2000/9 Rapporter Reports

Håvard Hungnes

Beregning av årsrelasjoner på grunnlag av økonometriske kvartalsrelasjoner

Statistisk sentralbyrå • Statistics Norway
Oslo–Kongsvinger

Rapporter

I denne serien publiseres statistiske analyser, metode- og modellbeskrivelser fra de enkelte forsknings- og statistikkområder. Også resultater av ulike enkeltundersøkelser publiseres her, oftest med utfyllende kommentarer og analyser.

Reports

This series contains statistical analyses and method and model descriptions from the different research and statistics areas. Results of various single surveys are also published here, usually with supplementary comments and analyses.

ISBN 82-537-4799-3
ISSN 0806-2056

Emnegruppe
09.90 Metoder, modeller, dokumentasjon

Design: Enzo Finger Design
Trykk: Statistisk sentralbyrå

Standardtegn i tabeller	Symbols in tables	Symbol
Tall kan ikke forekomme	Category not applicable	.
Oppgave mangler	Data not available	..
Oppgave mangler foreløpig	Data not yet available	...
Tall kan ikke offentliggjøres	Not for publication	:
Null	Nil	-
Mindre enn 0,5 av den brukte enheten	Less than 0.5 of unit employed	0
Mindre enn 0,05 av den brukte enheten	Less than 0.05 of unit employed	0,0
Foreløpige tall	Provisional or preliminary figure	*
Brudd i den loddrette serien	Break in the homogeneity of a vertical series	—
Brudd i den vannrette serien	Break in the homogeneity of a horizontal series	|
Rettet siden forrige utgave	Revised since the previous issue	r

Sammendrag

Håvard Hungnes

Beregning av årsrelasjoner på grunnlag av økonometriske kvartalsrelasjoner

Rapporter 2000/9 Statistisk sentralbyrå 2000

Et alternativ til å estimere sammenhenger mellom årlige observasjoner av økonometriske størrelser er å beregne disse fra estimerte kvartalsrelasjoner. Ved å bruke en slik fremgangsmåte utnytter man datasetet bedre, siden man benytter informasjonen som ligger i kvartalsdatene. Hvis relasjonene allerede er estimert på kvartal, er det en enda større fordel å kunne beregne årsrelasjonen fra kvartalsrelasjonen, siden estimeringsarbeidet allerede er gjort.

I rapporten viser vi hvordan kvartalsrelasjoner på logaritmisk form kan omregnes til årsrelasjoner. For å illustrere hvor effektiv metoden er presenteres en Monte Carlo simulering.

Innhold

1 Innledning[1]

Mange forfattere har undersøkt effekter av tidsaggregering av forskjellige typer ARMA-modeller, herunder Telser (1967), Amemiya og Wu (1972), Kendall (1973), Brewer (1973), Tiao og Wei (1976), Wei (1981), Weiss (1984), Stram og Wei (1986) og Wei (1989). Telser (1967) og Amemiya og Wu (1972) ser på problemet ved rene autoregressive AR(p) modeller. Resultatet de rapporterer er at rene autoregressive AR(p) modeller endres gjennom tidsaggregering til en ARMA(p,q)-modell. Brewer (1973) diskutere effekten av tidsaggregering av ARMA(p,q)-modeller og viser at ARMA(p,q)-modellene vil etter tids-aggregering blir også omformet til en ARMA(p,q*)-prosess, men med flere lag i MA-prosessen (q*>q).

Brewer (1973), Tiao og Wei (1976), Wei (1981) og Weiss (1984) har videre sett på ARMAX-modeller, dvs. ARMA-modeller med eksogene variable. Tiao og Wei (1976) foreslår en endring i metoden til Brewer (1973) som utnytter informasjonen om den eksogene variabelen bedre, men begrenset seg til en relasjon der tilbakedaterte verdier av den endogene variabelen ikke inngår (dvs. relasjoner uten AR-polynom). Wei (1981) og Weiss (1984) ser på ARMAX-modeller med metoden foreslått i Tiao og Wei (1976).

Starm og Wei (1986) og Wei (1989) viser at det kan være skjulte periodisiteter i AR-polynomet, og at antall lag i den tidsaggregerte relasjonen i så fall kan reduseres. Selv om Stram og Wei (1986) og Wei (1989) bare diskuterer tilfellet med skjulte periodisiteter innenfor relasjoner der den endogene variabelen er en beholdningsvariabel, kan det tilsvarende utledes for tilfellet der den endogene variabelen er en strømningsvariabel.

Med bakgrunn i den overfor nevnte teorien, viser vi i denne rapporten hvordan man kan omregne kvartalsrelasjoner om til årsrelasjoner. (Metoden kan enkelt modifiseres til å regne om fra f.eks. månedsrelasjoner til kvartalsrelasjoner.) Rapporten viser denne omregningen i detalj. I denne forbindelse behandles også problemer som sesongdummyer, trend og andre dummyer. I motsetning til i de ovenfornevnte arbeidene benyttes logaritmiske funksjoner, siden dette er mest vanlig for økonomiske relasjoner. Dette medfører at en tilnærming må gjøres hvis en eller flere av variablene i relasjonen er strømningsvariable.

I rapporten skiller vi mellom strømnings- og beholdningsvariabler. Årsverdien til en strømningsvariabel er lik summen av kvartalsverdiene over kalenderårets fire kvartaler; for eksempel er produksjonen i løpet av et år lik summen av produksjonen i årets fire kvartaler. Årsverdien for en beholdningsvariabel settes vanligvis lik beholdningsvariabelen ved utgangen av fjerde kvartal. Hvis årsvariabelen er lik gjennomsnittet av kvartalsvariabelen over kalenderåret, vil den her bli behandlet på samme måte som en strømningsvariabel.

I kapittel 2 ser vi først på tilfellet der både den endogene og den eksogene variabelen er beholdningsvariable (og årsverdien av beholdningsvariabelen er

[1]Forfatteren vil takke Knut Moum, Terje Skjerpen og Anders Rygh Swensen for kommentarer på tidligere utkast. I tillegg til å komme med mange nyttige kommentarer har Skjerpen fungert som en diskusjonspartner, og fortjener derfor en spesiell takk.

lik fjerde-kvartalsverdien). Kapittel 3 omhandler hvordan sesongsvingninger, trender og dummyer i kvartalsrelasjonen påvirker årsrelasjonen. I kapittel 4 tar vi for oss tilfellet der den eksogene variabelen er en strømningsvariabel, mens tilfellet der den endogene variabelen er en strømningsvariabel behandles i det femte kapittelet. I kapittel 6 viser vi hvordan konstantleddet og restleddsprosessen i årsrelasjonen kan bestemmes. Styrken ved metoden testes ved hjelp av en Monte Carlo simulering i kapittel 7. Kapittel 8 konkluderer. Flere av kapitlene avsluttes med et eksempel der metoden som beskrives i kapittelet illustreres.

