14. Prisindeks for melk, ost og egg med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100

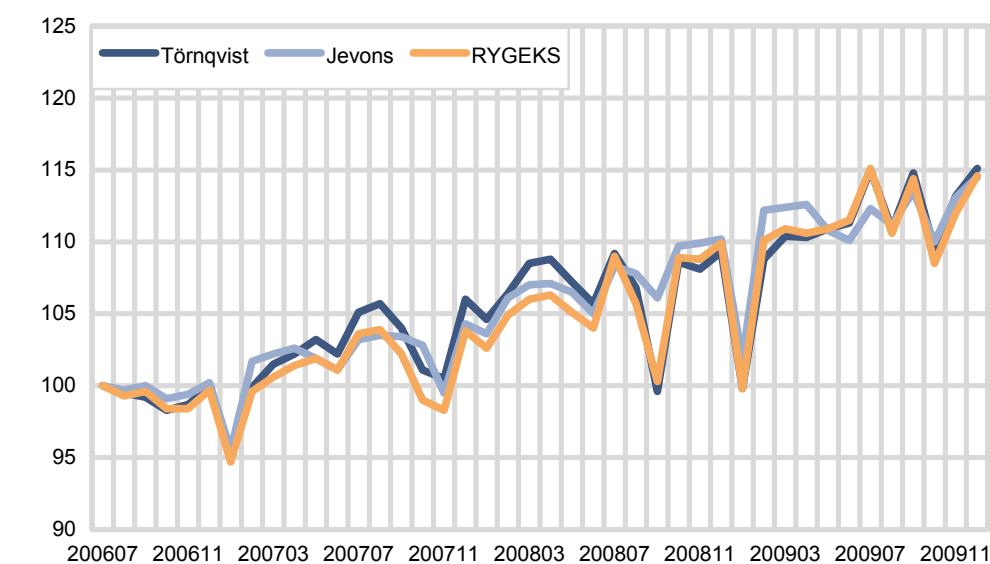


Analyser av prisindeksene på COICOP6-nivå viser i likhet med analyser på COICOP4-nivå at det er store forskjeller fra konsumgruppe til konsumgruppe. Matvareindeksen består av 138 COICOP6- grupper, se vedlegg A og B. For flere av seriene er avvikene små med tilnærmet lik pristrend gjennom hele perioden. De fleste av disse seriene består av homogene basisvarer som melk, fløte, margarin, enkelte

hermetikkvarer, supper og ris og annet gryn som i mindre grad knyttes opp mot tilbudskampanjer.

I prisindeksene for fisk ser det heller ikke ut til å være noen betydelig systematikk. Indeksene for fisk er nokså volatile, men følger hverandre likevel relativt godt over tid. Utviklingen i Törnqvist, Jevons og RYGEKS prisindeks for fisk og fiskevarer vises i figur 15.

Figur 15. Prisindeks for fisk med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100



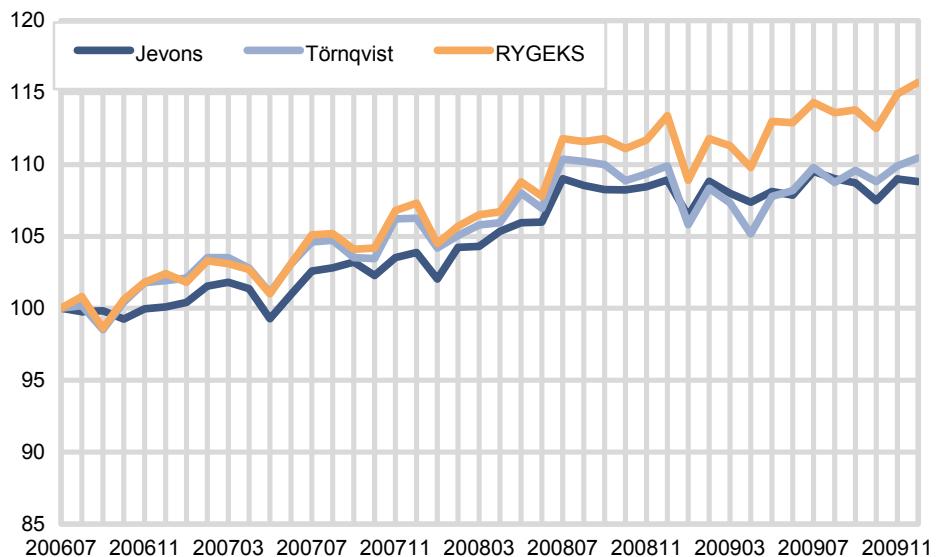
Indeksen for fisk er blant de mest volatile gruppene sammen med gruppene ”Frukt”, ”Grønnsaker” samt ”Kaffe, te og kakao”, se tabell 3.

