Interne notater

STATISTISK SENTRALBYRÅ

85/19 28. mai 1985

SESONGJUSTERING VED X11-METODEN

RAPPORT FRA EN ARBEIDSGRUPPE

AV

Morten Jensen, Vidar Knudsen, Hilde Olsen og Tore Schweder. *)

INNHOLD:

*) Vi takker Per Richard Johansen og Olav Bjerkholt for verdifulle kommentarer. Vi vil også takke Anne Thorstensen for utholdenhet i arbeidet med manuskriptet.

1. INNLEDNING.

Konjunkturovervåking er blitt et stadig mer utbredt arbeidsfelt i Norge. Mye av konjunkturovervåkingen baserer seg på den offisielle økonomiske korttidsstatistikken. Med korttidsstatistikk menes her statistikk for variable med periodelengde mindre enn ett år, f.eks. månedlig produksjonsindeks og kvartalsvis ordreindeks. Brukerne av denne statistikken ønsker å studere både trendmessige og sykliske bevegelser i økonomien, og har derfor behov for at de økonomiske tidsseriene renses for variasjoner som tilslører de underliggende tendenser. Dette fører bl.a. til et behov for å sesongjustere de aktuelle tidsseriene.

Statistisk Sentralbyrå har benyttet X11-metoden til å sesongjustere den offisielle kortidsstatistikken siden 1980, mens en metode utviklet i Byrået ble benyttet før den tid (se Brenna (1966)). Bakgrunnen for det foreliggende arbeidet er at i forbindelse med offentliggjøring av det løpende kvartalsvise nasjonalregnskapet (KNR) er det også ønskelig å sesongjustere noen av tallene for å bedre mulighetene til å gi endringene fra et kvartal til neste trendmessige tolkninger. Dette notatet er skrevet av en arbeidsgruppe som har hatt som oppgave å vurdere Byråets nåværende praksis hva angår sesongjustering - først og fremst med henblikk på behovet for sesongjustering av enkeltstående serier i KNR. Notatet er basert på en gjennomgang av en del av litteraturen på feltet og de vurderinger som har fremkommet der. Det er ment å danne grunnlaget for en videre diskusjon av sesongjustering i første rekke i Økonomisk Analysegruppe, men også i resten av Byrået da våre konklusjoner avviker fra Byråets nåværende praksis. De behov for ujusterte data som oppstår ved økonometrisk samvariasjonsanalyse, vil ikke bli særlig berørt og interesserte henvises til Biørn og Jensen (1983) for en slik drøfting og videre referanser. Se også avsnitt 4 i Biørn, Jensen og Reymert (1985) for en oversikt over behandlingen av sesongvariasjoner i kvartalsmodellen KVARTS.

Kapittel 2 inneholder noen allmene betraktninger om behovet og mulighetene for sesongjustering og referanser til noen sentrale artikler om emnet. I kapittel 3 har vi gitt noen forslag til terminologi. X11-metoden er svært uoversiktlig og det foreligger ikke, så vidt vi vet, noen lett tilgjengelig oversikt over detaljene i metoden. Kapittel 4 inneholder derfor en så nøyaktig beskrivelse av X11-metoden som vi har kunnet nå fram til. Kapitlet inneholder også endel kritiske merknader til X11-metoden generelt samt bruken av den i Byrået spesielt og er ment å være grunnlaget for konklusjonene i kapittel 5.

2. ALLMENT OM SESONGJUSTERING

Hensikten med å sesongjustere en økonomisk tidsserie er, løselig sagt, å fjerne den normale sesongvariasjonen slik at den ikke skal forstyrre vår tolkning av seriens trend, som i litteraturen om sesongjustering gjerne kalles den trendsykliske komponenten i tidsserien. Tanken er at den trendsykliske komponenten, C, varierer sakte over tid. Sesongkomponenten, S, representerer den variasjon som gjentar seg fra år til år, og den tilfeldige komponenten, I, representerer rett og slett restvariasjonen. Om man kan bryte den observerte serien, O, opp i sine tre komponenter,

$$O = C + S + I, \tag{2.1}$$

blir det mulig å studere f.eks. det sykliske forløpet, C, utviklingen i det sesongjusterte nivået C + I, og innslaget av tilfeldige begivenheter, I.

En mekanisk dekomponering av en tidsserie i elementene i (2.1), f.eks. ved hjelp av en glattingsprosedyre som X11-metoden, har ikke karakter av å være en økonometrisk eller statistisk modell. Når vi "estimerer" de tre komponentene ved X11-metoden bør en slik praksis derfor betraktes med samme skepsis som annen "measurement without theory". Likevel tyder erfaringene fra bruken av X11-metoden på at denne begrepsmessig enkle og analytisk billige dekomponeringen er formålstjenlig for mange serier . Dette gjelder først og fremst i situasjoner hvor man vil sesongjustere svært mange tidsserier slik tilfellet er med KNR. Av ressursmessige grunner vil det være nærmest umulig å foreta en tilfredsstillende modellbasert sesongjustering av hele det kvartalsvise nasjonalregnskapet. For visse viktige enkeltserier kan det imidlertid være på sin plass å benytte mer "skreddersydde" modellbaserte metoder for sesongjustering enn det X11 gir anledning til. Ved å integrere sesongkomponenten i modellbeskrivelsen, vil en kunne utnytte den apriori innsikt som ligger bak modellen til å gi bedre sesongjusterte tall. Ved hjelp av slik modellbasert økonometrisk sesongjustering vil en både kunne oppnå en klarere tolkning av den sesongjusterte serien og av den estimerte trendsykel, og en vil kunne oppnå bedre estimater for sesongkomponenten. Det kan imidlertid koste betydelig i modellutvikling og vedlikehold å få til en vellykket økonometrisk sesongjustering. Men for de mest sentrale tidsseriene bør det vurderes om ikke den gevinsten som kan oppnåes er verdt den ekstra ressursinnsatsen.

Et alternativ til den økonometriske formuleringen, er å tilpasse en rent statistisk tidsseriemodell til den enkelte serie. Ved hjelp av modeller innenfor ARIMA-klassen, vil en kunne oppnå bedre sesongjusterte tall enn ved en mekanisk bruk av X11-metoden (bedre i en nærmere presisert statistisk forstand). Det koster imidlertid også her betydelig innsats for å få til vellykket modellbasert sesongjustering. I tillegg vil den ARIMA-baserte sesongjusteringen trolig være lite robust dersom man feilspesifiserer modellen. Selv med betydelig innsats kan derfor likevel gevinstene ved å nytte ARIMA-modeller i sesongjusteringen være beskjedne.

Som en oppsummering av disse generelle betraktningene, kan en si at når mange serier skal sesongjusteres f.eks. for konjunkturanalytiske formål, er X11 eller en tilsvarende enkel, men robust, metode godt egnet. For noen få, men sentrale, serier bør en vurdere om ikke gevinstene ved en økonometrisk eller rent statistisk modellbasert sesongjustering oppveier merkostnadene. Se Pierce (1980), Pierce (1983) og Durbin & Kenny (1982) for en nærmere diskusjon av disse punktene. Det kan også være grunn til å nevne at da man tidlig i 60-årene diskuterte sesongjusteringsmetoder i Byrået, var konklusjonen at enkle filtreringsmetoder, grovt sett av samme type som X11, burde benyttes for de fleste seriene. Dette dog med det forbehold, som vi vil slutte oss til, at i den grad det er praktisk mulig bør forhåndskunnskap om den enkelte serien være medbestemmende for hvordan den skal filtreres. Se Brenna (1961 og 1966). Med de mange opsjoner standardvarianten av X11-metoden inneholder, skulle den gi tilstrekkelig rom for å utnytte slik forhåndskunnskap.

3. TERMINOLOGI

Det engelske ordet for sesongjustering er "seasonal adjustment". Det svarer derfor bedre til den engelske betegnelsen å benytte sesongjustering framfor det kanskje mer vanlige sesongkorrigering. En korrigert serie bør jo være korrekt, og det vil ofte være en lite treffende assosiasjon å knytte til den sesongjusterte serien.

Det vil som regel være ønskelig å revidere de løpende sesongjusterte seriene etter hvert som vi får flere observasjoner av tidsserien. Intuitivt henger det sammen med at den sesongvariasjon vi ønsker å trekke ut, er den som gjentar seg fra år til år i en årrekke omkring det aktuelle tidspunkt. Når den løpende sesongjusteringen foretas, vil vi imidlertid bare ha kunnskap om den tidligere variasjonen. Etter hvert som vi får flere observasjoner av tidserien, vil vi imidlertid i ettertid kunne trekke ut den sesongvariasjon som er typisk for en årrekke omkring det aktuelle tidspunkt. Vi vil velge å kalle disse endelige sesongjusterte tallene for sesongreviderte tall.