Vi ser her på omregning av kvartalsrelasjoner på "autoregressiv distribuert lag"-form (ADL-form) til årsrelasjoner på samme form. I økonomiske modeller er relasjonene ofte presentert på feiljusteringsform. I vedlegg A viser vi derfor hvordan en relasjon kan beregnes fra en modell på feiljusteringsform til en ADL-form, og omvendt.

Vi tar utgangspunkt i relasjon (1), som antas estimert på kvartalsdata. For enkelthets skyld ser vi på tilfellet med bare en eksogen variabel. Det er uproblematisk å innføre flere eksogene variable.

$$A_p(L) \ln y_t = \alpha_0 + B_q(L) \ln x_t + \varepsilon_t, \;\; t = 1, 2, 3, ..., \tau \tag{1}$$

Her er y_t og x_t henholdsvis verdien på den endogene og eksogene variabelen i kvartal t, $A_p(L) = (1 - \alpha_1 L - \alpha_2 L^2 - ... - \alpha_p L^p)$ og $B_q(L) = (\beta_0 + \beta_1 L + \beta_2 L^2 + ... + \beta_q L^q)$ er lag-polynomer med henholdsvis p og q lag, der lagoperatoren L opererer på kvartal som $L^s(\ln y_t) = \ln y_{t-s}$. Restleddet ε_t antas å ha forventning null ($E(\varepsilon_t) = 0$, $\forall t$), konstant varians ($Var(\varepsilon_t) = \sigma^2$, $\forall t$) og å være uten autokorrelasjon ($E(\varepsilon_t \varepsilon_s) = 0$, $\forall t, s$ der $t \neq s$).

2 Beholdningsvariable

I dette kapittelet antar vi at både den eksogene og den endogene variabelen er beholdningsvariable, og at verdien til årsvariabelen er lik verdien til kvartalsvariabelen i 4. kvartal. Problemet med å benytte relasjon (1) i en årsmodell er at den inneholder 'uobserverbare' verdier. Med uobserverbare verdier mener vi at relasjonen inneholder verdier av beholdningsvariablene for alle kvartaler, mens bare verdien av beholdningen i 4. kvartal benyttes (er 'observerbar') i årsmo-dellen. Brewer (1973) foreslår at hvert ledd i (1) multipliseres med et lagpolynom $C(L) = (c_0 + c_1 L + ...)$ som er slik at produktet av $C(L)A(L)$ gir et lagpolynom med nuller for de uobserverbare verdier, dvs. $C(L)A(L) = A^*(L)$, der $A^*(L) = (1 - \alpha_4^* L^4 - \alpha_8^* L^8 - ...)$. Brewer (1973) viser at C- og A^*-polynomene må ha henholdsvis $3p$ og $4p$ lag. Likningssystemet $C_{3p}(L)A_p(L) = A_{4p}^*(L)$ inneholder dermed $4p + 1$ relasjoner til å løse for like mange ukjente parametere ($3p + 1$ ukjente parametre i C-polynomet og p ukjente parametre i A^*-polynomet).

Ifølge Weiss (1984) har Wei (1981) utledet en formel for beregning av C-polynomet, som bygger på et tilsvarende polynom Amemiya og Wu (1972) utviklet for beholdningsvariable. La $\phi_1, ..., \phi_p$ være de inverse av røttene i $A_p(z) = 0$. Det forutsettes at røttene ligger utenfor enhetssirkelen, slik at $|\phi_i| < 1, i = 1, 2, ..., p$.[2] Lag-polynomet er da gitt i (2).

$$C_{3p}(L) = \prod_{j=1}^{p} \frac{1 - \phi_j^4 L^4}{1 - \phi_j L} = \prod_{j=1}^{p} \left(1 + \phi_j L + \phi_j^2 L^2 + \phi_j^3 L^3\right) \tag{2}$$

De inverse av røttene i $A^*(L)$ (beregnet på år) er $\phi_1^4, ..., \phi_p^4$, dvs. lik de inverse av røttene i $A(L)$ opphøyd i 4. Vi sier at relasjon (1) inneholder skjulte periodisiteter hvis det finnes et par av røtter der $\phi_i \neq \phi_j$ mens $\phi_i^4 = \phi_j^4$. Stram og Wei (1986) viser at hvis A-polynomet i relasjon (1) inneholder skjulte periodisiteter, vil det ikke være nødvendig å multiplisere med så mange lag som $3p$ for at A^*-polynomet skal bestå av bare 'observerbare' størrelser. I et teorem viser Stram og Wei (1986) følgende:

Teorem 1 *(Stram og Wei, 1986, teorem 4.1, side 288).*[3] *La $\phi_1^{-1}, \phi_2^{-1}, ..., \phi_{p'}^{-1}$ være de p' forskjellige (distinkte) røttene i $A(L)$, hver med multiplisitet s_i. La b være antall forskjellige verdier av $\phi_1^4, \phi_2^4, ..., \phi_{p'}^4$. Partisjoner videre numrene $s_1, s_2, ..., s_{p'}$ i b forskjellige sett A_i slik at $(s_i, s_j) \in A_i$ hvis og bare hvis $\phi_i^4 = \phi_j^4$. Da er det maksimale antall nødvendige lag i C-polynomet gitt ved $3p^*$, der $p^* = \left[\sum_{i=1}^{b} \max A_i\right]$ hvor $\max A_i$ er det største elementet i A_i.*

Bevis. Se Stram og Wei (1986, side 288-289). ■

Det minste antall lag i C-polynomet er dermed $4p^* - p$ (som er lik $3p$ når $p^* = p$). Dette polynomet finner vi ved å ordne røttene slik at de p^* første røttene ikke inneholder skjulte periodisiteter i $A(L)$. C-polynomet beregnes da

[2] Generelt kan man også ha komplekse røtter. Kravet blir da at modulus til røttene må være større enn 1.

[3] Teoremet er skrevet litt om.

ved

$$C_{4p^*-p}(L) = \frac{\prod_{j=1}^{p^*} \left(1 - \phi_j^4 L^4\right)}{\prod_{j=1}^{p} \left(1 - \phi_j L\right)},$$

og A^*-polynomet vil da få $4p^*$ kvartalslag $(p + (4p^* - p))$.[4]

I det videre vil vi anta at det ikke er skjulte periodisiteter i A-polynomet i relasjon (1), slik at C-polynomet har $3p$ lag. (Hvis det er skjulte periodisiteter i (1), vil de $4(p - p^*)$ siste lag-ene i C-polynomet være lik null, og den videre utregningen kan fortsatt benyttes.)

