

Den nye matvareindeksen: Bruk av strekkodedata i konsumprisindeksen

Joaquin Rodriguez og
Frank Haraldsen

Statistisk sentralbyrå benytter fra og med augustindeksen 2005 strekkodedata i full skala til å beregne delindeksen for matvarer og alkoholfrie drikkevarer i konsumprisindeksen (KPI). Dette innebærer at denne delindeksen går fra å måle prisene på omkring 250 representantvarer til å måle prisene på over 14 000 varer. Disse blir vektet sammen ut fra salget av de ulike varene, noe som gir mer detaljerte vekter enn tidligere i matvareindeksen. Hovedelementer i den nye matvareindeksen er et nytt utvalg der butikkene er stratifisert etter butikkjede og profil, hvor tilnærmet alle varer som omsettes i markedet inngår. Et nytt nivå er introdusert i konsumklassifiseringen og indekser på dette nivået beregnes ved Fisher-formelen med månedlig kjeding. En annen fordel er at oppgavebyrden for butikkene blir langt mindre enn tidligere på grunn av automatiserte rutiner for elektronisk rapportering. Den nye metoden gir større månedlige prisvariasjoner, spesielt for varer med mye sesongvariasjon i forbruk og priser.

Innledning

I de senere årene har det både i Norge og internasjonalt vært et økende fokus på bruk av strekkodedata ved beregning av prisindekser. Årsakene er flere, men den kanskje aller viktigste er at kvaliteten på indekserne blir bedre ved at man får tilgang på mer detaljert informasjon om transaksjonene. Gjennom strekkodedata får man tilgang på transaksjonsinformasjon om blant annet tidspunkt, mengde og pris for samtlige varer i de butikkene det rapporteres fra. Mange prinsipielle og metodiske valg har blitt tatt i prosessen med å innføre bruken av strekkodedata i full skala i matvareindeksen. Noen av disse har vært nokså opplagte, mens andre igjen er kontroversielle. Den nye metoden er tatt i bruk fra og med indeksen for august 2005. I denne artikkelen dokumenteres metoder og rutiner som er tatt i bruk, samtidig som vi presenterer egenskaper og noen resultater basert på denne type data. Delindeksen for matvarer og alkoholfrie drikkevarer i KPI var tidligere basert på et utvalg av varer fra et utvalg av butikker som skulle reflektere prisendringene i hele varepopulasjonen. Tilgang på strekkodedata skaper en helt ny situasjon, som blant annet gir oss flere muligheter til å beregne prisutviklingen for matvarer og alkoholfrie drikkevarer basert på tilnærmet alle varene som omsettes i det norske dagligvaremarkedet.

Strekkodedata er elektroniske data som genereres på salgsstedet og ved salgstidspunktet. Disse dataene inneholder informasjon både om transaksjonspris, mengde, sted, tidspunkt og en kort beskrivelse av produktene. Mulighetene for å beregne mer korrekte indekser blir med dette betraktelig større. For det første er det mulig å beregne mikroindekser hvor alle varene i en aktuell konsumgruppe inngår. Videre kan en ta hensyn til endringer i volumene når en vektet prisene, noe som er svært relevant ved beregning av matvareindekser hvor prisendringer på de mest omsette varene kan føre til en betydelig skift i etterspørsel og omsetning (substitusjonseffekt).

I en vurdering av beste metodevalg vil formålet med KPI være viktig. KPI brukes primært som en indikator for utviklingen i levekostnadene i samfunnet og som input for å deflatere verditall i nasjonalregnskapet. I tillegg til dette brukes KPI som en viktig inflasjonsindikator når Norges Bank fastsetter pengepolitikken. Disse faktorene alene sier hvor viktig det er å ta hensyn til både pris og volum når mikroindeksene beregnes, se boks 1.

Statistisk sentralbyrå har i løpet av de siste årene fått tilgang til og testet en stor mengde strekkodedata fra dagligvarekjedene, med positivt resultat for å beregne matvareindeksene. Beregninger med historiske data gir oss et grunnlag for å trekke følgende konklusjoner:

- Det finnes ingen klare indikasjoner på at prisutviklingen for matvarer er signifikant forskjellig i ulike geografiske områder i Norge. Derimot viser resultatene at det er primært butikkjede og butikkprofil (supermarked, nærbutikk, lavpris,

Joaquin Rodriguez er seniorrådgiver ved Seksjon for økonomiske indikatorer (joaquin.rodriguez@ssb.no)

Frank Haraldsen er konsulent ved Seksjon for økonomiske indikatorer (frank.haraldsen@ssb.no)

Noen fakta om den norske konsumprisindeksen

Metoder og rutiner brukt i den norske konsumprisindeksen er i stor grad basert på rammeverket som internasjonale eksperter og organisasjoner på området har definert, se for eksempel International Labour Organization (2004). Manualen er et omfattende dokument hvor både innsamlings-, bearbeidings- og publiseringsrutiner beskrives og drøftes. Med dette utgangspunkt må hvert enkelt land gjøre sine valg basert på egenskapene i sine data og tilgjengelige ressurser.

Den norske konsumprisindeksen er et resultat av aggregerte indekser fra et begrenset antall varer og tjenester som skal representere hele populasjonen av konsumvarer og -tjenester. Prisene hentes inn fra hele landet ved bruk av skjema (papir og web), telefon, innsamling fra internett og elektronisk innsending fra kjedekontorer. I motsetning til mange andre europeiske land brukes det ikke egne prisinnsamlere i Norge. På den annen side er Norge i en privilegert situasjon hva gjelder tilgangen til elektroniske data, spesielt når det gjelder strekkodedata innen dagligvarehandelen.