Differensen mellom (de løpende) sesongjusterte tallene og de tilsvarende sesongreviderte tallene vil vi kalle sesongrevisjonen. Ofte vil det være av interesse å knytte et usikkerhetsmål til det sesongjusterte tallet. Rimeligvis burde dette si noe om den forventede revisjonen. Dersom sesongrevisjonen oppfattes som en stokastisk variabel, vil en kunne angi revisjonsusikkerheten ved et prediksjonsintervall med gitt konfidensgrad eller ved et standardavvik, eventuelt målt i prosent av trendnivået.

Vårt forslag til terminologi er da

sesongjustere	=	løpende sesongjustering av nye tall
sesongrevidere	=	endelig sesongjustering av den historiske serie
sesongrevisjon	=	sesongrevidert tall minus sesongjustert tall
revisjonsusikkerhet	=	et mål på variasjonen omkring null i fordelingen til sesongrevisjonen

4. NÆRMERE OM X11-METODEN

I dette kapitlet skal vi i relativt stor detalj gjennomgå X11-metoden for sesongjustering. Denne metoden benyttes i svært mange land til produksjon av offisielle sesongjusterte tall. Den ble utarbeidet ved U.S. Bureau of the Census i slutten av 1960-årene og en detaljert teknisk dokumentasjon finnes i Shiskin, Young og Musgrave (1967).[1)]

Utgangspunktet for X11 er den enkle oppspaltingen (2.1): $O = C + S + I$, som altså deler den observerte serien O opp i de tre uobserverbare komponentene trendsykel: C, sesong: S, og tilfeldig komponent: I. Ofte vil

---

1) X11 står for eksperimentversjon nr. 11 av metode II for sesongjustering og er resultatet av en lang utviklingsprosess. Det første kjente forsøk på å sesongjustere økonomiske data ble foretatt av James W. Gillbart så tidlig som i 1854. Han studerte sesongvariasjon i pengeetterspørselen i England. Se Hylleberg (1984) avsnitt 1.2 for historikk omkring sesongjustering.

det være aktuelt å la O være logaritmen til den observerte serien, slik at (2.1) representerer en multiplikativ og ikke en additiv dekomponering av originalserien. I prinsippet kunne en tenke seg at denne dekomponeringen også kunne gjøres på andre transformasjoner av den opprinnelige serien. For eksempel kunne en benytte potenstransformasjon med eksponent γ. Da vil

$$O^{\gamma} = C + S + I,$$

og dersom $\hat{S}$ er et estimat for S, blir den sesongjusterte serien på opprinnelig skala

$$\widehat{C+I} = (O^{\gamma} - \hat{S})^{\frac{1}{\gamma}}.$$

Den logaritmiske transformasjonen er et grensetilfelle av potenstransformasjonene når $\gamma \rightarrow 0$. Verdier av γ mellom 0 og 1 innebærer at serien har en kombinasjon av et multiplikativt og et additivt sesongmønster. Hvis γ er nær 0, er den multiplikative komponenten dominerende. Tilsvarende er den additive komponenten dominerende hvis γ er nær 1.

Bruk av X11-metoden gir estimater for hver av komponentene i (2.1). Vi får derfor i en og samme operasjon utført sesongjustering og trendsykelestimering, og vi får estimert den tilfeldige eller irregulære komponenten, som kan gi en indikasjon på om det er et "overraskende" utslag i serien. Dersom vi har fått utført X11-oppspaltingen for et fortidig tidspunkt, slik at både trend og sesong har kunnet bli estimert ved hjelp av symmetriske og ikke ensidige filtere, vil vi oppnå sesongreviderte tall (og ikke det vi har kalt sesongjusterte tall).

For å forklare ideen med lineære filtere, eller synonymt, glidende gjennomsnitt, skal vi nå konsentrere oppmerksomheten om estimering av trendkomponenten, og vi skal gjøre det for et tidspunkt sentralt i tidsrekken. Bruk av X11 innebærer som nevnt at en ikke spesifiserer en økonometrisk eller statistisk tidsrekke-modell, men at en nøyer seg med en mekanisk dekomponering av originalserien. Dette betyr at en opererer med et trendbegrep som er omfattende nok for å fange inn mangeartede situasjoner, men som da til gjengjeld er vagt. Hovedsaken er at trendsykelen skal være noenlunde glatt og at den skal endre seg sakte i forhold til orginalserien. Slike glattede kurver kan over korte intervaller tilnærmes med et polynom av lav orden, og man sier da gjerne at man fastlegger en lokal trend i et observasjonsvindu. Koeffisientene i dette polynomet vil variere sakte ettersom en trekker intervallet eller vinduet gjennom hele den tidsperioden serien finnes. Vi snakker derfor om den lokale trendsykelen som verdien på det lokalt tilpassede polynomet i midtpunktet av intervallet. Det legges altså et vindu symmetrisk om hvert tidspunkt og koeffisientene i det lokale polynomet estimeres. Denne estimeringen kan gjøres ved hjelp av lineær regresjon. Siden regresjonsestimatene er lineære i observasjonene, ender vi opp med en lineær estimator for den lokale trenden. Det kan vises at denne framgangsmåten er ekvivalent med å "glatte" tidsserien v.h.a. et lineært filter eller et veiet glidende gjennomsnitt, som vi trekker serien gjennom. Estimatet på trendverdien på tidspunktet t blir således:

(4.1) $C_t = m_{-k}O_{t-k} + m_{-k+1}O_{t-k+1} + \dots + m_0O_t + \dots m_kO_{t+k}.$

Fotindeksen t står for tidspunkt. Dette glidende gjennomsnittet (vi skal la t løpe) er av lengde 2k+1 og det har $m_{-k}, \dots, m_k$ som vekter. Jo større vinduslengde, (2k+1), vi benytter i estimeringen av den lokale trenden, desto glattere blir trendestimatet.

Siden vi ønsker å legge et vindu symmetrisk omkring et gitt tidspunkt, får vi et odde antall, 2k+1, punkter i vinduet. Dette gir oss et lite problem fordi mange av de periodene vi ønsker å ta gjennomsnitt over, består av et like antall punkter; årets 12 måneder eller 4 kvartaler. Dersom en nå for eksempel skulle ønske et uveiet gjennomsnitt over en periode av lengde ett år i en månedlig serie, er det umulig å få dette gjennomsnittet sentrert i en måned. En løsning på problemet er å betrakte en periode på 13 måneder og la observasjonene i vinduets to ytterkanter ha halv vekt. Dette er ekvivalent med å ta gjennomsnittet av de to årsgjennomsnittene sentrert henholdsvis en halv måned før og en halv måned etter den måneden vi ser på. Vi vil kalle dette for et 13-måneders sentrert glidende gjennomsnitt. De symmetriske filtrene i X11 skal vi se nærmere på i avsnitt 4.1.

Dersom tidspunktet t er det siste vi har i tidsrekken, blir det naturligvis umulig å benytte et symmetrisk filter som (4.1). Den lokale trenden vil da måtte estimeres ved et ensidig filter, det vil si at vektene $m_{-1}, m_{-2}, \dots, m_{-k}$, alle må settes lik null. De andre vektene vil da måtte bestemmes med dette for øyet. Problemer omkring behandlingen av endepunkter ved løpende sesongjustering vil vi ta opp i avsnitt 4.3.

X11-metoden benytter seg ikke bare av lineære filtere. For at sterkt avvikende (muligens feilaktige) observasjoner ikke skal ha for stor innflytelse på de estimerte trendsykel- og sesongkomponentene, benyttes spesielle korreksjonsrutiner. Dette gjøres etter liknende prinsipper som i regresjonsanalysen (robust regresjon), der sterkt avvikende observasjoner, såkalte "outliers", a priori tillegges liten vekt eller lukes helt bort når strukturen skal estimeres. Dette er viktig for at ikke trendsykel og sesong skal preges for mye av de ekstreme observasjonene, slik at de mer ordinære observasjonene feilaktig vil framstå som mer avvikende enn de er, og ekstremobservasjonene som mindre avvikende. På grunn av denne neddempingen av de ekstreme observasjonene blir trend og sesong estimert mer robust (mot slike utslag) enn om rene lineære filtere var benyttet. Den prisen som må betales for dette, er at X11-metoden ikke er en lineær metode. Denne ikke-lineariteten skaper visse problemer både når det gjelder tolkning av metoden og når det gjelder forholdet mellom sesongjustering og aggregering.