Etter at vi har multiplisert relasjon (1) med C-polynomet står vi igjen med relasjonen

$$\begin{aligned} A_{4p}^*(L) \ln y_T &= C_{3p}(1)\alpha_0 + B_{3p+q}^*(L) \ln x_T + C_{3p}(L)\varepsilon_T \\ T &= 4, 8, 12, ..., \tau, \end{aligned} \tag{3}$$

der $B^*(L) = C(L)B(L)$ og T bare teller over observerbare kvartaler. Siden det er $3p$ lag i C-polynomet og q lag i B-polynomet vil B^*-polynomet ha $3p + q$ lag.

Multipliseringen av C-polynomet innebærer ikke at vi har pålagt null-restriksjoner for parametrene i B^*-polynomet som er knyttet til uobserverbare verdier av variabelen. Vi må derfor gjøre tilpasninger i B^*-polynomet for at alle verdiene av den eksogene variabelen som inngår i (3) skal være observerbare. For å kunne gjøre en slik tilpasning må vi kjenne den marginale prosessen til den eksogene variabelen. Det enkleste vil være å anta at den følger en enkel ikke-stasjonær prosess som i (4), hvor logaritmen av den eksogene variabelen følger en tilfeldig gang med drift. Denne prosessen innebærer at x_t er sterk eksogen, dvs. at den ikke avhenger av utviklingen i den endogene variabelen.

$$\ln x_t = \delta + \ln x_{t-1} + u_t, \; t = 1, 2, 3, .., \tau, \tag{4}$$

der restleddet u_t er normalfordelt og $E(u_t) = 0$, $Var(u_t) = \sigma_x^2$ og $E(u_t u_s) = 0$ for $\forall t, s$ der $s \neq t$. I tillegg er det per konstruksjon ingen korrelasjon mellom restleddet i (1) og restleddet i (4), dvs $E(\varepsilon_t u_t) = 0$, $\forall t$.

Relasjon (4) kan ved gjentatt innsetting omskrives til

$$\begin{aligned} \ln x_{T-j} &= \ln x_{T-4} + (4 - j) \cdot \delta + \sum_{i=j}^{3} u_{T-i} \\ T &= 4, 8, 12, ...\tau, \quad j = 0, 1, 2, 3 \end{aligned} \tag{5}$$

[4] Et eksempel på en A-polynom som inneholder skjult periodisitet er $A(L) = (1 - \alpha_2 L^2)$. De inverse av røttene til dette polynomet er $\pm\sqrt{\alpha_2}$, men polynomet har skjult periodisitet siden $(\sqrt{\alpha_2})^4 = (-\sqrt{\alpha_2})^4$. Vi kan dermed beregne C-polynomet med

$$C(L) = \frac{1 - (\sqrt{\alpha_2})^4 L^4}{(1 - \sqrt{\alpha_2} L)(1 + \sqrt{\alpha_2} L)} = \frac{1 - (\alpha_2)^2 L^4}{1 - \alpha_2 L^2}.$$

Multipliserer vi ut for å finne lagpolynomet i årsrelasjonen får vi $A^*(L) = C(L)A(L) = \left(1 - (\alpha_2)^2 L^4\right)$.

Brewer (1973) foreslår å benytte relasjon (5) til erstatning for 'uobserverte' verdier av $\ln x_{T-j}$. Men som påpekt av Tiao og Wei (1976) utnytter vi ikke all informasjon ved å benytte denne omskrivingen. Selv om vi ikke kjenner hvert enkelt restledd i relasjon (4), har vi likevel kunnskap om dem. Det ser vi hvis vi løser for $j = 0$ i (5), og omskriver til

$$u_T + u_{T-1} + u_{T-2} + u_{T-3} = \ln x_T - \ln x_{T-4} - 4\delta. \tag{6}$$

Siden hvert element på venstresiden i denne relasjonen er kjent, kjenner vi også summen av restleddene på høyresiden. Dette kan vi utnytte til å uttrykke en betinget forventning til restleddene. Siden $u_T, u_{T-1}, u_{T-2}, u_{T-3}$ er uavhengige normalfordelte, er

$$\begin{pmatrix} u_{T-j} \\ U_T \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & 4\sigma^2 \end{pmatrix}\right), \tag{7}$$

der $U_T = u_T + u_{T-1} + u_{T-2} + u_{T-3}$. Den betingede forventningen er da gitt ved

$$\begin{aligned} E\left[u_{T-j} \left| U_T \right.\right] &= E\left[u_{T-j}\right] + \frac{\sigma^2}{4\sigma^2}\left(U_T - E\left[U_T\right]\right) \\ &= \frac{1}{4} U_T \text{ eller} \\ E\left[u_{T-j} \left| \ln x_T, \ln x_{T-4} \right.\right] &= \frac{1}{4}\left(\ln x_T - \ln x_{T-4} - 4\delta\right), \end{aligned} \tag{8}$$

når vi benytter relasjon (6) og (7). Dette leder oss til følgende teorem.

Teorem 2: *Hvis den eksogene variabelen følger en marginal prosess som i (4), er den betingede forventningen til den eksogene variabelen gitt ved et veid gjennomsnitt av den siste observerte før og den første observerte etter, når vi betinger med hensyn på de observerbare verdiene av den eksogene variabelen. Vektene er omvendt proporsjonal med avstanden mellom den uobserverbare og de observerbare verdien av den eksogene variabelen.*

$$\ln x_{T-j} = \frac{4-j}{4} \ln x_T + \frac{j}{4} \ln x_{T-4} - \left(\frac{4-j}{4} \sum_{i=0}^{j-1} u_{T-i} - \frac{j}{4} \sum_{i=j}^{3} u_{T-i} \right) \tag{9}$$

Den betingede forventningen til dette restleddet er null, slik at

$$\begin{aligned} E(\ln x_{T-j} \left| \ln x_T, \ln x_{T-4} \right.) &= \frac{4-j}{4} \ln x_T + \frac{j}{4} \ln x_{T-4} \\ T &= 4, 8, 12, ..., \tau \quad j = 0, 1, 2, 3 \end{aligned} \tag{10}$$

Bevis. Vi tar den betingede forventningen av relasjon (5), og benytter relasjon (8).

$$
\begin{aligned}
& E\left[\ln x_{T-j}\left|\ln x_T, \ln x_{T-4}\right.\right] \\
= \;& \ln x_{T-4}+(4-j)\delta+E\left[\sum_{i=j}^{3} u_{T-i}\left|\ln x_T, \ln x_{T-4}\right.\right] \\
= \;& \ln x_{T-4}+(4-j)\delta+\frac{4-j}{4}\left(\ln x_T-\ln x_{T-4}-4\delta\right) \qquad (11) \\
= \;& \frac{4-j}{4}\ln x_T+\frac{j}{4}\ln x_{T-j},
\end{aligned}
$$

Vi har dermed vist hvordan vi kommer frem til relasjon (10). For å vise relasjon (9) benytter vi at ved gjentatt innsetting kan relasjon (4) også skrives som

$$\ln x_{T-j}=\ln x_T-j\cdot\delta-\sum_{i=0}^{j-1} u_{T-i}. \qquad (12)$$

Beregner vi et veid gjennomsnitt av relasjon (5) og (12), der vi benytter henholdsvis $j/4$ og $1-j/4$ som vekter, får vi relasjon (9). ■

Vi har dermed bevist både at den strukturelle delen av relasjonene (9) og (10) inneholder all tilgjengelig informasjon om restleddsprosessen og at den betingede forventningen til det sammensatte restleddet i (9) er null.