Varene og tjenestene (representantvarer) som inngår i konsumprisindeksen er offisielt delt inn i 12 hovedgrupper etter en offisiell konsumklassifisering (COICOP). Matvarer og alkoholfrie drikkevarer er en av disse 12 hovedgruppene og representerer omkring 12 prosent av konsumprisindeksen

totalt. Hver hovedgruppe er så delt i grupper og undergrupper innenfor det offisielle systemet, se Fløttum (1999).

Indeksene på representantvarenivå beregnes ved å sammenligne alle prisene (geometriske gjennomsnitt) i en måned med prisene som ble observert i juli måned. Disse indeksene beregnes etter geografiske områder og vektet ut fra omsetningsandeler innen de ulike områder, se Johannessen (2001). Fra august 2005 vil indeksen for matvarer og alkoholfrie drikkevarer skille seg ut fra resten av KPI på akkurat dette området bortsett fra alkoholholdige drikkevarer fra Vinmonopolet og for nye personbiler, se Henriksen (2004).

Indeksene på representantvarenivå er det vi kaller mikroindeksene. Mikroindeksene aggregeres til alle nivåene i COICOP ved bruk av budsjettsandeler fra forbruksundersøkelsen (FU).

I august hvert år oppdateres basisperioden og nye og utgåtte produkter tas henholdsvis inn og ut av beregningene. I august hvert år benyttes også resultatene fra den siste forbruksundersøkelsen til å oppdatere vektene i konsumprisindeksen. Rent teknisk innebærer rutine og metodene beskrevet ovenfor at dagens KPI bygges opp ved årlige kjeding av Laspeyres indekser, se Konsumprisindeksen 1995-2000.

kiosk) som er mest relevant med tanke på prisutvikling.

- Egenskapene og kvaliteten på strekkodedataene tillater oss å bruke disse i full skala på en slik måte at alle produktene som omsettes i markedet kan inngå i beregningene, gruppert etter standard KPI gruppering (COICOP).
- Mikroindeksene for matvarer og alkoholfrie drikkevarer kan ikke lenger beregnes på representantvarenivå. Strekkoden til en rekke varer vil av ulike årsaker endre seg over tid, noe som gjør en entydig vareidentifisering tilnærmet umulig. Dette gjør det nødvendig å beregne mikroindeksene til et nytt nivå som både er tilstrekkelig homogent og har nok produkter (priser) til at vi kan lage en indeks av høy kvalitet. Det er også viktig at det nye beregningsnivået gjør det mulig å beregne alle de ulike indeksene som vi også tidligere har formidlet.
- Mikroindeksene bør beregnes ved å bruke Fisher-formelen og månedlig kjeding. Dette betyr i praksis en månedlig oppdatering av «handlekurven» ved at vi alltid sammenligner priser på identiske produkter, samtidig som vi tar hensyn til omsetningsandeler i de to periodene.

Valget av månedlig kjedete Fisher-indekser er kanskje det mest kontroversielle av de valgene som er tatt, og det finnes både klare fordeler og ulemper med et slikt valg. Vi mener at det månedlige ressursbehovet for å overvåke de månedlige kjedete Fisher-indeksene er betydelige mindre enn det som kreves for å vedlike-

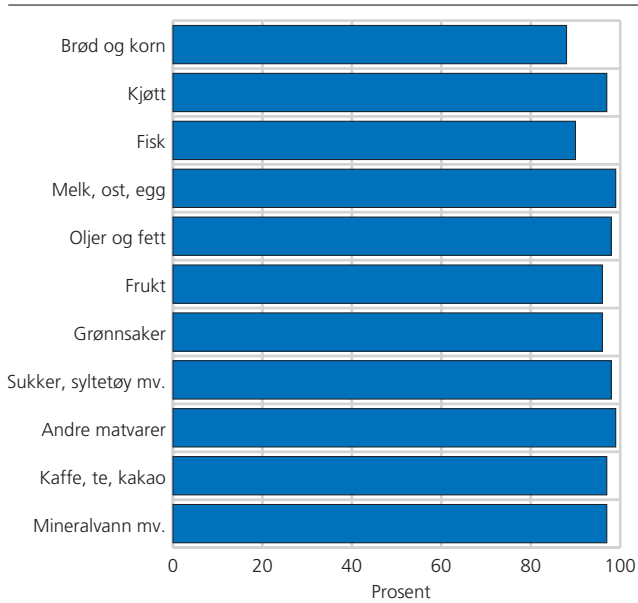
holde en fast kurv med flere tusen varer. Likevel viser erfaringer fra andre land at enkelte varegrupper, deriblant sesongvarer, bør behandles med spesiell varsomhet.

Vi skal i denne artikkelen utdype disse momentene og presentere enkelte av resultatene fra beregningene med strekkodedata.