En "outlier" kan skyldes flere forhold. Det kan ha skjedd noe uventet f.eks. spesielle værforhold eller en streik, eller det kan rett og slett være en feil. Videre kan det skyldes bevegelige helligdager (påsken) som kan kreve spesielle korreksjoner. Dette kommer vi tilbake til i avsnitt 4.2.

Til slutt i dette kapitlet skal vi diskutere noen problemer knyttet til sesongjustering av aggregerte tall (avsnitt 4.4) og presentere noen eksempler på bruk av X11-metoden (avsnitt 4.5).

4.1. Beregning av trend- og sesongverdien sentralt i serien.

Vi skal nå gjennomgå den serie av glidende gjennomsnitt som benyttes i X11. Vi starter med en råserie, O_t, som består av månedlige observasjoner. På de punkter hvor X11-metoden behandler kvartalsdata forskjellig fra månedstall, er dette nevnt. For enkelthets skyld benevner vi alle vektene med bokstaven m og lar det framgå av tallverdiene når de er ulike.

A.

Med grunnlag i originalserien beregnes et 13-måneders sentrert glidende gjennomsnitt som et første estimat på trendsykelen i serien:

$$C_t^1 = \sum_{k=-6}^{6} m_k O_{t-k} \tag{4.2}$$

Vektene i det glidende gjennomsnittet er like (unntatt i endene hvor de har halv verdi) og summerer seg til 1; dvs. (1/24, 1/12, ..., 1/12, 1/24). For kvartalsdata benyttes et 5-perioders sentrert glidende gjennomsnitt, som dermed har vektene (1/8, 1/4, 1/4, 1/4, 1/8). Et første estimat på den trendjusterte serien, dvs. summen av sesongkomponenten og den irregulære komponenten, får vi da ved

$$(4.3) \quad Y_t^1 = O_t - C_t^1$$

B.

Et første estimat på verdiene til sesongfaktorene beregnes deretter ved å bruke et veiet "fem-punkts periodisk" glidende gjennomsnitt på den foreløpige trendrensede serien Y_t^1. Dette innebærer at sesongkomponenten for f.eks. januar beregnes ved å ta et veiet gjennomsnitt av januar-verdiene i fem på hverandre følgende år. Vi får altså at

$$(4.4) \quad S_t^1 = \sum_{k=-2}^{2} m_k Y_{t-12k}^1$$

Vektene i dette gjennomsnittet er

(0.111, 0.222, 0.333, 0.222, 0.111)

(4.4) er et første anslag på sesongkomponenten siden (4.2) er et foreløpig estimat for trenden. (Vektene i gjennomsnittet framkommer som et (3x3) gjennomsnitt, dvs. et 3-leddet uveid gjennomsnitt av et 3-leddet uveid gjennomsnitt. Lengden på det bevegelige gjennomsnittet er opsjonelt i X11.) For å sentrere sesongen, dvs. sikre at summen av sesongfaktorene er null, foretar vi transformasjonen

$$(4.5) \quad S_t^2 = S_t^1 - \sum_{k=-6}^{6} m_k S_{t-k}^1$$

Vi bruker her samme type uveide sentrerte glidende gjennomsnitt som i A. S_t^2 vil nå tilnærmet ha egenskapen

$$\sum_{k=0}^{11} S_{t-k}^2 = 0$$

På dette punkt i X11-metoden foretas korrigeringen for ekstreme observasjoner. Dette er det gjort nærmere rede for i avsnitt 4.2.

Den første sesongjusterte versjon blir nå

$$(4.6) \quad O_t^{SJ1} = O_t - S_t^2$$

C.

I neste skritt (og altså etter ekstremverdikorreksjoner) beregnes et nytt estimat på trendfaktoren ved å bruke et såkalt Henderson's glidende gjennomsnitt, oppkalt etter en aktuar som i 1916 publiserte en artikkel der han utviklet dette spesielle glidende gjennomsnittet som vi nå skal se litt nærmere på.

Foran gjennomgikk vi den trendestimeringsmetoden som går ut på å føye et lokalt sett av observasjoner til et polynom av gitt grad. Det kan vises at når føyningen skjer ved vanlig minste kvadraters metode, så får vi entydig bestemt koeffisientene i det ekvivalente bevegelige gjennomsnittet. I de fleste tilfeller der denne metoden er benyttet, har en valgt å spesifisere et 3. grads-polynom. Dette gjelder f. eks. Spencer's glidende gjennomsnitt (se Kendall [1973] s. 36), som ble brukt i tidligere versjoner av sesongkorrigeringsmetoder i U.S. Bureau of the Census. Valg av et 3. grads polynom innebærer at det korresponderende filteret (det glidende gjennomsnittet) slipper en kubisk trendkurve uforandret igjennom, dvs. at vi både kan ha et maksimums- og et minimumspunkt i vinduet. Også Henderson's glidende gjennomsnitt bygger på føyning av observasjonene i vinduet til et 3. grads polynom, men i tillegg åpnes det for at vi benytter veiet regresjon ved føyningen. Dette virker intuitivt som en rimelig utvidelse. Det må jo være en fordel å kunne gi observasjonene X_{t-k}, X_{t-l+1}, ...X_t,...., X_{t+k-1}, X_{t+k} forskjellig vekt når vi skal anslå trendverdien på tidspunkt t, f. eks. kunne vi tenke oss å la de "midterste" observasjonene få større vekt enn de langt ut på sidene osv. Spørsmålet er da hvordan vi skal bestemme vektene. Vektene i Henderson's glidende gjennomsnitt bygger på et krav om at den resultatserien vi får når vi benytter gjennomsnittet (den "glattede" serien) skal være så glatt som mulig. Dette presiseres ved at variansen til tredjedifferensene av den glattede serien skal minimeres (og dette kan vises å være ekvivalent med å minimere kvadratsummen av koeffisientene i gjennomsnittet, jmf. punkt 4.3). Utledningen av formelen for koeffisientene i det glidende gjennomsnittet er vist i appendixet til Durbin og Kenny (1982). Lengden på det glidende gjennomsnittet eller vinduet følger imidlertid ikke av de kravene som er nevnt over. Denne må fastsettes ut fra andre vurderinger. I X11 blir lengden av Henderson's glidende gjennomsnitt bestemt som et kompromiss mellom behovet for å fjerne den irregulære komponenten ved å glatte den ukorrigerte serien og behovet for å reprodusere den systematiske komponenten ved å bruke et kort glidende gjennomsnitt. På en "glatt" originalserie blir det derfor brukt et kort filter, mens det på en serie som inneholder mye støy er nødvendig å bruke et langt filter for å få fjernet mest mulig av den irregulære komponenten. X11 velger for månedsserier mellom lengder på 9, 13 og 23 ledd.[1)] Koeffisientene i disse symmetriske gjennomsnittene er (den sen-

1) Det 9(23)-leddete Henderson's bevegelige gjennomsnittet benyttes hvis forholdet mellom gjennomsnittlig prosentvis endring i den irregulære komponenten og gjennomsnittlig prosentvis endring i trendkomponenten er mindre enn eller lik 0,99 (større enn eller lik 3,5), mens det 13-leddete gjennomsnittet benyttes ellers. Opsjonelt kan lengden fastsettes etter brukerens eget ønske.

trale koeffisienten er bakerst og koeffisientene på høyre side av dette sentrum er sløyfet):

9 ledd: - 0.041, - 0.010, 0.119, 0.267, 0.330
13 ledd: - 0,019, - 0.028, 0.0, 0.066, 0.147, 0.214, 0.240
23 ledd: - 0.004, - 0.011, - 0.016, -0.015, - 0.005, 0.013, 0.039, 0.068, 0.097, 0.122, 0.138, 0.148

For kvartalsserier brukes alltid et 5-leddet Henderson-glidende gjennomsnitt med vektene - 0.073, 0.294, 0.558.

Trendsykel-estimatet bestemmes ved et Henderson's glidende gjennomsnitt av den foreløpige sesongjusterte serien:

$$(4.7) \quad C_t^2 = \sum_{k=-K}^{K} m_k \, O_{t-k}^{SJ1} \qquad (K = 4,6 \text{ eller } 11)$$

0.

En ny trendjustert serie beregnes som

$$(4.8) \quad Y_t^2 = O_t - C_t^2$$

Det endelige anslaget på sesongkomponenten beregnes så ved et sju-punkts-periodisk glidende gjennomsnitt på serien Y_t^2 . Vi får da:

$$(4.9) \quad S_t^3 = \sum_{k=-3}^{3} m_k \, Y_{t-12k}^2$$

Vektene i dette symmetriske glidende gjennomsnittet er

0.067, 0.133, 0.200, 0.200.