Relasjon (10) vil bare uttrykke den betingede forventningen til 'uobserverbare' verdier av den eksogene variabelen hvis den eksogene variabelen følger en prosess som i relasjon (4). Hvis den eksogene variabelen følger en stasjonær prosess vil koeffisienten foran $\ln x_{t-1}$ i relasjon (4) være mindre enn 1. Uttrykket for den betingede forventningen til den eksogene variabelen vil likevel ikke avvike vesentlig fra uttrykket i (10) så lenge koeffisienten er nær 1. Dette er en rimelig antagelse, siden (den inverse av) enhetsroten ikke er langt fra 1 i dynamiske relasjoner på kvartalsdata.

Vi kan erstatte de uobserverbare verdiene av den eksogene variabelen i (3) med uttrykket i relasjon (9). Relasjonen vi ender opp med er en årsrelasjon med observerbare verdier av både den endogene og den eksogene variablene.

Generelt kan vi beregne lag-polynomet til den eksogene variabelen på følgende måte: La $B_{4n}^{*\prime}=(\beta_0^*,\beta_1^*,\beta_2^*,...,\beta_{4n-1}^*)$ være en vektor som består av alle koeffisientene i B^*-polynomet, der $n=1+\left\{\frac{3p+q}{4}\right\}^+$ hvor $\{\}^+$ betyr at innholdet i klammeparentesen avrundes oppover til nærmeste heltall. De siste $(4n-1)-(3p+q)$ elementene i B^*-vektoren settes lik 0.

Videre definerer vi matrisen

$$
H_{4n\times 4n} = \frac{1}{4}\left(\begin{array}{ccccccc}
4 & 0 & 0 & & & & \\
3 & 1 & 0 & & & & \\
2 & 2 & 0 & & & & \\
1 & 3 & 0 & & & & \\
0 & 4 & 0 & \cdots & & 0 & \\
0 & 3 & 1 & & & & \\
0 & 2 & 2 & & & & \\
0 & 1 & 3 & & & & \\
0 & 0 & 4 & & & & \\
 & \vdots & & \ddots & & \vdots & \\
 & & & & 4 & 0 & 0 \\
 & & & & 3 & 1 & 0 \\
 & & & & 2 & 2 & 0 \\
 & & & & 1 & 3 & 0 \\
 & & & & 0 & 4 & 0 \\
 & 0 & & \cdots & 0 & 3 & 1 \\
 & & & & 0 & 2 & 2 \\
 & & & & 0 & 1 & 3 \\
 & & & & 0 & 0 & 4 \\
 & & & & 0 & 0 & 3 \\
 & & & & 0 & 0 & 2 \\
 & & & & 0 & 0 & 1
\end{array}\right) \otimes \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}, \tag{13}
$$

der $\otimes$ er **Kronecker**-produktet.

Da kan vi beregne vektoren $B_{4n}^{**\prime} = (\beta_0^{**}, \beta_1^{**}, \beta_2^{**}, ..., \beta_{4n-1}^{**})$, ved

$$
B_{4n}^{**\prime} = B_{4n}^{*\prime} \cdot H_{4n\times 4n}. \tag{14}
$$

Elementene i vektoren $B^{**\prime}$ er koeffisientene vi benytter i åresrelasjonen. I lag-polynomet $B^{**}(L) = (\beta_0^{**} + \beta_1^{**}L + \beta_2^{**}L^2 + ... + \beta_{4n-1}^{**}L^{4n-1})$ vil bare hver fjerde koeffisient være ulik null, og lagpolynomet kan derfor skrives $B^{**}(L) = (\beta_0^{**} + \beta_4^{**}L^4 + ... + \beta_{4(n-1)}^{**}L^{4(n-1)})$.

Ved å sette inn fra relasjon (9) for uobserverbare verdier av den eksogene variabelen, må vi også ta hensyn til restleddet for den eksogene variabelen. Her viser vi hvordan vi utleder den delen av restleddsprosessen som skyldes at vi ikke kan observere alle verdiene av den eksogene variabelen. Vi tar utgangspunkt i relasjonen i Teorem 2, der vi ser at koeffisientene i denne delen av MA-prosessen til restleddet avhenger av to forhold; (i) koeffisientene i lagpolynomet $B^*(L)$ (siden disse vil stå foran de eksogene variablene) og (ii) restleddene i klammeparentesen i relasjonen i Teorem 2. Vi lar $r = \left\{\frac{3p+q}{4}\right\}^+$, og kan da skrive

$$\begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_{4r-1} \end{pmatrix}_{4r\times 1} = U_{4r\times 4r} \cdot \begin{pmatrix} \beta_0^* \\ \beta_1^* \\ \vdots \\ \beta_{4r-1}^* \end{pmatrix}_{4r\times 1}, \tag{15}$$

der

$$U_{4r\times 4r} = \frac{1}{4} \cdot I_r \otimes \begin{pmatrix} 0 & -3 & -2 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & 3 \end{pmatrix}. \tag{16}$$

La lagpolynomet $\Gamma(L)$ inneholdet elementene i vektoren vi beregnet i (15), dvs. $\Gamma(L) = (\gamma_0 + \gamma_1 L + ... + \gamma_{4r-1} L^{4r-1})$. Dette lagpolynomet uttrykker den delen av MA-prosessen i restleddet som følge av at vi ikke kan observere alle verdiene til den eksogene variabelen.

Vi har nå endt opp med en relasjon bestående av bare observerbare verdier av både den endogene og den eksogene variabelen, samt en dobbelt sammensatt MA-prosess i restleddet. Den ene delen av MA-prosessen stammer fra kvartals-relasjonen etter at vi har multiplisert med C-polynomet, se relasjon (3). Den andre følger av at uobserverbare verdier av den eksogene variabelen er erstattet ved hjelp av relasjon (9). De to MA-prosessene er gjengitt i relasjon (17).

$$\begin{aligned} A_{4p}^*(L) \ln y_T &= C_{3p}(1)\alpha_0 + B_{4(n-1)}^{**}(L) \ln x_T + C_{3p}(L)\varepsilon_T + \Gamma_{4r-1}(L)u_T, \\ T &= 4, 8, 12, ..., \tau. \end{aligned} \tag{17}$$

Det eneste som gjenstår nå er å omforme MA-prosessen i restleddet til en MA-prosess på år. Dette gjøres ved å matche kovariansene svarende til år i en prosess $C^*(L) = \left(1 + c_4 L^4 + ...\right)$ med tilsvarende kovarianser for $C_{3p}(L)\varepsilon_T + \Gamma_{4r-1}(L)u_T$. Dette er en tilnærmelse, og samsvarer med metoden foreslått i Brewer (1973).