Tabell 3. Gjennomsnittlig månedsendring og standardavvik for Törnqvist, Jevons og RYGEKS prisindeks. Juli 2006-desember 2009

	Gj.snittlig mnd.endring			Standardavvik		
	Törnqvist	Jevons	RYGEKS	Törnqvist	Jevons	RYGEKS
Brød og kornprodukter	0,39	0,42	0,47	1,07	0,97	1,12
Kjøtt	0,25	0,21	0,37	1,45	1,10	1,57
Fisk	0,40	0,37	0,40	3,50	2,84	3,59
Melk, ost og egg	0,51	0,53	0,53	1,23	1,09	1,23
Oljer og fett	0,51	0,51	0,53	1,56	1,60	1,55
Frukt	0,28	0,31	0,28	3,56	3,43	3,02
Grønnsaker	-0,05	-0,05	0,09	2,57	2,28	2,65
Sukker, syltetøy, sjokolade og andre sukker-varer	0,03	0,13	0,28	1,34	1,13	1,23
Andre matvarer	0,33	0,35	0,38	1,05	0,94	1,07
Kaffe, te og kakao	0,24	0,30	0,32	3,09	2,13	3,13
Mineralvann, leskedrikker og juice	0,37	0,36	0,43	1,93	1,44	2,00

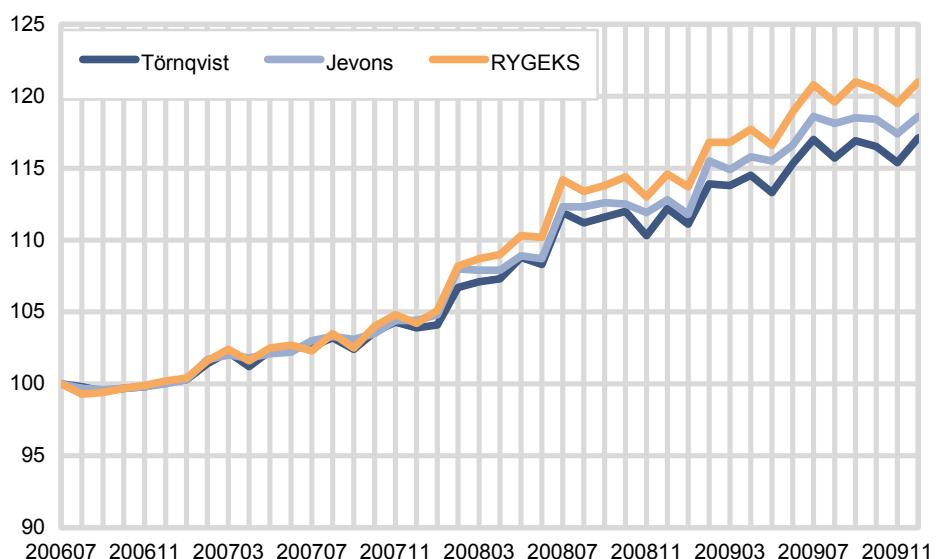
Törnqvist, Jevons og RYGEKS prisindeks for kjøtt følger også hverandre relativt godt fram til midten av 2008 (selv om Jevons prisindeks ligger noe under begge de to alternativene) da både Törnqvist og Jevons prisindeks blir liggende systematisk under RYGEKS-prisindeksen. Utover 2009 øker avviket særlig mellom Törnqvist/Jevons prisindeks og RYGEKS prisindeks. Dette skyldes blant annet prisutviklingen på ”Fjørfe”, ”Sideflesk og spek” og ”Svin”. Eksempelvis faller både RYGEKS og Törnqvist prisindeks for sideflesk og spek kraftig i desember 2008. RYGEKS prisindeks henter seg igjen, mens Törnqvist prisindeks ikke gjør det grunnet tilbudspriser som forsvinner fra datamaterialet måneden etter. Det samme gjelder indeksen for fjørfe der Törnqvist og RYGEKS prisindeks viser et kraftigere fall i november 2008 sammenlignet med Jevons prisindeks og der RYGEKS prisindeks henter seg inn måneden etter, men ikke Törnqvist prisindeks.

Figur 16. Prisindeks for kjøtt med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100



Fra og med starten av 2008 ligger Törnqvist prisindeks systematisk under både Jevons og RYGEKS prisindeks for konsumgruppen "Brød og kornprodukter". Totalt sett i analyseperioden er avviket mellom Törnqvist og RYGEKS prisindeks på i undertakn av 4 prosentpoeng. Brytes "Brød og kornprodukter" ned på 5-sifret COICOP-nivå er det mindre avvik mellom seriene for "Pasta" og "Brød" mens avvikene er større for gruppene "Mel, frokostblanding og gryn", "Kaker" og "Andre bakevarer". For gruppen "Andre bakevarer" er det særlig gruppen "Pizza" som viser store avvik mellom Törnqvist og RYGEKS prisindeks, jmf. figur 10.

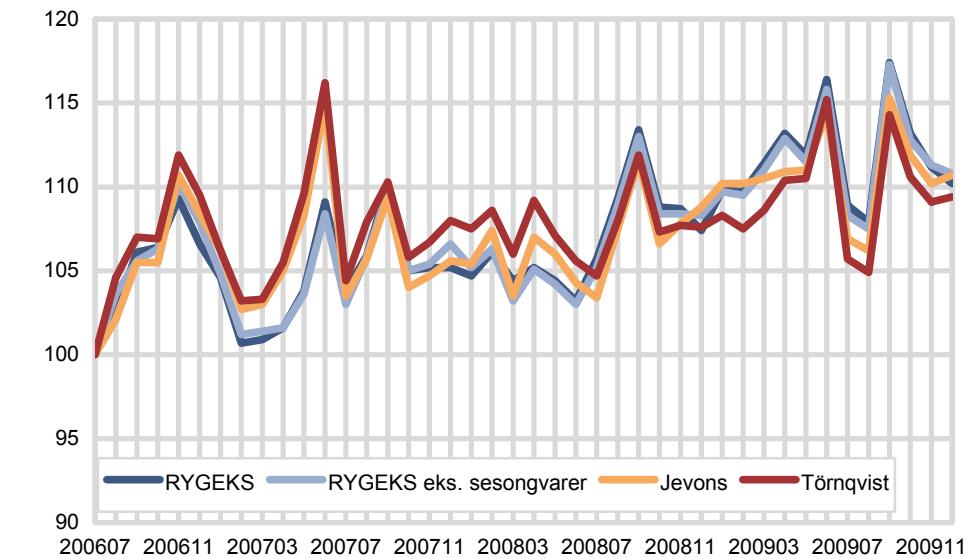
Figur 17. Prisindeks for brød og kornprodukter med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100