(Vektene fremkommer som et (3x5)-glidende gjennomsnitt, dvs. et 5-leddet glidende gjennomsnitt med like vekter av et 3-leddet glidende gjennomsnitt med like vekter. Lengen på det bevegelige gjennomsnittet er opsjonelt.) På samme måte som i punkt 8 foretas det en sentrering slik at

$$(4.10) \quad S_t = S_t^3 - \sum_{k=-6}^{6} m_k \, S_{t-k}^3$$

Den endelig sesongreviderte serien blir da:

$$(4.11) \quad O_t^{SJ} = O_t - S_t$$

Vi har dermed gjennomført en rekke lineære symmetriske filtere på originalserien, som alt i alt danner ett lineært symmetrisk filter. Wallis [1982] viser at filterkoeffisientene i den symmetriske delen av X11 strekker seg 84 måneder (7 år) til hver side. For kvartalsvise serier strekker filteret seg 28 kvartaler til hver side.

Som vi har sett, kan den symmetriske delen av X11-filteret spesifiseres litt forskjellig. For hver spesifikasjon kan det vises at det finnes en bestemt ARIMA-modell slik at X11-filteret gir den optimale sesongjustering innenfor modellen. Det viser seg at X11-metoden er relativt robust mot avvik fra den korresponderende ARIMA-modellen. Se Pierce (1980) og Durbin og Kenny (1982).

Vi vil til slutt i dette avsnittet nevne at X11-programmet også beregner to oppsummerende mål for styrken i dekomponeringen av originalserien. For det første utføres en F-test for eksistensen av et stabilt sesongmønster. Testen baserer seg på størrelsesforholdet mellom sesongkomponenten og den irregulære komponenten. Testen forutsetter bl.a. at sesongsvingningene er stabile, og er følsom overfor brudd på denne forutsetningen. For det andre beregner X11 såkalte "periods of cyclical dominance" (PCD). PCD er definert som det tidsrom som trengs for at endringen i den trendsykliske komponenten overstiger gjennomsnittlig (absolutt) utslag i den tilfeldige komponenten. Dersom trendutviklingen dominerer den irregulære komponenten vil PCD være et lite tall og man behøver bare noen få perioder for å avgjøre om trenden peker oppover. Tilsvarende vil et høyt tall for PCD tilsi at man skal være varsom med å gi utsagn om trendutviklingen bare på grunnlag av noen få perioders observasjoner.

4.2 Korreksjoner for ekstremverdier, variasjoner i antall virkedager o.l.

Det er ulike synspunkter på i hvilken utstrekning det bør korrigeres for ekstremverdier. Klein (1976) argumenterer for at det bare skal korrigeres for eksogene hendelser slik som forskjell i antall kalenderdager. Streiker og andre forhold som ikke er direkte kalenderavhengig skal således ikke behandles på noen særskilt måte. Dette har sammenheng med Klein's synspunkt om at offisiell statistikk bare skal renses for "rene" sesongeffekter. Trendsykelen skal ikke forsøkes tallfestet og offentliggjort på noen måte, fordi skillet mellom trendsykel og irregulær komponent ikke er entydig. I Pierce (1980, p. 131) argumenteres det derimot for det positive i å dempe den innflytelse som enkelte sterkt avvikende observasjoner kan ha på totalresultatet. Dette svarer til at en i regresjonsanalyse ofte vil ønske å dempe ned den innflytelse ekstreme observasjoner har på koeffisient-estimatene. Dette er av særlig betydning i den mer data-analytiske situasjon som vi befinner oss i når vi sesongjusterer, hvor vi ikke har spesifiserte antagelser hverken om sesong-og trendkomponentens form eller om feilfordelingen i den irregulære serien.

X11-programmet inneholder opsjoner som gjør det mulig å foreta korreksjon for spesielle trekk ved dataserien før selve sesongjusteringen begynner. Brukeren kan f. eks. be om en prekorrigering av tidsserien for å ta hensyn til varierende antall arbeidsdager i den perioden serien refererer seg til. Slik "trading-day adjustment" tar utgangspunkt i vekter for hver av ukedagene, som enten beregnes av X11 eller gis av brukeren. På grunnlag av disse vektene beregnes månedlige faktorer, som divideres inn i dataserien for å fjerne virkedag-variasjoner. X11 inneholder også en opsjon for å oppgi et sett av månedlige justeringsfaktorer som divideres inn i dataserien. "Trading-day adjustment" kan bare brukes i forbindelse med månedsdata. I programmet slik det vanligvis nyttes, inngår dessuten rutiner som korrigerer ekstremverdier på grunnlag av variansberegninger på dataserien. Vi skal se litt nærmere på dette.

For å vurdere om et utslag er "ekstremt" og behøver justering trekkes sesongfaktorene S_t^2 i (4.5) fra estimatet på den trendjusterte komponenten, slik at vi får et estimat for den irregulære komponenten, I_t^1. Deretter beregnes et 5-års glidende standardavvik (σ_I) på estimatet, og verdiene av den irregulære komponenten i det midterste av disse årene sammenliknes med σ_I. Verdier av den irregulære komponenten som er større enn 2.5 σ_I anses som "ekstreme" og fjernes. Deretter beregnes det glidende 5-års standardavviket en gang til, og de irregulære komponentene gis vekter etter hvor store de er i forhold til σ_I:

$|I^1| > 2.5\ \sigma_I$ får vekt 0

$1.5\ \sigma_I < |I^1| < 2.5\sigma_I$ får vekter som avtar lineært fra 1 til 0.

For verdier som får mindre enn full vekt, blir den korresponderende komponenten som består av sesong- og tilfeldig utslag, erstattet med et gjennomsnitt av den aktuelle verdien (multiplisert med sin vekt) og de to foregående og etterfølgende (fullvekts-) verdien av komponenten for måneden.

Korrigering for ekstremverdier introduserer ikke-linearitet i glattingsfilteret og ikke-lineariteten varierer med hyppigheten av ekstreme observasjoner i en gitt serie. Videre er det å foreta korreksjonene i seg selv, og størrelsene 1.5 og 2.5 på "cut-off"-verdiene gjenstand for valg; de er opsjoner på samme måte som en additiv eller multiplikativ sesongformulering er det. Dette innebærer et subjektivt element i sesongjusteringen, men gir samtidig brukeren en stor grad av fleksibilitet.

I tillegg til denne innebygde rutinen inneholder X11-programmet også opsjoner for å korrigere for spesielle unormale hendelser, f.eks. streik. Disse korrigeringene følger samme mønster som korrigeringen for ekstremverdier, men blir gjort seinere i programmet, dvs. på mer bearbeidede tall. Det tas da utgangspunkt i trendsykel-estimatet C_t^2 fra (4.7). C_t^2 trekkes fra det foreløpige sesongjusterte estimatet O_t^{SJ1} for å få et estimat på den irregulære komponenten. Denne behandles så på samme måte som nevnt over. For verdier som får mindre enn full vekt, blir den korresponderende sesongjusterte komponenten erstattet med et gjennomsnitt av den aktuelle verdien multiplisert med sin vekt og de to foregående og etterfølgende (fullvekts-) komponentene for samme måned (kvartal). Et nytt trend-estimat beregnes så ved å anvende Henderson's gjennomsnittet på den modifiserte sesongjusterte serien.

4.3 Løpende sesongjustering.

Den utstrakte bruken av lange symmetriske gjennomsnitt innebærer at det bare er mulig å sesongjustere de sentrale observasjonene i serien med de "høyverdige" symmetriske filtrene. Som nevnt i avsnitt 4.1 trenger programmet 7 år med observasjoner på hver side av den aktuelle perioden for å kunne utnyttes fullt ut. For brukere er det imidlertid først og fremst viktig å skaffe sesongjusterte verdier på slutten av serien.

Tidligere beskrev vi hvordan trendsykelen kan beregnes ved føyning av polynomer til lokale observasjonssett. Der ble det estimerte polynomet brukt til å bestemme trendverdien i det midterste punktet i observasjonssettet. Det er imidlertid klart at vi også kan bruke dette polynomet til å bestemme trendverdien i andre punkter i observasjonssettet. Spesielt kan vi da benytte et polynom estimert på de (2k+1) siste observasjonene i en serie til å bestemme trenden ikke bare i midtpunktet, men også i de m siste punktene i serien. Kendall (1973) viser at slik føyning er ekvivalent med å ta asymmetriske glidende gjennomsnitt av de siste observasjonene i serien. Vi taper noe ved denne trendestimeringen for endepunktene i forhold til det vi får for de sentrale verdiene. Det viser seg at jo nærmere vi kommer enden av serien, jo mer ulike blir koeffisientene i de glidende gjennomsnittene og jo større blir deres kvadratsummer. Dette betyr at trendbestemmelsen blir mindre pålitelig ved endepunktene, målt ved den såkalte "error-reducing power" (dvs. den reduksjonen i residualvariansen som følger av å bruke det glidende gjennomsnittet på en serie). Kendall uttrykker dette slik: "The fitted curve, it has been said, tends to wag its tail", Kendall (1973), s. 38.