I restleddsprosessen i relasjon (17) er variansen og kovariansene gitt ved

$$Var(V_T) = \sigma^2 \sum_{i=0}^{m} \left(c_i^2 + \frac{\sigma_x^2}{\sigma^2} \gamma_i^2 \right) \tag{18}$$

$$Cov(V_T, V_{T-j}) = \sigma^2 \sum_{i=4}^{m} \left(c_i c_{i-j} + \frac{\sigma_x^2}{\sigma^2} \gamma_i \gamma_{i-j} \right), \ j = 4, 8, ..., 4n, \tag{19}$$

der σ^2 er variansen til restleddet i kvartalsrelasjonen i (1), σ_x^2 er variansen i den marginale prosessen (4), og $m = \max(3p, r)$ og $n = \left\{\frac{m}{4}\right\}^-$, hvor $\{\}^-$ uttrykker at innholdet i klammeparentesen avrundes nedover til nærmeste heltall.

La restleddsprosessen vi vil benytte i årsrelasjonen være

$$C^*(L) = (1 + c_4^* L^4 + c_8^* L^8 + ... + c_{4n}^* L^{4n}). \tag{20}$$

Denne restleddsprosessen har følgende varians og kovarianser

$$Var(V_T) = \sigma_\eta^2 \left(\sum_{i=0}^{n} (c_{4\cdot i}^*)^2 \right) \tag{21}$$

$$Cov(V_T, V_{T-j}) = \sigma_\eta^2 \left(\sum_{i=1}^{n} c_{4i}^* c_{4i-j}^* \right), \ j = 4, 8, ..., 4n, \tag{22}$$

hvor σ_η^2 er variansen i årsrelasjonen. Ved å sette uttrykkene for variansen og kovariansene i (18) og (19) lik de tilsvarende uttrykkene i (21) og (22) får vi $n+1$ relasjoner til å løse for variansen i årsrelasjonen (σ_η^2) og de n lagene i MA-prosessen på år $(c_4^*, c_8^*, ..., c_{4n}^*)$. Beregningen av MA-prosessen i restleddet illustreres ved et eksempel 2 i kapittel 2.2.

2.1 Eksempel 1

Vi skal her illustrere metoden ved hjelp av et eksempel der vi tar utgangspunkt i følgende kvartalssammenheng av beholdningsvariable.

$$(1 - \alpha_1 L) \ln y_t = \alpha_0 + (\beta_0 + \beta_1 L) \ln x_t + \varepsilon_t, \ t = 1, 2, 3, ..., \tau \tag{23}$$

Som vist i Wei (1981) kan vi multiplisere alle ledd i (23) med lagpolynomet $C(L) = \left(1 - (\alpha_1)^4 L^4\right) / (1 - \alpha_1 L) = \left(1 + \alpha_1 L + (\alpha_1)^2 L^2 + (\alpha_1)^3 L^3\right)$, jf. relasjon (2). Vi får da følgende relasjon der bare observerbare verdier av den endogene variabelen inngår.

$$\begin{aligned}
& (1 - (\alpha_1)^4 L) \ln y_T \\
= \ & C(1)\alpha_0 + \\
& \left(1 + \alpha_1 L + (\alpha_1)^2 L^2 + (\alpha_1)^3 L^3\right) (\beta_0 + \beta_1 L) \ln x_T \\
& + C(L)\varepsilon_T \qquad (24) \\
= \ & C(1)\alpha_0 + \beta_0 \ln x_T + (\alpha_1 \beta_0 + \beta_1) \ln x_{T-1} + ((\alpha_1)^2 \beta_0 + \alpha_1 \beta_1) \ln x_{T-2} \\
& + ((\alpha_1)^3 \beta_0 + (\alpha_1)^2 \beta_1) \ln x_{T-3} + (\alpha_1)^3 \beta_1 \ln x_{T-4} + C(L)\varepsilon_T \\
T = \ & 4, 8, 12, ..., \tau
\end{aligned}$$

Relasjonen i (24) består av både observerbare og uobserverbare verdier av den eksogene variabelen. Vi må derfor modellere de uobserverbare verdiene som funksjoner av de observerbare. Som vi har vist i Teorem 2 kan vi uttrykke den betingede forventningen til de uobserverbare eksogene variablene som et veid gjennomsnitt av den siste observerbare verdien før og den første observerbare etterpå. Men vi får også med oss restleddet i den marginale prosessen, se relasjon (9).

Ved å sette inn (9) i (24) for de uobserverbare verdiene for den eksogene variabelen, får vi følgende årsrelasjon:

$$
\begin{aligned}
& (1-(\alpha_1)^4 L)\ln y_T \\
= \; & C(1)\alpha_0 \\
& +\beta_0 \ln x_T \\
& +(\alpha_1\beta_0+\beta_1)\left[\frac{3}{4}\ln x_T+\frac{1}{4}\ln x_{T-4}\right] \\
& +((\alpha_1)^2\beta_0+\alpha_1\beta_1)\left[\frac{2}{4}\ln x_T+\frac{2}{4}\ln x_{T-4}\right] \\
& +((\alpha_1)^3\beta_0+(\alpha_1)^2\beta_1)\left[\frac{1}{4}\ln x_T+\frac{3}{4}\ln x_{T-4}\right] \\
& +(\alpha_1)^2\beta_1 \ln x_{T-4} \\
& -(\alpha_1\beta_0+\beta_1)\left[\frac{3}{4}u_T-\frac{1}{4}(u_{T-1}+u_{T-2}+u_{T-3})\right] \\
& -((\alpha_1)^2\beta_0+\alpha_1\beta_1)\left[\frac{2}{4}(u_T+u_{T-1})-\frac{2}{4}(u_{T-2}+u_{T-3})\right] \\
& -((\alpha_1)^3\beta_0+(\alpha_1)^2\beta_1)\left[\frac{1}{4}(u_T+u_{T-1}+u_{T-2})-\frac{3}{4}u_{T-3}\right] \\
& +C(L)\varepsilon_T
\end{aligned}
\tag{25}
$$