Butikkutvalget - nye prinsipper

Konsumprisindeksen benytter et utvalg av bedrifter for å estimere prisutviklingen i alle delindeksene. I forbindelse med innføringen av den nye beregningsmetoden i matvareindeksen har det blitt trukket et helt nytt bedriftsutvalg for innhenting av priser på matvarer og alkoholfrie drikkevarer. Det nye utvalget består kun av bedrifter som kan rapportere strekkodedata via hovedkontorene. Bedrifter som kan levere strekkodedata på denne måten står for omkring 98 prosent av omsetningen av alle matvarer og alkoholfrie drikkevarer. Figur 1 viser tall fra SSBs Avanseundersøkelse for detaljhandel fra 2003. Vi ser at bortsett fra varegruppene fisk og brød og korn er omsetningsandelene som strekkodedatabedriftene representerer klart over 95 prosent.

Dette innebærer at skjevheten som nødvendigvis oppstår i utvalget som følge av at vi kun får informasjon via strekkodedata er helt marginal. Vi har dessuten ingen indikasjon på at prisutviklingen kan være vesentlig forskjellig i de butikkene som ikke kan rap-

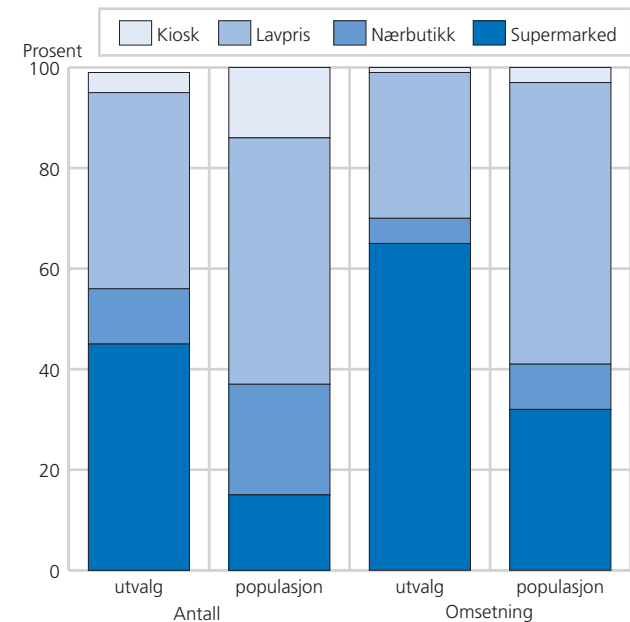
Figur 1. Omsetningsandeler som strekkodedatabedrifter representerer

Kilde: SSB Avanseundersøkelsen 2003.

portere via strekkodedata (bensinstasjoner, bakere og spesialforretninger for kjøtt og fisk).

Det nye utvalget består av 158 enheter fra en total trekkpopulasjon på 3078 bedrifter. Et utvalg på 158 bedrifter vil sikre representativitet samtidig som vi har en håndterbar størrelse på utvalget. For å være garantert en viss utvalgsstørrelse innen hver delpopulasjon, og for å redusere usikkerheten i estimatene for totalomsetningen, har utvalget blitt stratifisert. Det gamle utvalget var stratifisert etter geografisk beliggenhet, der Norge var delt inn i åtte geografiske regioner. Våre beregninger basert på strekkodedata viser imidlertid at kjedekonsept har større betydning for prisutviklingen enn hvor bedriften befinner seg rent geografisk. Dette skyldes antakelig fremveksten av sterke kjedekonsepter og at disse over noen år har hatt en relativt stabil markedssituasjon. Med dette som utgangspunkt har vi i det nye bedriftsutvalget valgt å prioritere stratifisering etter butikkjede og kjedekonsept fremfor stratifisering etter geografisk område. Konseptene vi har valgt å stratifisere etter er kiosk, lavpris, supermarked og nærbutikk, se figur 2.

Trekkepopulasjonen for hver enkelt butikkjede ble stratifisert etter kjedekonsept og enhetene i utvalget ble allokert etter en såkalt Neymans allokeringmetode. Omsetningstall fra Bedrifts- og foretakregistret ble brukt for å estimere spredning innad i strataene og utvalget ble trukket enkelt tilfeldig uten tilbakelegging innen hvert enkelt stratum¹.

Figur 2. Oversikt over konseptfordeling i utvalg og populasjon

Kilde: Bedrifts- og foretakregistret. Internett.

Fra figur 2 kan vi se at supermarkedene er overrepresentert, mens lavprisbutikkene er underrepresentert i utvalget. Dette skyldes i hovedsak at Neymans allokeringmetode tar hensyn til spredningen innenfor hvert stratum. Spredningen i omsetningstallene er stor innen stratumet supermarkeder, noe som gjør at bedrifter i dette konseptet relativt sett har stor sannsynlighet for å bli trekt ut. Det motsatte er tilfellet for omsetningsspredningen innenfor lavpriskonseptet. Konseptet er preget av butikker med nokså lik størrelse og homogene vareutvalg, noe som gjør at også bedriftenes omsetning er relativt stabil innenfor stratumet.

Et annet viktig moment med innføringen av en matvareindeks basert utelukkende på strekkodedata, er at oppgavebyrden til bedriftene reduseres betraktelig. Tidligere ble papirskjema sendt ut til hver enkelt butikk og disse butikkene måtte manuelt fylle ut prisene på en rekke varer. Elektronisk rapportering av strekkodedata innebærer at SSB får prisene samlet fra de respektive hovedkontorene. Hovedkontorenes oppgavebyrde ved å lage uttrekk av data til matvareindeksen er helt marginale sammenlignet med situasjonen tidligere.