I prisindeksen for frukt er det blant annet gruppene "Steinfrukt", "Melon" og "Nøtter, frø og kjerner" som viser store avvik mellom prisindeksene. For gruppen "Steinfrukt" ligger RYGEKS prisindeks under de to andre alternativene i starten av analyseperioden, men fra januar 2008 blir Törnqvist prisindeks liggende systematisk under både RYGEKS og Jevons prisindeks og avviket forsterkes utover perioden. For gruppen "Melon" følger de tre prisindeksene hverandre fram til juli 2007. Fra og med juli 2007 blir RYGEKS prisindeks liggende godt over både Törnqvist og Jevons prisindeks ut perioden med et økende avvik. Avviket skyldes hovedsake-

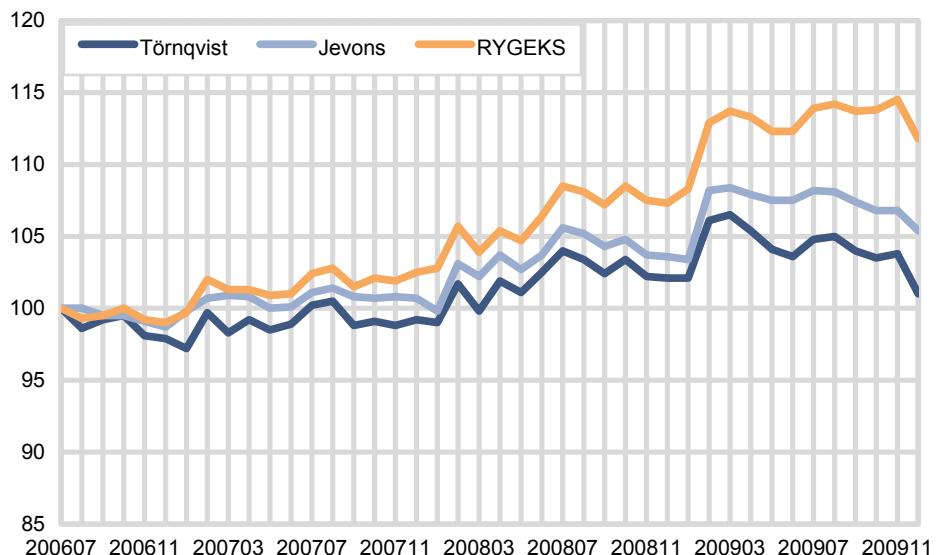
lig sesongproblematikk. Sesongvarene skaper ingen systematisk skjevhet i RY-GEKS prisindeksen slik de gjør i Törnqvist og Jevons prisindeks. Det samme gjelder gruppen ”Nøtter, frø og kjerner” som generelt sett har et sterkt innslag av sesongvarer. I indeksen for bearbeidet frukt derimot, ligger Törnqvist prisindeks noe over både RYGEKS og Jevons prisindeks i hele analyseperioden.

Figur 18. Prisindeks for frukt med Törnqvist, Jevons og RYGEKS formel. Juli 2006=100



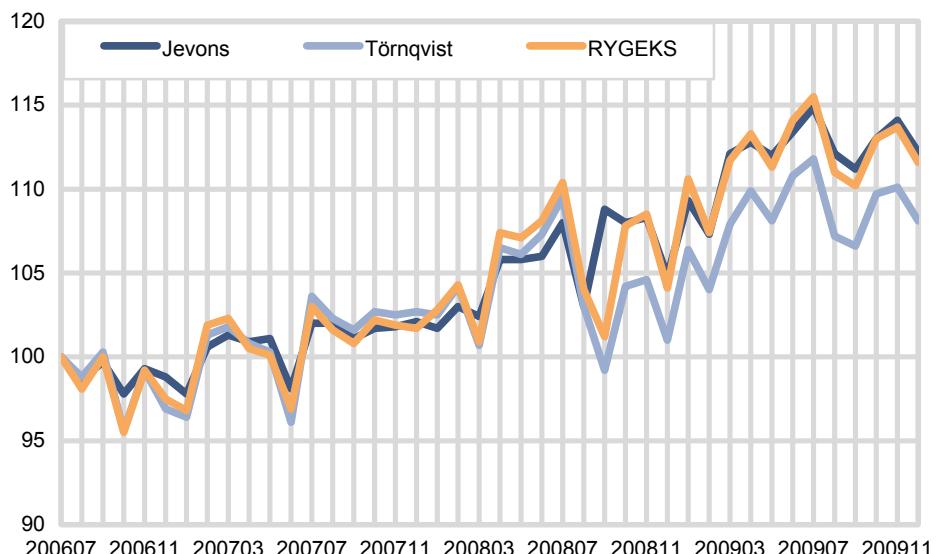
Størst avvik finner vi i gruppen ”Sukker, syltetøy, sjokolade og andre sukkervarer”. RYGEKS prisindeks ligger systematisk over både Törnqvist og Jevons prisindeks i tilnærmet hele analyseperioden. Avviket mellom Törnqvist og RYGEKS prisindeks er på hele 10,8 prosentpoeng i denne perioden. Avviket oppstår i slutten av 2006, og tiltar gradvis, særlig i begynnelsen av 2009 øker avstanden mellom prisindeksene. Det er særlig gruppen ”sjokolade, konfekt, drops” og ”Iskrem” der RYGEKS prisindeks ligger klart over både Törnqvist og Jevons prisindeks. For gruppen ”sjokolade, konfekt, drops” ligger også Jevons prisindeks klart over Törnqvist prisindeks. Avviket i gruppen skyldes flere forhold, både manglende prisobservasjoner etter tilbudsperioder, sesongvarer og hyppig utskifting av produkter. Bildet i indeksen for sukker, syltetøy, sjokolade og andre sukkervarer er imidlertid noe diversifisert - for gruppene ”sukker og søtningsmidler” er det svært små avvik mellom Törnqvist og RYGEKS prisindeks med Jevons prisindeks liggende noe under begge de andre indeksene. For ”syltetøy, marmelade og kompott” er det også svært små forskjeller mellom RYGEKS og Jevons prisindeks med Törnqvist prisindeks liggende noe under de andre prisindeksene fra 2. halvår 2008.