$$
\begin{aligned}
= \; & C(1)\alpha_0 \\
& +\left[\begin{array}{c} \frac{4}{4}\beta_0+\frac{3}{4}(\alpha_1\beta_0+\beta_1)+\frac{2}{4}((\alpha_1)^2\beta_0+\alpha_1\beta_1) \\ +\frac{1}{4}((\alpha_1)^3\beta_0+(\alpha_1)^2\beta_1)+\frac{0}{4}(\alpha_1)^3\beta_1 \end{array}\right]\ln x_T \\
& +\left[\begin{array}{c} \frac{0}{4}\beta_0+\frac{1}{4}(\alpha_1\beta_0+\beta_1)+\frac{2}{4}((\alpha_1)^2\beta_0+\alpha_1\beta_1) \\ +\frac{3}{4}((\alpha_1)^3\beta_0+(\alpha_1)^2\beta_1)+\frac{4}{4}(\alpha_1)^3\beta_1 \end{array}\right]\ln x_{T-4} \\
& -\left[\frac{3}{4}(\alpha_1\beta_0+\beta_1)+\frac{2}{4}((\alpha_1)^2\beta_0+\alpha_1\beta_1)+\frac{1}{4}((\alpha_1)^3\beta_0+(\alpha_1)^2\beta_1)\right]u_T \\
& -\left[-\frac{1}{4}(\alpha_1\beta_0+\beta_1)+\frac{2}{4}((\alpha_1)^2\beta_0+\alpha_1\beta_1)+\frac{1}{4}((\alpha_1)^3\beta_0+(\alpha_1)^2\beta_1)\right]u_{T-1} \\
& -\left[-\frac{1}{4}(\alpha_1\beta_0+\beta_1)-\frac{2}{4}((\alpha_1)^2\beta_0+\alpha_1\beta_1)+\frac{1}{4}((\alpha_1)^3\beta_0+(\alpha_1)^2\beta_1)\right]u_{T-2} \\
& +C(L)\varepsilon_T
\end{aligned}
$$

Ved å definere vektoren

$$
B^* = \begin{pmatrix} \beta_0 \\ \alpha_1\beta_0+\beta_1 \\ (\alpha_1)^2\beta_0+\alpha_1\beta_1 \\ (\alpha_1)^3\beta_0+(\alpha_1)^2\beta_1 \\ (\alpha_1)^3\beta_1 \end{pmatrix}, \tag{26}
$$

kan vi skrive (25) mer kompakt som

$$
\begin{aligned}
& (1-(\alpha_1)^4 L)\ln y_T \\
= \; & S(1)\alpha_0 + B^{*\prime}\begin{pmatrix} 1 \\ 3/4 \\ 1/2 \\ 1/4 \\ 0 \end{pmatrix}\ln x_T + B^{*\prime}\begin{pmatrix} 0 \\ 1/4 \\ 1/2 \\ 3/4 \\ 1 \end{pmatrix}\ln x_{T-4} \qquad (27) \\
& -B^{*\prime}\begin{pmatrix} 0 \\ 3/4 \\ 1/2 \\ 1/4 \\ 0 \end{pmatrix}u_T - B^{*\prime}\begin{pmatrix} 0 \\ -1/4 \\ 1/2 \\ 1/4 \\ 0 \end{pmatrix}u_{T-1} - B^{*\prime}\begin{pmatrix} 0 \\ -1/4 \\ -1/2 \\ 1/4 \\ 0 \end{pmatrix}u_{T-2} \\
& +C(L)\varepsilon_T.
\end{aligned}
$$

Restleddet ε_T i (24)-(27) har bare 3 kvartalslag, og restleddet u_T i (25)-(27) har bare 2 kvartalslag. Det betyr at det ikke vil være autokorrelasjon i restleddene i årsrelasjonen. Vi kan derfor skrive om (27) til

$$
\begin{aligned}
& (1-(\alpha_1)^4 L)\ln y_T \\
= \; & S(1)\alpha_0 + B^{*\prime}\begin{pmatrix} 1 \\ 3/4 \\ 1/2 \\ 1/4 \\ 0 \end{pmatrix}\ln x_T + B^{*\prime}\begin{pmatrix} 0 \\ 1/4 \\ 1/2 \\ 3/4 \\ 1 \end{pmatrix}\ln x_{T-4} + \epsilon_T, \qquad (28)
\end{aligned}
$$

der forventningen til restleddet er null og det ikke er autokorrelasjon i restleddet. Relasjon (28) er årsrelasjonen utledet fra kvartalsrelasjonen (23) gitt den marginale prosessen i (4).

2.2 Eksempel 2

Vi illustrerer her metoden for å beregne MA-prosessen i restleddet i årsrelasjonen når $p = q = 2$. C-polynomet i (17) har da $3p = 6$ lag, mens Γ-polynomet har $r = 7$ lag. Vi kan dermed beregne variansen og én kovarians.

$$
\begin{aligned}
Var(V_T) &= \sigma^2 v_0 \qquad (29) \\
Cov(V_T, V_{T-4}) &= \sigma^2 v_4, \qquad (30)
\end{aligned}
$$

der $v_0 = 1 + c_1^2 + c_2^2 + c_3^2 + c_4^2 + c_5^2 + c_6^2 + \frac{\sigma_x^2}{\sigma^2}(\gamma_0^2 + \gamma_1^2 + \gamma_2^2 + \gamma_3^2 + \gamma_4^2 + \gamma_5^2 + \gamma_6^2 + \gamma_7^2)$ og $v_4 = c_4 + c_1 c_5 + c_2 c_6 + \frac{\sigma_x^2}{\sigma^2}(\gamma_0\gamma_4 + \gamma_1\gamma_5 + \gamma_2\gamma_6 + \gamma_3\gamma_7)$. Denne variansen og kovariansen ønsker vi å overføre til restleddsprosessen i årsrelasjonen, som må ha ett årslag.

$$
C^*(L) = (1 + c_4^* L^4) \qquad (31)
$$

Variansen og kovariansen i denne prosessen er

$$
\begin{aligned}
Var(V_T) &= \sigma_\eta^2 \left(1 + (c_4^*)^2\right) &(32)\\
Cov(V_T, V_{T-4}) &= \sigma_\eta^2 c_4^*. &(33)
\end{aligned}
$$

Når vi setter både variansene og kovariansene like finner vi løsningene for σ_η^2 og c_4^*:

$$
\begin{aligned}
c_4^* &= \frac{v_0 \pm \sqrt{v_0^2 - 4v_4^2}}{2v_4} &(34)\\
\sigma_\eta^2 &= \frac{\sigma^2 v_4}{c_4^*} &(35)
\end{aligned}
$$

Som vi ser er det to mulige løsninger for c_4^* (og derned også for σ_η^2). Dette er vanlig for MA-prosesser. I en MA-prosess kan en hvilken som helst rot i lagpolynomet erstattes med den inverse av roten uten at autokorrelasjonen endres. (Multipliser sammen de to løsningene for c_4^* for å se at produktet deres er én.) Det sies ofte at de "deler den samme autokorrelasjonsfunksjonen". Det vanlige er i slike tilfeller å velge den verdien av c_4^* som gjør at røttene i lagpolynomet til MA-prosessen ligger utenfor (evt. på) enhetssirkelen. Ved bare ett lag i lagpolynomet er dette det samme som å velge den verdien av c_4^* som ligger innenfor enhetssirkelen (siden c_4^* i så fall er den inverse av roten). Fordelen med å benytte denne verdien er at relasjonen vil være invertibel.