Egenskaper ved strekkodedata for matvarer

SSB har i de senere år fått gradvis bedre tilgang til strekkodedata for de viktigste aktørene i detaljhandelen. Vårt utvalg er som nevnt trukket blant butikkjedene som sto for mer enn 95 prosent av omsetningen

¹ Et detaljert beskrivelse av arbeidet ved etableringen av utvalget skal publiseres på et senere tidspunkt.

Tabell 1. Matvarer i alt: Oversikt over månedlige priser og varer

COICOP	KPI fra august 2005 (Strekkode data)		KPI før august 2005 (representantvarer)	
	Priser	Produkter	Priser	Produkter
0111: Brød og kornprodukter	42 705	2 496	3 297	36
0112: Kjøtt	29 225	1 932	2 666	46
0113: Fisk	13 926	1 096	1 370	36
0114: Melk, ost og egg	32 751	896	2 077	22
0115: Oljer og fett	6 007	193	520	6
0116: Frukt	8 649	546	1 300	17
0117: Grønnsaker	21 828	1 230	1 825	21
0118: Sukker, syltetøy mv.	47 368	1 470	3 074	29
0119: Andre matvarer	73 134	2 978	2 126	23
0121: Kaffe, te, og kakao	9 044	368	639	6
0122: Mineralvann mv.	25 982	827	1 383	10
Total	310 619	14 032	20 277	252

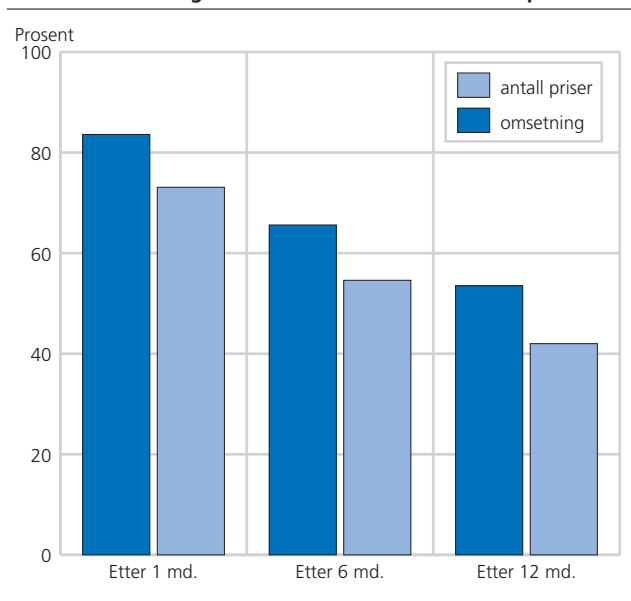
Kilde: SSB Konsumprisindeksen

innen dagligvarehandelen i 2004, og vi får hver måned totalt ca 300 000 prisobservasjoner fordelt på 14 000 varer innen gruppen matvarer og alkoholfrie drikkevarer. Tabell 1 viser datamaterialet som danner grunnlaget for å beregne indeksene for matvarer i en enkelt måned før og etter at strekkodedata ble tatt i bruk.

Varene som rapporteres gjennom strekkodedata identifiseres via EAN- og PLU-koder. EAN er en standardisert identifiseringskode som gjelder for et produkt uavhengig av i hvilken butikk produktet selges. PLU er identifiseringskoder som også gjelder for enkeltprodukter, men kun for butikker innen samme kjede. Ved hjelp av disse kodene, interne varegrupperinger i kjedene og produktbeskrivelser, har vi bygget opp en varekatalog som brukes til å koble COICOP (KPIs konsumgruppering) til hver enkelt prisobservasjon.

Strekcodedata inneholder informasjon om pris, mengde, sted, tidspunkt og en kort beskrivelse av produktene. Vareprisene er en gjennomsnittspris for hver vare i månedens «midtuke». Gjennomsnittsprisen for midt-ukene er definert som total omsetningsverdi delt på total mengde for den enkelte vare i denne uken. Dette gjelder for hver enkelt butikk og de rapporterte prisene er med andre ord beregnede transaksjonspriser. I datamaterialet blir det rapportert om varene omsettes til normalpris eller på tilbud, og dersom prisen endres gjennom uken vil en vektet gjennomsnittspris for denne varen bli reflektert i KPI.

Bruk av strekkoder som identifikasjon av varer har sine fordeler og ulemper. Kodene består av en sammensetning av 13 eller 8 siffer som inneholder informasjon om både produksjonsland, leverandør og artikkeltype. Artikkelnnummeret fastsettes av vareleverandøren, noe som betyr at når vi sammenlikner pri-

Figur 3. Matvarer i alt: Antall prisobservasjoner og omsetningsandeler som er felles med basisperiode

Kilde: SSB Konsumprisindeksen.

sene for samme EAN-kode i to forskjellige måneder er vi garantert at dette gjelder for helt identiske produkter. Et problem er imidlertid at samme type produkter kan få forskjellige EAN-koder dersom butikkene enten skifter produsent eller at produsenter skifter artikkelnummer. Figur 3 illustrerer problemet i 2004, hvor vi ser at avgangen og tilgangen av varer er meget stor i løpet av 12 måneder. I figuren vises andelen av prisobservasjoner som kommer fra identiske produkter etter hvor langt tilbake man velger basisperioden for sammenligningen.