Figur 19. Prisindeks for sukker, syltetøy, sjokolade og andre sukkervarer med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100



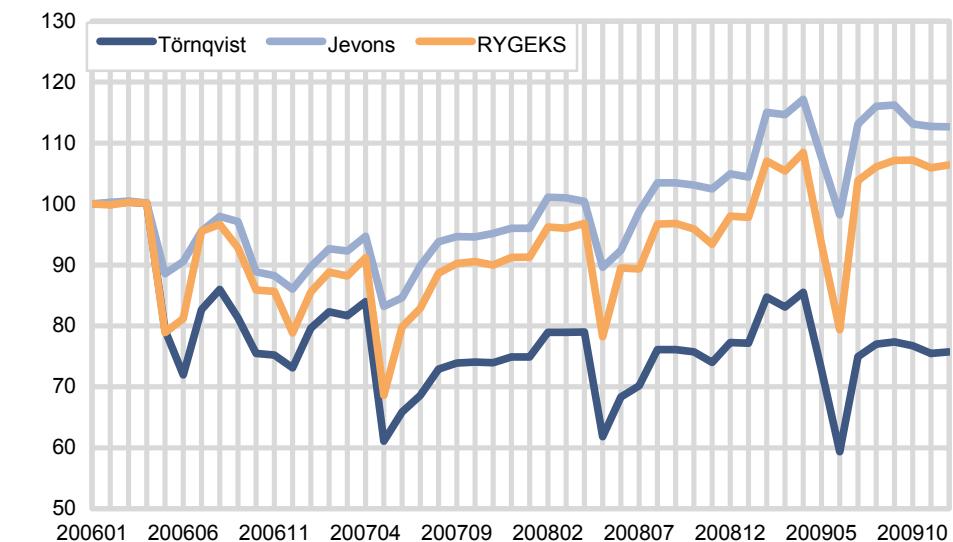
For ”Kaffe, te og kakao” derimot følger prisindeksene hverandre rimelig godt frem til og med august 2008, da alle prisindeksene faller på grunn av tilbudsaktivitet. I måneden etter henter Jevons prisindeks seg inn igjen mens de to andre prisindeksene fortsetter å falle. I oktober 2008 er RYGEKS prisindeks nesten tilbake på nivået før salg, mens Törnqvist-prisindeksen blir liggende godt under de to andre alternativene resten av analyseperioden. Dette skyldes langt på vei prisindeksen for kaffe og asymmetrisk vektlegging av prisoppgang og prisnedgang. I indeksen for mineralvann, leskedrikker og juice er det små avvik mellom Törnqvist og Jevons prisindeks med RYGEKS prisindeks liggende systematisk over fra og med starten av 2008.

Figur 20. Prisindeks for kaffe, te og kakao med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100



Prisindeksene blir i offisiell metode beregnet rett til COICOP6-nivå, det mest detaljerte aggregeringsnivået. I figur 21 har vi beregnet prisindeks ned på EAN-nivå for rømmegrøt med Törnqvist, Jevons og RYGEKS formel i perioden januar 2006 til desember 2009, jmf. figur 1 og 2 i kapittel 3.1.

Figur 21. Prisindeks for rømmegrøt med Törnqvist, Jevons og RYGEKS formel. Januar 2006-desember 2009. Indeks januar 2006=100



I figur 21 ser vi tydelig at Törnqvist prisindeks drifter nedover og dermed undervurderer den faktiske prisutviklingen. Driftingen er hovedsaklig forårsaket av tilbudsaktivitet over flere måneder noe som bidrar til å skape asymmetrisk vektlegging av prisnedgang og prisoppgang.

6. Oppsummering og konklusjon

Det ble i 2009 startet et evalueringssarbeid av KPI's prisindeks for matvarer og alkoholfrie drikkevarer. På grunn av internasjonal fokus på drifting ved bruk av kjedede superlative prisindekser, har et av formålene i dette arbeidet vært å kontrollere vår beregningsmetode for drifting. Ivancic, Fox og Diewert (2009) peker på mulig nedadgående drifting i kjedede superlative prisindekser forårsaket av konsumenter som hamstrar når varer er på tilbud og handler mindre enn normalt i en periode etter tilbud. Vi ser av norske strekkodedata at konsumentene reagerer raskt på tilbudspriser, men vi kan ikke se at denne "hamstringseffekten" er et veldig typisk trekk i norske strekkodedata. Det vi kan se er at lengden på tilbudsperioden får stor betydning. Det er ikke alltid slik at en tilbudsvarer går rett tilbake til "normalpris", prisene kan hente seg gradvis inn igjen. Slike forløp kombinert med fallende omsetningsandeler gjør at vi får asymmetrisk vektlegging av prisnedgang og prisoppgang og dermed en prisindeks som drifter nedover og dermed undervurderer den faktiske prisutviklingen. Før vi endret vår beregningsmetode i 2005 ble det gjennomført mange analyser og beregninger basert på norske strekkodedata og konklusjonene våre den gang var at vi fant klare eksempler på drifting ned på vare-nivå (EAN-nivå), men at på konsumgruppenivå var effektene mindre.