Som vi ser kreves det at vi kjenner variansen i den marginale prosessen til den eksogene variabelen for å benytte denne metoden til å beregne MA-prosessen i restleddet, jf. relasjon (18) og (19). (Strengt tatt er det bare forholdet mellom variansen i den marginale prosessen og variansen i kvartalsrelasjonen som det er nødvendig å kjenne.) I forbindelse med omregningen av kvartalsrelasjoner til årsrelasjoner har vi ikke kjennskap til denne variansen. Et alternativ er da å estimere MA-prosessen fritt på årsvariablene. Siden man da ikke pålegger koeffisientrestriksjoner på restleddsprosessen, kan man være restriktiv med antall lag.

3 Sesong og trend (og andre dummyer)

I mange kvartalsrelasjoner er det med både trend og sesongdummyer. I tillegg kan kvartalsrelasjonen også inneholde dummyer. I årsrelasjonen vil sesongdummyene forsvinne, mens vi fortsatt vil ha en trend. I dette kapittelet viser vi hvordan vi omregner kvartalsrelasjoner til årsrelasjoner når relasjonen har trend og dummyer. For enkelhets skyld antar vi fortsatt at både den endogene og den eksogene variabelen er beholdningsvariable, men tilsvarende metode kan benyttes hvis en eller flere av variablene er strømningsvariable. Kvartalsrelasjonen er

$$\begin{aligned} A_p(L)\ln y_t &= \alpha_0 + S_2(L)s_t + d\cdot D + B_q(L)\ln x_t + \gamma t + \varepsilon_t, \\ t &= 1,2,3,...,\tau, \end{aligned} \tag{36}$$

der $S_2(L) = (s_1 + s_2L + s_3L^2)$, hvor s_i er sesongdummyen i kvartal i, mens D er en impulsdummy. Vektoren s har enere i første kvartal hvert år og nuller ellers, dvs $s = (1,0,0,0,1,0,0,0,...,1,0,0,0)$. Som tidligere multipliserer vi med C-polynomet for at alle verdiene av den endogene variabelen skal være observerbare i årsrelasjonen.

$$\begin{aligned} A^*_{4p}(L)\ln y_T &= a^*_0 + B^*_{3p+q}(L)\ln x_T + D^* + C(1)\gamma T + C_{3p}(L)\varepsilon_T \\ T &= 4,8,12,...,\tau, \end{aligned} \tag{37}$$

der $A^*(L) = C(L)A(L)$ og $B^*(L) = C(L)B(L)$. Legg merke til at $a^*_0 = C_{3p}(L)\left[\alpha_0 + S_2(L)s_T\right] + \left[C(L) - C(1)\right]\gamma T$ er en konstant. Det ser vi ved å splitte den opp. Første del av konstanten, $C_{3p}(L)\alpha_0 = C_{3p}(1)\alpha_0$, ser vi enkelt er konstant. I andre del gjentas $S_2(L)s_T$ hver fjerde periode siden $s_T = s_{T-4}$. Dermed er $C_{3p}(L)S_2(L)s_T$ konstant. Siste del, $\left[C(L) - C(1)\right]\gamma T = -\gamma\sum_{j=0}^{3p} c_j \cdot j$, ser vi også er uavhengig av tiden og dermed konstant.

I relasjon (37) har vi definert dummyvariabelen $D^* = d\cdot C(L)\cdot D$. Dummyvariabelen D^* kan normeres, slik at koeffisienten den multipliseres med er lik én. Vi forklarer dette ved et eksempel. Anta at D er en impulsdummy som er 1 i 1. kvartal 1990, og null ellers. Anta videre at A-polynomet har ett kvartalslag slik at C-polynomet har tre kvartalslag, dvs. $C(L) = (1 + c_1L + c_2L^2 + c_3L^3)$. Vektoren D^* vil da inneholde nuller for alle år bortsett fra 1990. I 1990 (som tilsvarer 4. kvartal 1990 i kvartalsrelasjonen) har vektoren verdien $d\cdot c_3$.

Hvis det er flere dummyer i kvartalsrelasjonen kan samme metode benyttes. Anta for enkelthets skyld at det er to impulsdummyer i kvartalsrelasjonen, D_1 og D_2 med koeffisientene d_1 og d_2. I årsrelasjonen kan vi da definere dummyvariabelen $D^* = C(L)\left[d_1D_1 + d_2D_2\right]$. Ved å definere dummyen i årsrelasjonen på denne måten ender vi opp med bare én dummy i årsrelasjonen.

Bortsett fra trenden og dummyvariabelen, står vi altså igjen med samme relasjon som før. Ved beregning av trenden er det også viktig å huske at trenden i årsrelasjonen øker med én enhet per år, og ikke fire enheter. Uttrykket for trenden bør derfor omskrives til $4C(1)\gamma\left(\frac{T}{4}\right)$.

Det kan også være sesongvariasjoner og trend i den marginale modellen til den eksogene variabelen, dvs.

$$(1-L)\ln x_t = \delta + S_2^x(L)s_t + \gamma^x t + u_t,\ t = 1, 2, 3, .., \tau \tag{38}$$

der $S_2^x(L) = (s_1^x + s_2^x L + s_3^x L^2)$, hvor s_i^x er sesongdummyen i kvartal i for den eksogene variabelen. Ved å benytte Teorem 2, får vi at forventningen til de uobserverbare verdiene av den eksogene variabelen (på en konstant nær) er et veid gjennomsnitt av den første observerbare verdien før og den første observerbare etter, dvs

$$\begin{aligned} E(\ln x_{T+j}\,|\ln x_T, \ln x_{T+4}) &= \alpha_j^x + \frac{4-j}{4}\ln x_T + \frac{j}{4}\ln x_{T+4} \\ \text{der } \alpha_0^x &= 0 \\ \alpha_1^x &= \frac{3}{4}s_1^x - \frac{1}{4}(s_2^x + s_3^x) - \frac{6}{4}\gamma^x \\ \alpha_2^x &= \frac{2}{4}(s_1^x + s_2^x) - \frac{2}{4}s_3^x - \frac{8}{4}\gamma^x \\ \alpha_3^x &= \frac{1}{4}(s_1^x + s_2^x + s_3^x) - \frac{6}{4}\gamma^x \\ \text{og } T &= 4, 8, 12, ..., \tau \quad j = 0, 1, 2, 3. \end{aligned} \tag{39}$$

Vi kan derfor fortsatt erstatte uobserverbare verdier av den eksogene variabelen med et veid gjennomsnitt av to observerbare verdier av variabelen, og justere konstantleddet i relasjonen for den endogene variabelen. Hvis vi ikke kjenner parametrene for sesongsvingningen i den marginale modellen (dvs. til den eksogene variabelen), kan vi ikke klare å beregne konstantleddet i relasjonen for den endogene variabelen. Men vi kan beregne alle de andre koeffisientene, og til slutt kalibrere konstantleddet.