Figuren ovenfor viser at dersom vi har en løpende basisperiode som alltid er måneden i forkant av beregningsperioden, vil om lag 73 prosent av prisobservasjonene ha samme EAN-kode i beregningsperioden som i basisperioden. Settes basisperioden til 1 år tidligere enn beregningsmåneden er det kun 42 prosent av prisobservasjonene som kobles mot basismåneden. Det betyr at dersom vi ville benytte hele datamaterialet og samtidig benytte en fast basismåned måtte vi imputere priser for 58 prosent av observasjonene etter 12 måneder. Dette resultatet er av stor betydning for valg av basisperiode for å beregne mikroindekser. Resultatene for totalen gjelder stor sett for samtlige konsumgrupper.

Et annet interessant og viktig moment ved strekkodedataene er hvorvidt strekkodedata støtter vektene som kommer frem via forbruksundersøkelsen (FU). I praksis beregnes vektene fra FU til KPI som et aritmetisk gjennomsnitt for de tre siste års data for å unngå tilfeldige store utslag i budsjettandeler for enkelte varer i enkelte perioder. Tabell 3 viser sammenhengen mellom FUs budsjettandeler og omsetningsandeler fra våre strekkodedata. Tallene fra strekkodedataene vi-

Tabell 2. Sammenligning av vektene: Forbrukundersøkelsen (FU)* mot strekkodedata

	2002		2003		2004	
	Strekkode	KPI	Strekkode	KPI	Strekkode	KPI
011: Matvarer	87,2	89,7	87,2	88,9	87,4	88,3
012: Alkohol-frie drikkevarer	12,8	10,3	12,8	11,1	12,6	11,7
K0111: Brød og kornprodukter	13,0	13,6	12,9	13,6	12,9	13,8
K0112: Kjøtt	17,6	21,5	17,3	21,2	17,9	21,0
K0113: Fisk	4,7	6,0	4,5	6,0	4,4	6,0
K0114: Melk, ost og egg	17,1	15,7	17,1	15,1	17,2	14,4
K0115: Oljer og fett	2,5	2,4	2,5	2,3	2,4	2,1
K0116: Frukt	5,0	6,8	5,0	6,7	4,8	6,6
K0117: Grønnsaker, inkludert poteter og andre rotvekster	8,6	8,6	9,1	8,6	8,7	8,8
K0118: Sukker, syltetøy, sjokolade og andre sukkervarer	10,0	9,8	10,0	10,3	10,1	10,3
K0119: Andre matvarer	8,7	5,2	8,8	5,9	9,0	6,6
K0121: Kaffe, te, og kakao	2,8	2,8	2,8	2,6	2,9	2,5
K0122: Mineralvann, leskedrikker og juice	10,0	7,5	10,0	7,7	9,7	7,9

Kilde: SSB Konsumprisindeksen og Forbrukundersøkelsen. Tallene fra Forbrukundersøkelsen kommer frem som et aritmetiske gjennomsnitt av de tre siste årene.

De nye mikroindeksene for matvarer og alkohol-frie drikkevarer

Det finnes mange mulige alternativer for å beregne mikroindekser, noe som gjelder både ved beregning av rene prisindekser, og blir enda mer aktuelt når man skal ta hensyn til både pris og volum.

Det er tre grunnleggende spørsmål man må ta stilling til. For det første må det velges et nivå for beregning av indekser, for det andre hvilken formel som skal benyttes og til slutt må basisperioden og vektene bestemmes. Tidligere ble mikroindeksene for matvarer i KPI beregnet på representantvare-nivå og var rene prisindekser. Formelen som ble brukt kalles i litteraturen Jevons-indeks og er gitt ved:

$$(1) \quad P_J(p^0, p^1) \equiv \prod_{m=1}^M \sqrt{\frac{p_m^1}{p_m^0}}$$

Det betyr at indeksen P for representantvare J er en funksjon kun av prisene p for denne varen i basisperioden 0 og beregningsperioden 1. Indeksen er gitt ved det geometriske gjennomsnitt av de M prisrelativene som er observert for denne varen. Basispris er prisen som observeres i foregående juli måned.

Utgangspunktet for beregningene av mikroindekser basert på strekkodedata er at metoden og rutinene skal være så like som mulig de som ble brukt tidligere. Som grunnleggende formel valgte vi derfor følgende vektete Jevons-indeks:

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

Formelen sier at logaritmen av en indeks P for varene som inngår i nivå K er en funksjon av prisene p i basisperioden 0 og i beregningsperioden 1, og omsetningsandelene s i periode i; i=0, 1 (eller en kombinasjon av begge)

Merk at formel (2) kommer frem ved å ta logaritmen av:

$$(3) \quad P_K(p^0, p^1, s^i) \equiv \prod_{m=1}^M \left(\frac{p_m^1}{p_m^0} \right)^{s_m}$$

Valg av perioden, i, for å beregne omsetningsandeler, s, gir oss følgende muligheter:

Paasche (IP): s_m står for omsetningsandel som vare m representerer i beregningsperioden, 1

Laspeyre (IL): s_m står for omsetningsandel som vare m representerer i basisperioden, 0

Fisher (IF): er gitt ved $(IL \times IP)^{1/2}$

Ved månedlige kjedede Fisher indekser mener vi IF hvor periode 1 står for beregningsmåned og 0 for måneden før beregningsmåned.

Indeksene på de offisielle COICOP-nivåene er et resultat av å aggregere mikroindekser ved hjelp av vektene fra forbruksundersøkelsen (FU). For eksempel er indeksen for COICOP 01, matvarer og alkohol-frie drikkevarer, gitt ved:

$$(5) \quad I_{01} = \sum_{K=1}^N P_K w_K$$

hvor N står for antall mikroindekser som inngår i konsum-gruppe 01 og w_K står for vektene som de representerer ifølge informasjon fra FU.