Vi ser også at manglende prisobservasjoner enten på grunn av permanent eller temporært frafall, gir skjevhet i prisindeksen og (som oftest) resulterer i for lavt målt prisvekst. Utilfredsstillende behandling av visse sesongvarer bidrar også til hovedsakelig nedadgående skjevhet i prisindeksen. Vi har siden omleggingen av matvareindeksen i 2005 vært innforstått med at vi ikke har en fullgod behandling av sesongvarer. Det har vi fortsatt ikke, men så langt forsøker vi å eliminere skjevheter forårsaket av sesongvarene ved å fjerne de mest vesentlige (i form av høye omsetningsandeler) fra prisindeksen.

For å avdekke skjevheter som drifting i vår matvareindeks har vi sammenlignet en reberegning av vår offisielle metode, en superlativ Törnqvist prisindeks, med en modifisert Jevons og RYGEKS prisindeks. Den nyutviklede beregningsmetoden RYGEKS prisindeks utnytter multilateral indeksteori for å skape superlative kjede-

indeks er uten drifting. RYGEKS prisindeks er transitiv av oppbygging som betyr at kjedeindeksen er lik den direkte indeksen og dermed per definisjon driftfri. Anses RYGEKS prisindeks som en ideell indeks eller ”benchmark” prisindeks viser våre sammenligninger klar indikasjon på nedadgående skjevheter både på aggregert og detaljert nivå ved bruk av dagens metode i den norske matvareindeksen. Differansen mellom vår offisielle metode - Törnqvist prisindeks og RYGEKS prisindeks er på i overkant av 1 prosentpoeng i gjennomsnitt i året hele analyseperioden sett under ett. Det er viktig å understreke at sammenligningene med RYGEKS prisindeks forteller oss ikke bare noe om betydningen av drifting forårsaket av svingninger i pris og omsatt mengde, men de viser oss en *samlet* effekt forårsaket av både variasjon i pris og mengde, manglende prisobservasjoner og utilfredsstillende behandling av sesongvarer (merk at RYGEKS prisindeks har alle sesongvarene *inkludert* i beregningene). Sammenligner vi RYGEKS prisindeks med vår *publiserte* Törnqvist prisindeks for matvarer og alkoholfrie drikkevarer øker differansen noe. Dette skyldes små skjevheter vi har hatt i vår revisjonsmetode.

RYGEKS prisindeks er på et tidlig stadium og som beregningsmetode må vi fortsatt klassifisere den som eksperimentell. De empiriske resultatene er likevel interessante og lovende, og det vil trolig bli mye fokus på denne beregningsmetoden i tiden framover både nasjonalt og internasjonalt. Nå er det også flere land som tester ut metoden. Til tross for sterke empiriske resultater, er det ikke aktuelt for SSB å innføre denne metoden med det første. Vi behøver mer erfaring med beregningsmetoden både nasjonalt og internasjonalt og ikke minst en internasjonal aksept av denne metoden som god rutine for bruk av strekkodedata i KPI. På kort sikt vil derimot SSB gjøre andre endringer i matvareindeksen for å redusere dagens skjevheter, som å innføre imputering av manglende prisobservasjoner og en bedre behandling av sesongvarer.

Referanser

- Eltö, Ö, og Köves, P. (1964): "On a Problem of Index Number Construction Relating to International Comparisons" (på ungarsk), *Statisztikai Szemle* 42, 507-518.
- Feenstra, R.C. og M.D. Shapiro (2003): "High Frequency Substitution and the Measurement of Price Indexes", s. 123-146 i Feenstra og Shapiro, *"Scanner Data and Price Indexes"*, Chicago: University of Chicago Press.
- Gini, C. (1931): "On the Circular Test of Index Numbers", *Metron* 9:9, 3-24.
- Haan, J. de og Grient, H. van der (2009): "Eliminating Chain Drift in Price Indexes Based on Scanner Data", artikkel presentert på Ottawa-konferansen, mai 2009.
- Haan, J. de og Opperdoes, D. (1997): "Estimation of the Coffee Price Index Using Scanner Data: The Choice of the Micro Index", artikkel presentert på internasjonal arbeidsgruppe for prisindeks, Voorburg, april 1997.
- Henriksen, K. (2006): "Utvalgsplan til konsumprisindeksens nye matvareindeks – Basert på strekkodedata". Serien Notater, Statistisk sentralbyrå.
- ILO, International Labour Office (2004): "Consumer Price Index Manual: Theory and Practice", utvidet versjon av "Consumer Price Indices: An ILO manual" (1989).
- Ivancic, L., Fox, K.J. og Diewert, W.E. (2009): "Scanner Data, Time Aggregation and the Construction of Price Indexes", artikkel presentert på Ottawa-konferansen, mai 2009.
- Norges Offisielle Statistikk (NOS C680): "Konsumprisindeksen 1995-2000".
- Pollak, R.A. (1989): "The Theory of the cost of living Index" New York: Oxford University Press.
- Rodriguez, J. og Haraldsen, F. (2005): "Den nye matvareindeksen: Bruk av strekkodedata i konsumprisindeksen", *Økonomiske analyser* 4/2005, Statistisk sentralbyrå.
- Rodriguez, J. (2004): "Strekkodedata for å beregne prisutviklingen på matvarer i KPI: Metoder, muligheter og resultater", internt notat, Statistisk sentralbyrå.
- Szulc, B. (1964): "Indices for Multiregional Comparisons" (på polsk), *Przeglad Statystyczny* 3, 239-254.
- Triplett, J.E. (2003): "Using Scanner Data in Consumer Price Indexes: Some Neglected Conceptual Considerations", s. 151-162 i Feenstra R.C. og Shapiro M.D.: *"Scanner Data and Price Indexes"*, Chicago: University of Chicago Press.
- <http://www.acnielsen.no/news/documents/PressemeldingDagligvarerapporten2010.pdf>