Som nevnt i avsnitt 4.1, pkt. A, beregnes først et 13-perioders sentrert glidende gjennomsnitt av den opprinnelige serien, dvs. C_t^1, gitt ved (4.2). Et slikt gjennomsnitt trenger 6 verdier på hver side av den aktuelle måneden, slik at vi her mister 6 verdier i hver ende av serien. Dersom siste observasjon i originalserien er tidspunkt T, mister vi verdier for tidspunktene T-5, T-4, ..., T-1, T. Erstatning av disse verdiene skal vi komme tilbake til; det gjøres ikke før mot slutten av prosessen. Siden

Siden C_t^1 mister 6 verdier i hver ende,vil det samme også gjelde Y_t^1, det første estimatet på sesong + irregulær komponent, som er gitt ved $O_t - C_{t_1}$.

Videre beregnes et fem-punkts periodisk glidende gjennomsnitt av Y_t i (4.4). Dette gjennomsnittet trenger verdier for 2 år på hver ende av serien, dvs. at vi mister verdiene for tidspunktene T-29, T-28, ..., T-7, T-6. Disse verdiene erstattes på følgende måte. For de 12 første av de 24 verdiene, dvs. de som gjelder tidspunktene T-29, T-28, ..., T-18, erstattes det fem-punkts glidende gjennomsnittet med et asymmetrisk fire-punkts glidende gjennomsnitt. Det asymmetriske gjennomsnittet benytter verdier for for ett år i forkant av den aktuelle måneden og to år i bakkant. For de 12 siste av de 24 verdiene som mistes ved å benytte det 5-punkts glidende gjennomsnittet, dvs. de som gjelder tidspunktene T-17, T-16, ..., T-6, benyttes et ensidig 3-års glidende gjennomsnitt. Det benytter verdier for to år i bakkant av den aktuelle måneden. Merk at vi fremdeles mangler verdier for tidspunktene T-5, T-4, ..., T.

Når vi så sentrerer sesongen, jfr. (4.5), mister vi igjen verdiene for de seks tidspunktene T-11, T-10, ..., T-6. For å skaffe verdier for disse periodene repeteres den sist tilgjengelige bevegeliggjennomsnittsverdien 6 ganger. Dette betyr altså at den siste verdien for det glidende gjennomsnittet i (4.5), for tidspunkt T-12, repeteres for de 6 neste tidspunktene slik at S_t^2 blir ført fram til og med tidspunkt T-6. For å forlenge S_t^2 med 6 tidspunkter, slik at den blir like lang som den opprinnelige serien, repeteres den nærmest tilgjengelige sesongfaktor, dvs. S_t^2-verdien, for de respektive månedene. Dette betyr at for tidspunkt T-5 benyttes S_t^2-verdien for tidspunkt T-5-12 = T-17, for T-4 benyttes S_t^2-verdien for tidspunkt T-4-12 = T-16, osv.

Ved beregning av det Henderson-glidende gjennomsnittet i punkt C i (4.1) mister vi et antall verdier på hver side avhengig av hvilken lengde av gjennomsnittet som velges. For å erstatte disse verdiene benyttes det Wallis (1982) kaller "ekvivalente" asymmetriske glidende gjennomsnitt. Det asymmetriske gjennomsnittet kan nemlig også tolkes som at vi benytter det symmetriske filteret på en serie som er forlenget ved en eller annen framskrivningsmetode, slik at serien består dels av faktiske observasjoner og dels av framskrevne verdier. Når X11 benytter Henderson's glidende gjennomsnitt på verdier nær og ved enden av serien, benyttes en metode som er ekvivalent med å framskrive serien ved å repetere gjennomsnittet av de seks tidligere månedsverdiene. Et tilsvarende "ekvivalent" asymmetrisk glidende gjennomsnitt benyttes også for å erstatte de manglende endeverdiene i det sju-års periodisk glidende gjennomsnittet som brukes ved den endelige estimeringen av sesongkomponenten S_t^3 i (4.9).

Generelt vil det for hvert trinn i X11 som involverer bruk av et symmetrisk bevegelig gjennomsnitt - gitt tilstrekkelig mange observasjoner - korrespondere et asymmetrisk filter som brukes på de siste observasjonene. Disse asymmetriske gjennomsnittene bruker likevel samme antall tidligere verdier som de symmetriske filtrene. Derfor er nettoeffekten av X11-metoden et sett av k+1 lineære filtere (bortsett fra korrigering for ekstremverdier o.l., som innfører ikke-lineariteter) med h.h.v. 0, 1,...,m koeffisienter for løpende og framtidige verdier, og k koeffisienter for tidligere verdier (k = 84 jmf. avsnitt 2.1).

De løpende sesongjusterte tallene vil gi et dårligere bilde av både trendsykel og irregulær komponent enn om en hadde hatt data på begge sider slik at tosidige symmetriske filtere kunne benyttes. Ettersom nye observasjoner kommer til, vil det derfor være naturlig å revidere de sesongjusterte tallene. Dette kan oppfattes som at de (summarisk) framskrevne verdiene blir erstattet med observasjoner. Det er nærmest et paradoks at den offentlige oppmerksomhet vanligvis er fokusert på det først publiserte sesongjusterte anslaget, dvs. det sesongjusterte anslaget for inneværende måned (kvartal) når dette er det mest upålitelige anslaget i den sesongjusterte serien for vedkommende variabel. Tallet vil normalt bli revidert betydelig når flere datapunkter blir tilgjengelig. For at de løpende sesongjusterte tallene skal bli oppfattet med den rette skepsis, kunne man utstyre dem med et anslag på standardavviket til revisjonen fra løpende til endelig sesongreviderte tall. Se Burridge og Wallis (1984), som gir en metode for å beregne slike standardavvik. En annen måte å ta konsekvensen av usikkerheten i de sesongjusterte tallene på er å publisere slike tall først og fremst i figurer. Trolig vil da forholdet mellom den tilfeldige og den systematiske variasjonen komme klart fram.

Det er to viktige problemer knyttet til endepunktene ved sesongjustering. Først må man ta stilling til hvordan en gitt serie eksplisitt eller implisitt bør framskrives. Som vi har sett gjøres dette i X11-programmet ved å la trenden flate ut, hvilket i mange tilfeller kan være en svært utilfredsstillende løsning. I følge Durbin og Kenny (1982) er det mye å vinne på å framskrive serien f.eks. v.h.a. ARIMA-metoder eller eksponensiell glatting, og benytte X11 på den "kunstig" forlengede serien. Dagum (1984) inneholder en beskrivelse av en slik framgangsmåte som nyttes av det canadiske statistiske sentralbyrå. Det andre problemet har å gjøre med hvor ofte man bør reestimere sesongfaktorene etter hvert som nye observasjoner kommer til. Durbin og Kenny fant at dersom man er interessert i å minimere de nødvendige revisjoner av publiserte sesongjusterte tall, bør man reestimere (i praksis kjøre X11) pány hver gang nye observasjoner foreligger. Det vil si at man for hver måned i året kjører X11 på nytt for å få sesongjusterte tall for siste måned. Dette er egentlig ikke overraskende, men i sterk kontrast til Byråets nåværende praksis med for inneværende år å nytte X11-programmets "seasonal factors one year ahead" estimert på grunnlag av observasjoner ut fjoråret. Disse beregnes etter følgende enkle framskrivningsformel: $S_{n+1} = S_n + 1/2\,(S_n - S_{n-1})$ der S_n er vektoren av sesongfaktorer i år n.

4.4 Aggregering.

Behandlingen av ekstremverdier innebærer at filtrene i X11-programmet er ikke-lineære. Det innebærer at en sesongjustert totalsum vanligvis vil være forskjellig fra summen av de sesongjusterte enkeltkomponentene.