Med andre ord vil de nye matvareindeksene fremkomme som følger: Vi vil først beregne månedlige kjedede Fisher-mikroindekser, for så å aggregere disse videre ved bruk av vektene fra FU. Det er altså kun indeksene på mikronivå som er Fisher-indekser, og indeksene fra og med COICOP-5 nivå vil som tidligere være Laspeyre-indekser.

ser årlige gjennomsnitt som fremkommer ved å beregne månedlige omsetningsandeler.

Det er flere momenter som er interessante å peke på i sammenligningen ovenfor. For det første ser vi at andelen på undergruppenivå har holdt seg stabile gjennom disse tre årene. Vi ser også noe som er meget fordelaktig, nemlig at det ikke er betydelige forskjeller mellom omsetningsandelene fra våre strekkodedata og budsjetandelene fra FU.

Bortsett fra forskjellene i gruppene kjøtt og andre matvarer viser tabell 2 at strekkodedataenes rangering av forbrukernes preferanser og innkjøpsatferd, er svært lik det som fremkommer av FU. Dette er en klar indikasjon på at kvaliteten på datamaterialet er god, og at koblingene mellom EAN-koder og COICOP er robuste.

Fisher indekser og sesongvarer - en utfordring

Den grunnleggende forskjellen mellom de nye og de tidligere publiserte matvareindeksene er metoden for å beregne mikroindeksene. Valget av Fisher-formelen (2) i boksen oppfatter vi som uproblematisk, siden det er en klar fordel at vi tar hensyn til vektene i de to periodene som prisene måles.

Det er et faktum at omsetningsandelene for hver enkelt matvare ikke er konstant over tid. Disse andelenes ofte sesongpreget og sterkt påvirket av tilbudsaktiviteter som butikkjedene setter i gang for å øke sine markedsandeler. Det er da også slik at Laspeyre-indeks (vektene er fra basisperioden) generelt overestimerer prisutviklingen og Paasche-indeks (vektene er fra beregningsperioden) generelt underestimerer prisutviklingen når de sammenlignes med Fisher-indeks. Årsaken til dette er at selv om pris-substitusjonseffekten for matvarene er ganske stor er det ikke slik at produktene er perfekt priselastiske. Én prosent endring i prisene gir ikke en perfekt negativ korrelasjon med én prosent endring i volum. Tanken er at ved bruk av Fisher-indeks utlikner man de uheldige sidene ved Laspeyres og Paasche (I_F er et estimat mellom I_p og I_L).

Noe mer problematisk er valget av basisperiode, det vil si perioden som skal brukes for å sammenligne beregningsmånedens priser. I utgangspunkt er de fremste ekspertene innen indeksteori lite begeistret for å bruke månedlige kjeding i de tilfeller hvor både priser og volum varierer sterkt over tid, se International Labour Organization (2004), avsnitt 20.25. Hovedinnvendingen mot månedlig kjeding er den såkalte drifteffekten, som kan oppstå idet indeksene kjedes. Dersom kjedingsperioden er kort, altså at det er kort tid mellom beregningsmåned og en løpende basismåned, kan ulike indekser i to etterfølgende måneder forekomme til tross for at prisene er helt iden-

tiske. Dette ville vært et uakseptabelt resultat dersom utgangspunktet for å beregne indekser hadde vært priser på varer i en fast handlekurv. Dette er imidlertid ikke tilfellet når vi skal beregne indekser med bakgrunn i strekkodedata der handlekurven stadig blir oppdatert på bakgrunn av rapporterte omsetningsandeler.

Problemet ovenfor viser en mulig uheldig side ved bruk av månedlige kjedeindekser. Det er likevel andre momenter som gir oss grunnlag for å hevde at månedlig kjeding av Fisher-mikroindekser er et godt valg når det gjelder strekkodedata for matvarer. I utgangspunktet er tanken med kjedeindekser å dempe de problemer som skyldes tilgang og frafall av varer. Kvalitetsendringer er også en faktor som bidrar til at mange mener at tidsrommet prisene sammenliknes over bør være så kort som mulig.

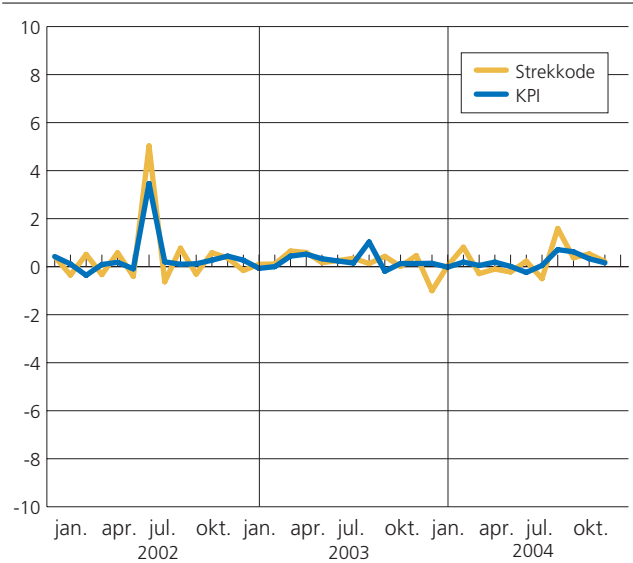
Vi har tidligere vist i figur 3 at bruk av EAN som identifiseringskode fører til at over 20 prosent av prisene ikke lar seg sammenligne fra en måned til neste. Begreper som «nye varer» og «kvalitetsendringer» får en annen dimensjon når vi snakker om strekkodedata, og bruk av månedlige kjedeindekser innebærer å fokusere på det som er hovedfordelen ved strekkodedata, nemlig sammenhengen mellom pris og mengde av helt identiske produkter. Konkurransen mellom aktørene i matvaremarkedet er stor, og vi vet at produktutvalget i butikkene er under kontinuerlig revisjon. Konsumentenes forbruksmønster endrer seg også hyppig, og månedlig kjeding av strekkodedata gir oss mulighet til å bygge opp mikroindekser som tar hensyn til løpende endringer i tilbud og etterspørsel.