Vedlegg A: Klassifisering av matvarer og alkoholfrie drikkevarer, etter COICOP

COICOP	KONSUMGRUPPE
01	Matvarer og alkoholfrie drikkevarer
011	Matvarer
0111	Brød og kornprodukter
01111	Mel, frokostblanding og gryn
01111	011111 Mel
01111	011112 Ris og andre gryn
01111	011113 Frokostblanding
01111	011119 Resten av mel, frokostblanding og gryn
01112	Brød
01112	011121 Brød
01112	011122 Baguetter og rundstykker
01113	Pasta
01113	011131 Pasta
01113	011132 Nudler
01114	Kaker
01114	011141 Boller
01114	011142 Wienerbrød
01114	011143 Tarter, pai
01114	011144 Muffin, brownies
01114	011145 Kremkaker
01114	011146 Tørre kaker
01114	011147 Vafler, pannekaker og smurte lefser
01114	011148 Kakemix
01114	011149 Andre kaker
01115	Andre bakevarer
01115	011151 Flatbrød og knekkebrød
01115	011152 Pizza
01115	011153 Kjeks
01115	011154 Lomper og lefser
01115	011155 Hamburger- og pølsebrød
01115	011156 Tacos, fajitas og burritos

0112 Kjøtt		
01121 Ferskt og frosset kjøtt og flesk		
01121	011211	Storfe
01121	011212	Svin
01121	011213	Sau, lam, geit
01121	011214	Vilt
01121	011215	Fjørfe
01121	011219	Resten av ferskt og frosset kjøtt og flesk
01122 Bearbeidet kjøtt og flesk		
01122	011220	Posteier og pateer
01122	011221	Spekeskinke og spekepølse
01122	011222	Sideflesk og spek
01122	011223	Knoker, haler og kraftben
01122	011224	Bacon
01122	011225	Kjøtthermetikk
01122	011226	Pølser
01122	011227	Kjøttdeig, kjøttkaker og karbonadekaker
01122	011228	Pålegg, vakumpakket
01122	011229	Resten av bearbeidet kjøtt og flesk
0113 Fisk		
01131 Fersk og frossen fisk		
01131	011311	Torsk
01131	011312	Sei
01131	011313	Laks/ørret
01131	011314	Steinbit, sild og flyndre
01131	011315	Skalldyr
01131	011319	Resten av fersk og frossen fisk
01132 Bearbeidet fisk		
01132	011321	Saltet, tørket og røkt fisk
01132	011322	Fiskehermetikk
01132	011323	Farse av fisk og skalldyr
01132	011324	Fritert og patinert fisk
01132	011329	Resten av bearbeidet fisk

0114	Melk, ost og egg
01141	Melk, yoghurt og fløte
01141	011411 Melk
01141	011412 Yoghurt
01141	011413 Fløte
01141	011414 Rømme og kesam
01144	Ost
01144	011440 Ost
01145	Desserter og andre melkeprodukter
01145	011450 Desserter og andre melkeprodukter
01146	Egg
01146	011460 Egg
0115	Oljer og fett
01151	Smør
01151	011510 Smør
01152	Margarin og olje
01152	011521 Margarin
01152	011522 Olje
0116	Frukt
01163	Frisk frukt
01163	011631 Sitrus
01163	011632 Bananer
01163	011633 Epler
01163	011634 Pærer
01163	011635 Steinfrukt
01163	011636 Melon
01163	011639 Resten av frisk frukt
01165	Bearbeidet frukt
01165	011651 Tørket frukt
01165	011652 Hermetisk frukt
01165	011659 Resten av bearbeidet frukt
01166	Friske bær
01166	011661 Friske bær
01166	011662 Frosne bær

01167	Nøtter, frø og kjerner	
01167	011670 Nøtter, frø og kjerner	
0117	Grønnsaker, inkludert poteter og andre rotvekster	
01171	Friske grønnsaker	
01171	011711 Poteter	
01171	011712 Kål	
01171	011713 Rotgrønnsaker	
01171	011714 Tomater	
01171	011715 Agurk	
01171	011716 Salat	
01171	011717 Fersk sopp	
01171	011719 Resten av friske grønnsaker	
01177	Konserverte grønnsaker	
01177	011771 Bønner, linser og erter	
01177	011772 Potetprodukter	
01177	011773 Konserverte maisprodukter	
01177	011774 Konserverte tomater	
01177	011775 Konservert sopp	
01177	011779 Resten av konserverte grønnsaker	
0118	Sukker, syltetøy, sjokolade og andre sukkervarer	
01181	Sukker og søtningsmidler	
01181	011811 Sukkerbiter og raffinade	
01181	011812 Farin	
01181	011813 Melis og kandisukker	
01181	011814 Søtningsmidler	
01182	Syltetøy, marmelade og kompott	
01182	011821 Syltetøy/marmelade	
01182	011822 Kompotter/fruktgrøt	
01182	011829 Resten av syltetøy, marmelade og kompott	
01183	Annet søtt pålegg	
01183	011831 Honning	
01183	011832 Sirup	
01183	011833 Sjokoladepålegg	
01183	011839 Resten av annet søtt pålegg	
01184	Iskrem	
01184	011840 Iskrem	