Problemet kan formuleres som å justere

(4.12) $x_i \qquad (i = 1,\ldots\ldots,n)$

og

(4.13) $y = \sum_{i=1}^{n} x_i$

slik at

(4.14) $y^s = \sum_{i=1}^{n} x_i^s$

der toppskrift s indikerer sesongjusterte størrelser. En simultan justering av (4.12) og (4.13) med (4.14) som bibetingelse, ville være det mest tilfredsstillende. I praksis er det en uoverkommelig oppgave å gjøre dette for et stort antall serier hver måned. Geweke (1976) viser at gevinsten ved simultan justering er størst for heterogene dataserier og serier med ulik struktur i sesongkomponenten og i andre faktorer og at den er størst for de siste observasjonene i en gitt serie. I homogene serier er imidlertid gevinsten liten. I Geweke (1976, p. 411) sier han likevel, når man ikke kan foreta simultan justering, at "for virtually every conceivable time series and a reasonably inclusive class of potential adjustment procedures..... seasonal adjustment should always precede temporal or sectoral adjustment". I prinsippet bør derfor sesongjusteringen foretas disaggregert. Denne framgangsmåten er dessuten analog til den måten prisdeflatorer vanligvis blir definert på. Når f. eks. nasjonalregnskapstall aggregeres, gjøres dette ved å aggregere både i faste og løpende priser. Deretter blir "nye" aggregerte prisdeflatorer definert ved å dividere det aggregerte verditallet med det aggregerte fastpristallet.

4.5. Noen eksempler

I dette avsnittet vil vi gi noen eksempler på virkningen av bruk av X11-metoden. Først vil vi sesongjustere serier som er konstruert slik at vi har "kontroll" over hvordan sesongen i serien faktisk er. Vi får dermed rendyrket noen egenskaper ved X11. Deretter skal vi kommentere sesongjusteringen av konsumprisindeksen og indeksen for total industriproduksjon; begge tall som er mye omtalt i den offentlige debatt. Hensikten med eksemplene er både å vise tilfeller hvor X11-metoden virker bra og å belyse i praksis noen av de svakhetene som er blitt berørt i de foregående avsnittene. Vi vil også reise problemstillingen hvorvidt en gitt tidsserie overhodet bør sesongjusteres - i alle fall med X11.

I figur 1 er vist en konstruert tidsserie som består av lineær stigende trend som er blitt pålagt en multiplikativ sesongkomponent. Høysesongen er i 1. kvartal og lavsesongen i 2. kvartal. I absolutte termer vil størrelsen på sesongkomponenten stige med nivået på trenden. Vi ser videre av figur 1 at det å bruke multiplikativ X11 resulterer i en meget nøyaktig rensing for sesong idet den faktiske trenden og X11-beregnet trend ligger svært nær hverandre. I figur 2 hvor vi har benyttet additiv X11, ser vi at å nytte "feil" sesongjusteringsmetode kan gi store svakheter i sesongrensingen. Den X11-beregnede trenden svinger rundt den faktiske og gir et ganske feilaktig forløp samtidig som feilene har innslag av sesong. Figuren viser også at X11 bommer mest nærmest enden på serien, fordi der er sesongkomponentens størrelse i stor grad basert på å repetere tidligere verdier, som i absolutte tall er mindre enn verdiene nær enden, jmf. avsnitt 4.3. Uheldig spesifikasjon gir altså større feil i de sesongjusterte tallene, og følgelig større revisjonsusikkerhet.

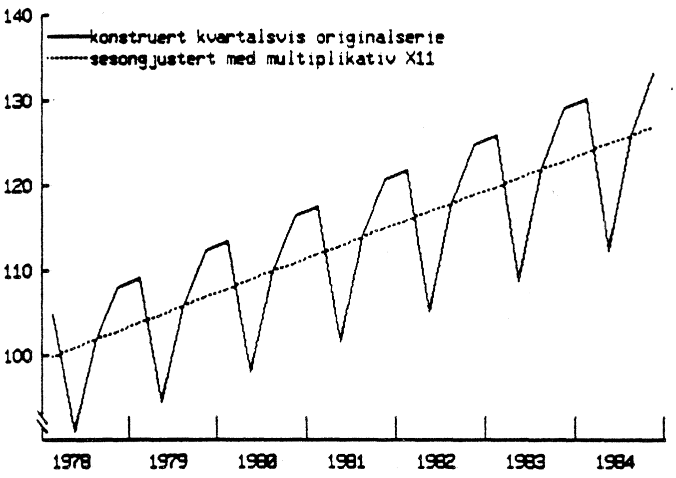
I det første eksemplet på sesongjustering kan vi si at som additivt tillegg var ikke sesongkomponenten konstant, mens X11 forutsetter konstant sesong for å fungere godt. Figur 3 viser et mer dramatisk eksempel på endringer eller brudd i det additive sesongmønsteret til en lineært stigende trend. Slike brudd har bl.a. oppstått ved sammenstilling av "gammelt" og "nytt" kvartalsvis nasjonalregnskap, og vi skal nedenfor komme tilbake til et annet praktisk eksempel. Vi ser av figur 4 at X11 klarer seg dårlig i bruddområdet. Når avstanden til hver side for bruddet i sesongmønsteret øker, treffer metoden trenden bedre og bedre fordi det andre sesongmønsteret får mindre innflytelse i de glidende gjennomsnittene. Man kan si at X11 klarer å lære om det endrede sesongmønsteret fordi gjennomsnittene beveger seg. Det er også verdt å merke seg at før bruddtidspunktet fjernes ikke hele sesongkomponenten, mens etter bruddtidspunktet fjernes mer en sesongeffekten. Dette skyldes at størrelsen på sesongutslagene reduseres betydelig, og at det gjennom vinduet trekkes inn betydelige virkninger fra den delen av serien som har et annet sesongmønster.

X11-programmet gir også, som nevnt i avsnitt 4.3, prognoseverdier for sesongfaktorene ett år fram. For å unngå hyppige revisjoner av de sesongjusterte tallene, blir disse ofte brukt til å beregne sesongjusterte tall for inneværende år. Figur 5 viser trenden og to sesongjusterte serier; den ene beregnet ved bruk av "seasonal factors one year ahead" og den andre v.h.a. løpende oppdatering av sesongfaktorene. Vi ser at den siste ligger nærmest trenden både i 1982 og i 1983. Begge bommer fordi sesongmønsteret endrer seg, men poenget i denne sammenhengen er at vi bommer mest ved å nytte "seasonal factors one year ahead", fordi vi da avskjærer X11 fra å "lære" om det endrede sesongmønsteret.

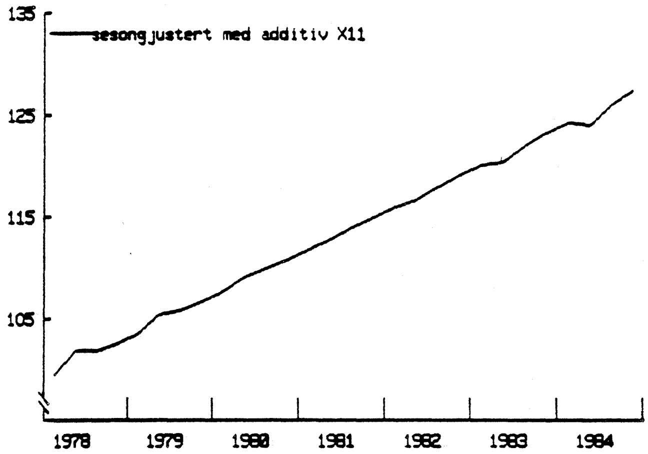
Figur 6 viser vekstraten pr. år i et tremåneders uveiet glidende gjennomsnitt av den sesongjusterte konsumprisindeksen. På grunnlag av kurven som går fram til april 1983 sluttet både Finansdepartementet og Bankforeningen at prisstigningstakten falt dramatisk ved inngangen til 1983. Vi vil imidlertid hevde at fallet slik det framtrer, i svært stor grad skyldes sesongjusteringsmetoden og endring i sesongmønsteret. Ved årsskiftene 80/81 og 81/82 var det nemlig meget store økninger i offentlige avgifter og reduksjoner i subsidier samt økt pris på offentlige gebyrvarer. X11-metoden har dermed estimert store sesongkoeffisienter - prishøysesong - i januar, og disse blir brukt til å nedjustere originalserien for å komme fram til den sesongjusterte verdien. Når det så ved årsskiftet 82/83 bare gjøres små avgiftsøkninger, vil de høye sesongfaktorene - tillært av X11 da det var store avgiftspåleggelser - føre til en stor nedjustering for å komme fram til den sesongjusterte prisindeksen. Sannsynligvis har det derfor liten mening å sesongjustere konsumprisindeksen, i alle fall for slike hendelser som det blir fattet "nye" beslutninger om hvert år. Spesielt vil det for Finansdepartementet ha liten mening å sesongjustere sine egne handlingsparametre v.h.a. X11-metoden. I figur 7 har vi vist konsumprisindeksens delindeks for klær og skotøy, og denne må sies å ha et meget klart sesongmønster. Det er sesongsalg i januar og august. En slik prisindeks vil det antakelig ha en klarere mening å sesongjustere enn totalindeksen. Men selv her er det ikke opplagt at man bør vente en konstant sesong; vi skulle snarere vente at sesongutslagene vil avhenge av etterspørselspresset i økonomien.