Erfaringer fra andre land viser imidlertid at enkelte varegrupper må behandles med stor forsiktighet for å unngå drift i indeksen. Særlig sesongvarer og grupper som er mye preget av tilbudsaktiviteter bør behandles på en spesiell måte i den forstand at enkeltobservasjoner ikke tillates å bestemme prisutviklingen for hele gruppen. I praksis har vi bygget opp produksjonsrutiner som demper utslagene fra observasjoner som avviker mye fra hovedtrenden i de enkelte gruppene. Prosessen med å identifisere ekstremer starter med å beregne bidraget som en prisobservasjon har på endringsraten for sin gruppe. Bidraget er en funksjon av både pris og mengde og ved hjelp av en rangering² av bidragene innenfor hver enkelt gruppe identifiseres og fjernes de observasjonene som blir oppfattet som ekstremer. Hva som oppfattes som ekstremer er ulikt fra gruppe til gruppe, men det er et absolutt krav at de observasjonene som fjernes kun utgjør en svært liten andel av totalt antall prisobservasjoner i gruppen.

Resultater – noe lavere prisvekst?

Vi har testet den nye metoden for beregning av matvareindeksen med data tilbake til januar 2002. Resultat-

² Det brukes Chebyshev metode.

Figur 4. Melk, ost og egg: Strekkode data mot KPI. Prosentvis endring fra måneden før

Kilde: SSB Konsumprisindeksen.

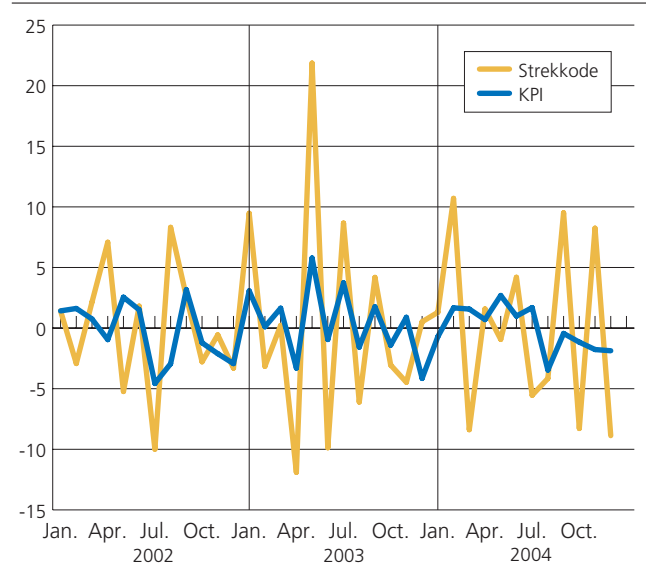
tatene som fremkommer kan ikke sammenlignes fullt ut med tallene som har blitt publisert fordi både utgangspunktet for utvalget og revisjonsrutinene av datamaterialet er noe ulikt. Resultatene med den nye beregningsmetoden viser likevel at veksten i matvareprisene i perioden januar 2002 til desember 2004 ville blitt tilnærmet lik uavhengig av om man hadde benyttet strekkodedata eller representantvarer.

Fra slutten av 2004 er bildet imidlertid noe annerledes, og prisveksten i første halvår 2005 ville sannsynligvis blitt noe lavere dersom man benyttet strekkodedata. Dette skyldes antakelig at tilbudsaktiviteter fra butikkjedene har økt i omfang og at pris-substitusjonseffekten på de mest omsatte varene har vært mye større enn tidligere antatt.

Strekkodedataene fra Norge viser at nordmenns respons på tilbudskampanjer er meget stor. Generelt er det slik at produktene som velges for tilbudskampanjer øker sine omsetningsandeler betraktelig og prisnedgangen vil dermed forsterkes på grunn av at økt volum gjør vektene til disse produktene større.

Tilbud på mindre omsatte varer vil ikke slå ut i en indeks som er basert på representantvarer, men i en indeks basert på skannerdata vil alle tilbud komme med i beregningene og prisveksten vil med dette dempes noe. Et resultat som viser at inflasjonen dempes og deflasjonen forsterkes ved bruk av vektete strekkodedata vil være i overensstemmelse med analyser i andre land, se blant annet Silver (1995), Reinsdorf (1996), Bradley, Cook, Leaver and Moulton (1997), Dalén (1997) og de Haan and Opperdoes (1997).