01185	Sjokolade, konfekt, drops etc.
01185	011851 Kokesjokolade
01185	011852 Spisesjokolade
01185	011853 Konfekt
01185	011854 Drops/pastiller
01185	011855 Tyggegummi
01185	011856 Sjokoladeposer
01185	011857 Vingummi og seigmenn
01185	011859 Resten av sjokolade etc.
0119	Andre matvarer
01191	Saus, krydder og tilbehør
01191	011911 Sauser
01191	011912 Salt/krydder/urter
01191	011913 Balsamiko/eddkik
01191	011914 Dressing og dip
01191	011915 Ketchup/sennep
01191	011916 Majones og remulade
01191	011919 Resten av saus etc.
01192	Ferdige salater og middager
01192	011922 Ferdige salater
01192	011923 Ferdige middager
01192	011924 Ferdige smurte baguetter og rundstykker
01192	011929 Resten av ferdige middager etc.
01193	Supper, gele og stuing
01193	011931 Supper
01193	011932 Gele
01193	011933 Stuing
01193	011934 Middagsbasiser
01193	011939 Resten av supper etc.
01194	Andre snacks
01194	011940 Div snacks
01195	Barnemat
01195	011951 Barnemat
01195	011952 Morsmelkerstatning
01195	011953 Diettprodukter
01196	Tilbehør baking
01196	011960 Tilbehør baking

012	Alkoholfrie drikkevarer
0121	Kaffe, te og kakao
01210	012101 Kaffe
01210	012102 Te
01210	012103 Kakao
01210	012109 Resten av kaffe, te og kakao
0122	Mineralvann, leskedrikker og juice
01221	Mineralvann og brus
01221	012211 Mineralvann
01221	012212 Brus og leskedrikker
01221	012219 Resten av mineralvann og brus
01222	Saft og juice
01222	012221 Saft
01222	012222 Juice
01222	012229 Resten av saft og juice

Vedlegg B: Andel som fjernes fra COICOP6-nivå i Jevons prisindeks med "cut-off" metode

011111	Mel	60,4
011112	Ris og andre gryn	40,3
011113	Frokostblandinger	38,7
011119	Resten av mel, frokostblandinger og gryn	54,1
011121	Brød	48,9
011122	Baguetter og rundstykker	45,3
011131	Pasta	37,8
011132	Nudler	38,6
011141	Boller	43,3
011142	Wienerbrød	40,5
011143	Terter, pai	44,4
011144	Muffin, brownies	37,1
011145	Kremkaker	42,7
011146	Tørre kaker	40,0
011147	Vafler, pannekaker og smurte lefser	43,4
011148	Kakemix	43,7
011149	Andre kaker	46,0
011151	Flatbrød og knekkebrød	43,1
011152	Pizza	48,3
011153	Kjeks	39,1
011154	Lomper og lefser	53,6
011155	Hamburger- og pølsebrød	41,8
011156	Tacos, fajitas og burritos	49,3
011211	Storfle	49,1
011212	Svin	51,3
011213	Sau, lam, geit	51,4
011214	Vilt	42,1
011215	Fjørfe	64,8
011219	Resten av ferskt og frosset kjøtt og flesk	68,3
011220	Posteier og pateer	46,5
011221	Spekeskinke og spekepølse	48,8
011222	Sideflesk og spek	11,1
011223	Knoker, haler og kraftben	37,3
011224	Bacon	47,7
011225	Kjøtthermetikk	42,4
011226	Pølser	54,6
011227	Kjøttdeig, kjøttkaker og karbonadekaker	63,7
011228	Pålegg, vakumpakket	47,2
011229	Resten av bearbeidet kjøtt og flesk	55,9
011311	Torsk	49,3
011312	Sei	48,5
011313	Laks/ørret	51,6
011314	Steinbit, sild og flyndre	54,9
011315	Skalldyr	66,7
011319	Resten av fersk og frossen fisk	55,8
011321	Saltet, tørket og røkt fisk	52,5
011322	Fiskehermetikk	54,6
011323	Farse av fisk og skalldyr	49,0
011324	Fritert og patinert fisk	44,4
011329	Resten av bearbeidet fisk	43,2