Figur 8 viser den ujusterte indeksen for total industriproduksjon. Den har et svært markert sesongmønster med en meget klar sesongbunn i juli da det er felles-ferie. Siden produksjonsindeksen er en månedsindeks, vil også andre og mindre dramatiske variasjoner i antall virkedager i en måned gi merkbare utslag i indeksens nivå. Spesielt viktig i denne sammenheng er hvorvidt påsken faller i mars eller april eller fordeler seg på begge månedene. Statistisk Sentralbyrå foretar derfor en prekorrigering av sériene før X11 benyttes til sesongjustering. Prekorrigeringen går ut på å fordele påskens normale produksjonstap etter faste koeffisienter på mars eller april avhengig av i hvilken måned den faller. Det er imidlertid vanskelig å reflektere i koeffisientene de mange måtene påsken kan fordele seg på mars og april når den kommer i månedsskiftet. Av figur 9 ser vi at nivået til mars- og apriltallene for produksjonsindeksen både i 1980 til 1982 og i 1984 ser rimelige ut til tross for at påsken fordelte seg ulikt på mars og april. Mai og juni 1981 og 1984 ser derimot meget spesielle ut. Årsaken er at pinsen falt i juni i disse årene i motsetning til mai i de andre og at det ikke er foretatt noen prekorrigering for å ta hensyn til dette. X11 vil her utfra tidligere "erfaring" ha estimert mai som en lav-sesong og sesongjustere opp maitallet. Når pinsen faller i juni vil en mai-måned med unormalt mange virkedager bli justert opp og derav den høye sesongjusterte verdien i mai. Juni på sin side vil ha unormalt få virkedager og et tilsvarende

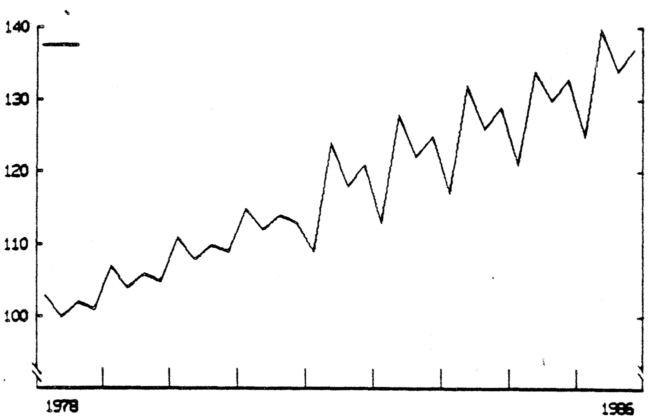
Figur 1. Konstruert kvartalsserie med multiplikativt sesongmønster



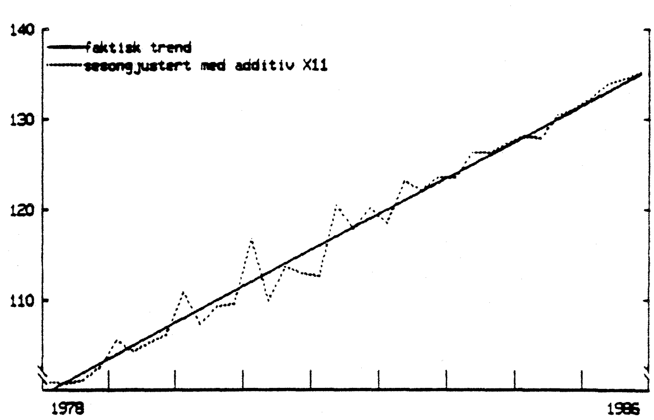
Figur 2. Konstruert kvartalsserie med multiplikativt sesongmønster



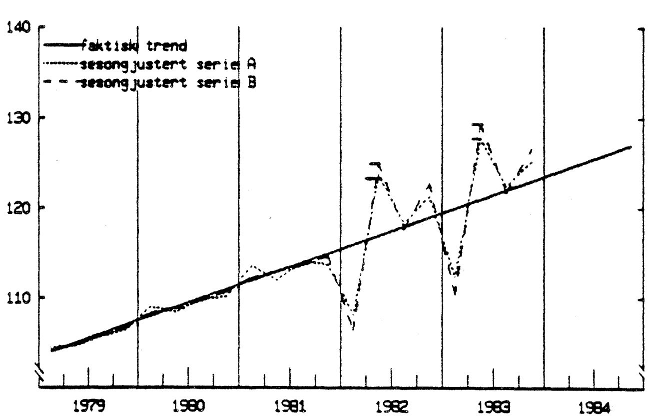
Figur 3. Konstruert kvartalsserie med brudd i sesongmønsteret i 1.kvartal 1982



Figur 4. Sesongjustert tidsserie når originalserien har brudd i sesongmønsteret

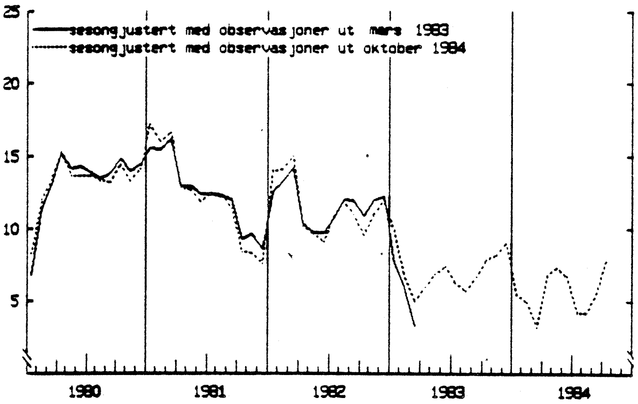


Figur 5. Bruk av "seasonal factors one year ahead" ved løpende sesongjustering

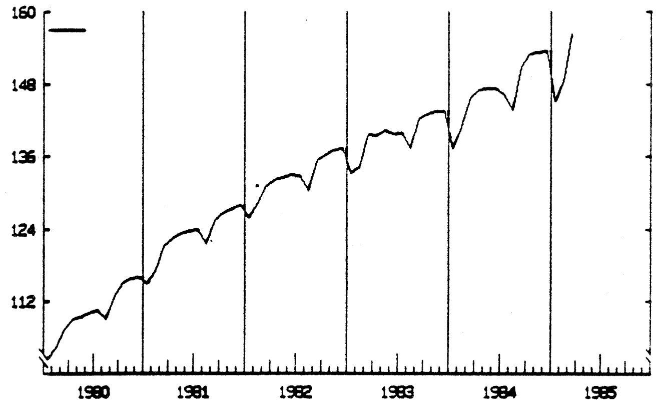


A: serie justert med additiv X11 for 1978 1 - 1981 4.
B: additivt justert X11-serie for 1978 1 - 1981 - 4.
I 1982 er serien sesongjustert ved å bruke X11's "seasonal factors".

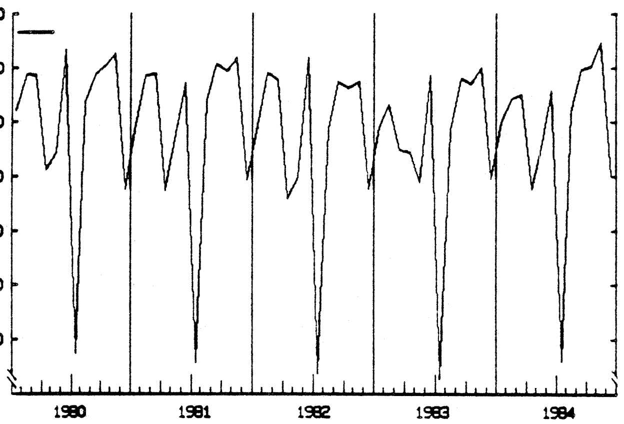
Figur 6. Tremåneders uveiet glidende gjennomsnitt av sesongjustert total konsumprisindeks. Årlig endringsrate



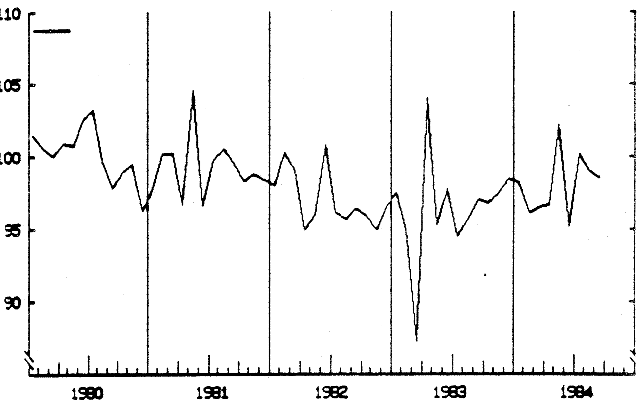
Figur 7. Konsumprisindeksens delindeks for tekstil- og bekledningsvarer. 1979=100.