Bildet kan være ganske forskjellige når vi ser på prisutviklingen etter konsumgrupper. Et felles fenomen er

Figur 5. Frukt: Strekkode data mot KPI. Prosentvis endring fra måneden før

Kilde: SSB Konsumprisindeksen.

at bruken av vektete strekkodedata fører til store variasjoner over tid. Denne effekten er tydeligst for fisk, kjøtt og frukt.

Vi viser i figurene 4 og 5 to eksempler på hvordan resultatene fra den nye metoden med strekkodedata kan sammenlignes med det som er blitt beregnet via representantvarer. Melk, ost og egg er en gruppe hvor både produktene som er med og prisene typisk er stabile gjennom året. I et slikt tilfellet vil vi få nokså like resultater helt uavhengig av metode. Situasjonen er imidlertid en annen når det gjelder frukt. Gruppen har mange produkter som er sterkt sesongpreget og både priser og omsetning varierer mye gjennom året. Totalresultatet av å benytte strekkodedata og månedlig kjeding blir følgelig en mer volatil indeks, enn hva den ville vært dersom vi fortsatte å benytte tradisjonelle representantvarer.

Konklusjoner

Vi har bygget opp et system for innhenting og bearbeiding av strekkodedata for matvarer og alkoholfrie drikkevarer. I denne studien har vi analysert forskjellige egenskaper med slike data, med særlig fokus på hvordan de har blitt integrert i beregningene av KPI. En god bruk av slike data krever at de viktigste butikkjedene er riktig representert i utvalget, og at stratifisering etter butikkjede og prisprofil prioriteres fremfor stratifisering etter geografisk område.

Vår bruk av månedlige kjedede Fisher-mikroindekser tar hensyn til følgende momenter:

- Løpende endringer i etterspørselen.
- Endringer i relative priser mellom merker.
- Endringer i relative priser mellom butikker
- Endringer i relative priser mellom produkter som inngår i samme beregningsnivå

Det viser seg at omsetningsandelene er sterkt påvirket av kjedenes regelmessige tilbudskampanjer. For enkelte konsumgrupper kan dette slå uheldig ut når indekserne kjedes månedlig, noe som krever at enkelte varegrupper blir behandlet etter spesielle prosedyrer.

Med utgangspunkt i resultatene bruker SSB månedlig kjedede Fisher-mikroindekser for å beregne prisutviklingen på matvarer og alkoholfrie drikkevarer fra og med augustindeksen 2005 .

Følgende momenter har spilt en sentral rolle for de valgene som er gjort:

- Tilgang og avgang av produkter er betydelig og vi vurderer det som vesentlig for resultatene at så mye som mulig av prismaterialet benyttes - derav månedlig kjeding.
- Det er ønskelig å ta hensyn til endringene i både priser og kvantum - derav Fisher-indekser i mikroindeksene.
- Ressursbehovet for å produsere indekserne er betydelig mindre enn for de øvrige alternativer - derav månedlig Fisher. Dette er imidlertid klart betinget av at kvaliteten på fremtidens rapporteringer holder samme nivå som i dag.
- Mulighetene for å måle endringer i relative priser i flere dimensjoner (butikkjede, butikkprofil, økologiske produkter osv.)

Arbeidet med å utvide bruken av strekkoder fra bransjer som per i dag leverer priser per skjema vil fortsette. Aktuelle områder er apotekvarer og teknologivarer. For klær har det vist seg å være vanskelig å utnytte det elektroniske datamaterialet direkte i prismålinger. Fokus her vil være å utnytte dataene til andre formål, som for eksempel å vurdere vareutvalget i KPI, samt å opprettholde kontakten med kjedene med tanke på fremtidige løsninger for levering av strekkodedata.

Referanser

Bradley, R., B. Cook, S.E. Leaver and B.R. Moulton (1997), «An Overview of Research on Potential Uses of Scanner Data in the U.S. CPI». Third Meeting of the International Working Group on Price Indices, Statistics Netherlands, Voorburg, April 16-18.

Dalén, J. (1997), «Experiments with Swedish Scanner Data», Third Meeting of the International Working Group on Price Indices, Statistics Netherlands, Voorburg, April 16-18.

de Haan, J. and E. Opperdoes (1997), «Estimation of the coffee Price Index Using Scanner Data: Simulation of Official Practices», Third Meeting of the International Working Group on Price Indices, Statistics Netherlands, Voorburg, April 16-18.

Fløttum, Erling Joar (1999), «Konsumgrupperinger i offisiell statistikk». Notater 1999/59. Statistisk sentralbyrå.

Henriksen, Katharina (2004), «Ny metode for pris-måling av personbiler i konsumprisindeksen». Notater 2004/58. Statistisk sentralbyrå.

International Labour Organization (2004), *Consumer price Index Manual: Theory and Practice*, Peter Hill (ed.), Geneva: The ILO.

Johannessen, Randi (2001), «Mikroindeksformel i konsumprisindeksen». Notater 2001/64. Statistisk sentralbyrå.

Konsumprisindeksen 1995-2000. NOS C 680. Statistisk sentralbyrå.

Pollak, Robert A. (1989) «The Theory of the cost of living Index» New York: Oxford University Press.

Reinsdorf, M. (1996), «Constructing Basic Component Indexes for the U.S. CPI from Scanner Data: A Test Using Data on Coffee», BLS working Paper 277, Bureau of Labor Statistics, Washington D.C., April.

Silver, M (1995), «Elementary Aggregates, Micro-Indices and Scanner Data: Some Issues in the Compilation of Consumer Price Indices», *Review of Income and Wealth* 41, 427-438.