011411	Melk	68,3
011412	Yoghurt	38,1
011413	Fløte	57,0
011414	Rømme og kesam	58,0
011440	Ost	56,3
011450	Desserter og andre melkeprodukter	45,6
011460	Egg	44,3
011510	Smør	54,5
011521	Margarin	46,9
011522	Olje	36,4
011631	Sitrus	60,7
011632	Bananer	48,0
011633	Epler	54,9
011634	Pærer	50,0
011635	Steinfrukt	54,9
011636	Melon	51,8
011639	Resten av frisk frukt	51,3
011651	Tørket frukt	43,8
011652	Hermetisk frukt	43,5
011659	Resten av bearbeidet frukt	0,0
011661	Friske bær	64,2
011662	Frosne bær	58,9
011670	Nøtter, frø og kjerner	46,5
011711	Poteter	56,0
011712	Kål	48,3
011713	Rotgrønnsaker	53,0
011714	Tomater	51,6
011715	Agurk	53,9
011716	Salat	52,9
011717	Fersk sopp	58,9
011719	Resten av friske grønnsaker	55,1
011771	Bønner, linser og erter	47,7
011772	Potetprodukter	42,2
011773	Konserverte maisprodukter	47,6
011774	Konserverte tomater	42,0
011775	Konservert sopp	34,0
011779	Resten av konserverte grønnsaker	44,4
011811	Sukkerbiter og raffinade	28,1
011812	Farin	60,0
011813	Melis og kandisukker	42,4
011814	Søtningsmidler	26,8
011821	Syltetøy/marmelade	38,8
011822	Kompotter/fruktgrøt	41,6
011829	Resten av syltetøy, marmelade og kompott	35,4
011831	Honnig	49,2
011832	Sirup	41,1
011833	Sjokoladepålegg	49,7
011839	Resten av annet søtt pålegg	47,5
011840	Iskrem	39,5
011851	Kokesjokolade	38,6
011852	Spisesjokolade	43,0
011853	Konfekt	43,3
011854	Drops/pastiller	38,0
011855	Tyggegummi	42,9
011856	Sjokoladeposser	46,3
011857	Vingummi og seigmenn	37,8
011859	Resten av sjokolade etc.	80,7

011911	Sauser	41,6
011912	Salt/krydder/urter	46,8
011913	Balsamiko/eddkik	31,3
011914	Dressing og dip	35,6
011915	Ketchup/sennep	51,0
011916	Majones og remulade	54,2
011919	Resten av saus etc.	12,8
011922	Ferdige salater	38,2
011923	Ferdige middager	41,7
011924	Ferdige smurte baguetter og rundstykker	62,8
011929	Resten av ferdige middager etc.	47,8
011931	Supper	40,0
011932	Gele	35,0
011933	Stuing	43,1
011934	Middagsbasiser	40,4
011939	Resten av supper etc.	0,0
011940	Div snacks	40,5
011951	Barnemat	35,9
011952	Morsmelkerstatning	38,6
011953	Diettprodukter	37,4
011960	Tilbehør baking	45,0
012101	Kaffe	63,7
012102	Te	32,8
012103	Kakao	35,8
012109	Resten av kaffe, te og kakao	24,8
012211	Mineralvann	48,8
012212	Brus og leskedrikker	56,6
012219	Resten av mineralvann og brus	61,0
012221	Saft	39,8
012222	Juice	56,1
012229	Resten av saft og juice	54,2

Figurregister

1.	Rømmegrøt 900g. Pris i kroner. Desember 2005-desember 2006.....	13
2.	Rømmegrøt 900g. Omsetning i kroner. Desember 2005-desember 2006.....	13
3.	Skivet sjampinjong i boks. Pris i kroner. Mars 2007-mars 2008	14
4.	Skivet sjampinjong i boks. Omsetning i kroner. Mars 2007-mars 2008	14
5.	Prisindeks for matvarer og alkoholfrie drikkevarer, Törnqvist prisindeks delvis uten sesongvarer og Törnqvist prisindeks uten sesongvarer. Juli 2005-september 2009. Indeks juli 2005=100	16
6.	Prisindeks for saltet, tørket og røkt fisk, Törnqvist prisindeks med og uten sesongvarer. Juli 2005-september 2009. Indeks juli 2005=100	17
7.	Prisindeks for sitrusfrukter. Publisert prisindeks, Törnqvist prisindeks delvis uten sesongvarer og Törnqvist prisindeks uten sesongvarer. Juli 2005-september 2009. Indeks juli 2005=100	18
8.	Kaffe filtermalt 250g. Pris i kroner. Januar 2005-november 2009.....	19
9.	Kaffe filtermalt 250g. Omsetning i kroner. Januar 2005-november 2009	19
10.	Prisindeks for pizza med Törnqvist og Jevons formel. Juli 2005-desember 2009. Indeks juli 2005=100	20
11.	Prisindeks for storfe med Törnqvist og Jevons formel. Juli 2005-desember 2009. Indeks juli 2005=100	21
12.	Prisindeks for Resten av sjokolade med Törnqvist og Jevons formel. Juli 2005-desember 2009. Indeks juli 2005=100.....	22
13.	Prisindeks for matvarer og alkoholfrie drikkevarer med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100	25
14.	Prisindeks for melk, ost og egg med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100	25
15.	Prisindeks for fisk med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100	26
16..	Prisindeks for kjøtt med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100	27
17.	Prisindeks for brød og kornprodukter med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100	27
18.	Prisindeks for frukt med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100.....	28
19.	Prisindeks for sukker, syltetøy, sjokolade og andre sukkervarer med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100	29
20.	Prisindeks for kaffe, te og kakao med Törnqvist, Jevons og RYGEKS formel. Juli 2006-desember 2009. Indeks juli 2006=100	29
21.	Prisindeks for rømmegrøt med Törnqvist, Jevons og RYGEKS formel. Januar 2006-desember 2009. Indeks januar 2006=100	30

Tabellregister

1.	Antall prisobserasjoner per måned i matvareindeksen etter COICOP	7
2.	Antall prisobserasjoner i prosent som er felles med basisperiode etter 1, 6 og 12 måneder. 2003	10
3.	Gjennomsnittlig månedsendring og standardavvik for Törnqvist, Jevons og RYGEKS prisindeks. Juli 2006-desember 2009	26