Figur 8. Indeks for total industriproduksjon. 1980=100



Figur 9. Sesongjustert indeks for total industriproduksjon



lavt tall[2]. Vi ser dessuten av figur 9 at månedene mars og april 1983 er høyst unormale. Dette skyldes, så vidt vi har kunnet fastslå, at påskens faktiske fordeling på de to månedene samsvarer dårlig med de koeffisientene Byrået bruker i sin nåværende prekorrigeringsrutine.

5. KONKLUSJON OG NOEN TILRÅDINGER.

Vi vil nå oppsummere diskusjonen og samtidig komme med noen tilrådinger for endringer av Byråets praksis for sesongjustering. Sett på bakgrunn av det meget store antall tidsserier Byrået sesongjusterer, er det lite trolig at mer tilfredsstillende modellbaserte justeringsmetoder er et realistisk alternativ for alle seriene. De følgende vurderinger tar derfor utgangspunkt i at Byrået skal nytte X11-metoden i en eller annen form.

1. Byrået må fortsette å gjøre ujusterte tall tilgjengelig da disse fyller et viktig behov bl.a. ved økonometrisk samvariasjonsanalyse. Etter vår mening er det også mulig at Byrået i større grad enn til nå bør gjøre det klart at sesongjusterte tall er en analytisk bearbeiding av primærstatistikk, og langt fra kan ha samme status som statistisk primærmateriale.

2. La oss slå fast at det symmetriske bevegelige gjennomsnittet som danner kjernen i X11, har gode egenskaper m.h.p. å rense ulike typer tidsserier for sesongsvingninger av ulik periodisitet. Vi vil her sitere Hylleberg (1984) s. 134, som etter en grundig evaluering av ulike justeringsmetoder sier at "In concluding this part of the evaluation of the application of officially adjusted data, it is probably fair to say that .....

---

2) Mangelen på prekorrigering var for øvrig opphav til en interessant avisdiskusjon om utviklingen i industriproduksjonen våren 1984. Industriminister Syse sa ifølge Aftenposten 19. juli med bakgrunn bl.a. i det høye maitallet for industriproduksjonen (sesongjustert) at Industriforbundets pessimistiske prognose basert bare på januar - apriltall var forfeilet. Det var sterk vekst i industriproduksjonen. Når det lave juni-tallet forelå en måned senere, tok selvfølgelig Industriforbundet sitt monn igjen. Merk at PCD er 6 måneder!

the X11 method performs second to none when the comparisons are made based on the characteristics of the single adjusted or unadjusted series". Metoden er også relativt robust overfor innflytelsen til ekstreme datapunkter. X11-metoden er dessuten fleksibel i den forstand at det er mange opsjoner til standardvarianten. En viktig konklusjon av vårt arbeid er at Byrået bør utnytte disse opsjonene mer enn tilfellet har vært til nå.

3. Byråets praksis er idag å benytte bare den additive varianten av X11. Vår vurdering er at man bør velge multiplikativ eller additiv sesongjusteringsmetode etter en nærmere vurdering av egenskapene til hver enkelt dataserie eller gruppe av serier. Vårt inntrykk er at da vil ofte den multiplikative metoden være å foretrekke framfor den additive.

4. X11-metoden utfører ikke-lineære operasjoner. Dette impliserer at man får ulike resultater om man først aggregerer - i tid eller over sektor - og deretter sesongjusterer eller motsatt. Vi mener det korrekte i dette tilfellet er, analogt med deflatering, først å sesongjustere og deretter aggregere seg fram til det sesongjusterte aggregatet. Byråets praksis idag er å sesongjustere aggregerte og disaggregerte tall hver for seg. Det sesongjusterte aggregatet vil da bare ved en tilfeldighet være lik summen av tallene som ble sesongjustert på disaggregert nivå. Dersom det på et disaggregert nivå er problemer med en stor irregulær komponent, kan dette forsterke ikke-linearitetene i X11-metoden. Dette kan imidlertid kompenseres ved å øke de kritiske verdiene i justeringsrutinen for ekstremverdier. På denne måten vil mer av tidsserien slippe gjennom og tilfeldighetene i de disaggregerte seriene kan jevne seg ut ved aggregeringen.

5. Ved estimeringen av de løpende sesongjusterte tallene benytter X11-metoden seg av et glidende gjennomsnitt som impliserer bestemte forutsetninger om dataseriens videre forløp, bl.a. at trenden "flater ut". Dette kan i mange tilfeller være urimelig. Vi har her ingen tilråding, men det bør for spesielt viktige tidsserier (f. eks. arbeidsledighetstallet) eksperimenteres noe med måter å "forlenge" den faktiske serien på før X11

benyttes; f. eks. ARIMA-modeller. Dette kan gi en forbedring av kvaliteten på de løpende publiserte sesongjusterte tallene.

6. Ved estimering av de vektene som benyttes ved løpende sesongjustering er det ineffektiv bruk av data å estimere faste sesongfaktorer for det inneværende år v.h.a. bare den informasjon som var til stede ved utgangen av fjoråret, slik som Byråets praksis er i dag. Ved utgangen av f. eks. 1983 lagde man v.h.a. X11 sesongfaktorprognoser for 1984 og disse faktorene ble benyttet uten revisjon etter hvert som nye data kom til gjennom 1984. Alt i alt blir den nødvendige revisjonen av de sesongjusterte tallene størst på denne måten. Vi vil tilrå at i alle fall for kvartalsseriene, og for månedsseriene hvert kvartal, bør sesongjusteringen utføres ved at X11 benyttes på hele den tilgjengelige serien. Problemet forsterkes ved at X11-metodens prognoser for sesongfaktorene ikke kan antas å være av særlig høy kvalitet.

7. For å understreke den klare usikkerheten som er forbundet med de sesongjusterte tallene, er det mulig man først og fremst burde basere presentasjoner av dem på figurer. Det vil da gå klart fram fra figurene at også de sesongjusterte tallene er erratiske, og de vil derfor tillegges mindre vekt ved kortvarige endringer.

8. For en del formål vil trolig et større innslag av prekorrigering av serien være gunstig; f.eks. i forbindelse med bevegelige helligdager.

REFERANSER

Biørn, K.E. og Jensen, M. (1983): Varige konsumgoder i et komplett system av etterspørselsrelasjoner. Rapporter nr. 83/16 fra Statistisk Sentralbyrå.

Biørn, K.E., Jensen, M. og Reymert, M. (1985): KVARTS - a Quarterly Model of the Norwegian Economy. Under utgivelse i serien Discussion Papers fra Statistisk Sentralbyrå.

Brenna, S (1961): Metoder for sesongutjevning I. Internt notat. Statistisk Sentralbyrå.

Brenna, S. (1966): Sesongutjevning av kvartalsserier. Internt notat. Statistisk Sentralbyrå.

Burridge, P. og Wallis, K.F. (1984): Calculating the Variance of Seasonally Adjusted Series. Paper presented at the Econometric Society Meeting in Madrid, september 1984.

Dagum, E.B. (1980): The X11-ARIMA seasonal adjustment method. Statistics Canada.

Durbin, J. og Kenny, P.B. (1982): Local trend estimation and seasonal adjustment of economic and social time series. Journal of the Royal Statistical Society A, 1982.

Geweke, J. (1979): The temporal and sectoral aggregation of seasonally adjusted time series,i Seasonal Analysis of Economic Time Series. Economic Research Report, ER-1, U.S. Department of Commerce, Bureau of the Census.

Hylleberg, S. (1984): Seasonality in Regression. Økonomisk Institutt. Aarhus Universitet.

Kendall M.C.(1973): Time Series, Griffin, London 1973.

Klein, L.R. (1976): Comments on "An overview of the objectives and framework of seasonal adjustmet" by Shirley Kallek, i Seasonal Analysis of Economic Time Series. Economic Research Report, ER-1, U.S. Department of Commerce, Bureau of the Census.

Pierce, D. (1980): A Survey of Recent Developments in Seasonal Adjustment. The American Statistician, August 1980, Vol. 34, No. 3.

Pierce, D. (1983): Seasonal adjustment of the monetary aggregates: Summary of the Federal Reserve's committee report. Journal of Bus. Econ. Statist., 1, 1983, 37-42.

Shiskin, J., Young, A.H. og Musgrave, J.C. (1967): The X11 variant of the Census method II seasonal adjustment program. Technical paper no. 15. Bureau of Census, February 1967.

Wallis, K.F. (1980): Seasonal adjustment and revision of current data: Linear filters for the X11 method. Journal of the Royal Statistical Society A, 145, 1982.