4 Eksogen variabel er strømningsvariabel

Vi skal her vise hvordan vi kan tilpasse metoden for beregning av årsrelasjon når den eksogene variabelen er en strømningsvariabel. (Denne metoden må vi også benytte hvis årsverdien av variabelen er lik gjennomsnittet av kvartalsverdiene. Kvartalsgjennomsnittet er da $X_T/4$, der $X_T = x_T + x_{T-1} + x_{T-2} + x_{T-3}$ er årssummen. Når vi tar logaritmen av kvartalgjennomsnittet får vi $\ln(X_T/4) = \ln X_T - \ln 4$, som bare adskiller seg fra logaritmen av årssummen med en konstant.)

I årsrelasjonen er det bare årssummen av den endogene (strømnings-) variabelen $X_T = x_T + x_{T-1} + x_{T-2} + x_{T-3}$ som er observerbar. I beregningen videre vil vi benytte en førsteordens Taylor-approksimasjon. For x_{T-i} $(i = 0, 1, 2, 3)$ gjør vi en approksimasjon rundt $X_T/4$.

$$\begin{aligned}
\ln x_{T-i} &= \ln\left(\frac{1}{4}X_T + \left(x_{T-i} - \frac{1}{4}X_T\right)\right) \\
&\approx \ln\left(\frac{X_T}{4}\right) + \frac{x_{T-i} - \frac{1}{4}X_T}{\frac{1}{4}X_T} \\
&= \ln X_T - \ln 4 - 1 + \frac{4x_{T-i}}{X_T} \\
&= -(1 + \ln 4) + \ln X_T + \frac{4x_{T-i}}{X_T}
\end{aligned} \tag{40}$$

Det siste leddet i relasjonen over er ikke observerbart, og vi approksimerer derfor rundt dette punktet. Vi må derfor benytte egenskapene til den marginale prosessen, se relasjon (4). Fra den marginale prosessen vet vi at

$$\begin{aligned}
x_{T-1} &= \Lambda^{-1} x_T e^{-u_T} \\
x_{T-2} &= \Lambda^{-2} x_T e^{-u_T - u_{T-1}} \\
x_{T-3} &= \Lambda^{-3} x_T e^{-u_T - u_{T-1} - u_{T-2}},
\end{aligned} \tag{41}$$

der $\Lambda = e^{\delta}$. Ved å benytte disse sammenhengene kan vi skrive

$$\frac{4x_{T-i}}{X_T} = \frac{4\Lambda^{3-i} e^{\sum_{m=i}^{3} u_{T-i}}}{\Lambda^3 e^{u_T + u_{T-1} + u_{T-2}} + \Lambda^2 e^{u_{T-1} + u_{T-2}} + \Lambda e^{u_{T-2}} + 1}, \tag{42}$$

og

$$\frac{\partial\left(\frac{4x_{T-i}}{X_T}\right)}{\partial u_{T-j}} \tag{43}$$

$$= \begin{cases}
\dfrac{\left(4\Lambda^{3-i} e^{\sum_{k=i}^{2} u_{T-k}}\right)\left(\sum\limits_{l=0}^{2-j} \Lambda^l e^{\sum_{m=l+1}^{2} u_{T-m}}\right)}{\left(\Lambda^3 e^{u_T + u_{T-1} + u_{T-2}} + \Lambda^2 e^{u_{T-1} + u_{T-2}} + \Lambda e^{u_{T-2}} + 1\right)^2} & \text{for} \quad i \leq j = 0, 1, 2 \\
-\dfrac{\left(4\Lambda^{3-i} e^{\sum_{k=i}^{2} u_{T-k}}\right)\left(\sum\limits_{l=3-j}^{3} \Lambda^l e^{\sum_{m=3-l}^{2} u_{T-m}}\right)}{\left(\Lambda^3 e^{u_T + u_{T-1} + u_{T-2}} + \Lambda^2 e^{u_{T-1} + u_{T-2}} + \Lambda e^{u_{T-2}} + 1\right)^2} & \text{for} \quad i > j = 0, 1, 2 \\
0 & \text{for} \quad j = 3.
\end{cases}$$

Definerer så følgende hjelpevariable:

$$\begin{aligned}
\theta_{i,j} &= \left.\frac{\partial\left(\frac{4x_{T-i}}{X_T}\right)}{\partial u_{T-j}}\right|_{u_T=u_{T-1}=u_{T-2}=0} \\
&= \begin{cases} \frac{\left(4\Lambda^{3-i}\right)\left(\sum\limits_{l=0}^{2-j}\Lambda^l\right)}{(\Lambda^3+\Lambda^2+\Lambda+1)^2} & \text{for } i \leq j = 0,1,2 \\ -\frac{\left(4\Lambda^{3-i}\right)\left(\sum\limits_{l=3-j}^{3}\Lambda^l\right)}{(\Lambda^3+\Lambda^2+\Lambda+1)^2} & \text{for } i > j = 0,1,2 \end{cases} \qquad (44) \\
\theta_i &= \left.\frac{4x_{T-i}}{X_T}\right|_{u_T=u_{T-1}=u_{T-2}=0} = \frac{4\Lambda^{3-i}}{\sum\limits_{k=0}^{3}\Lambda^k}
\end{aligned}$$

Ved å benytter overstående kan skrive:

$$\begin{aligned}
&\ln x_{T-i} \\
=\ & -(1+\ln 4) + \ln X_T + \frac{4x_{T-i}}{X_T} \\
=\ & -(1+\ln 4) + \ln X_T \qquad (45) \\
&+ \left(\left.\frac{4x_{T-i}}{X_T}\right|_{u_T=u_{T-1}=u_{T-2}=0} + \sum_{j=0}^{2} \left.\frac{\partial\left(\frac{4x_{T-i}}{X_T}\right)}{\partial u_{T-j}}\right|_{u_T=u_{T-1}=u_{T-2}=0} \cdot u_{T-j} \right) \\
=\ & -(1+\ln 4 - \theta_i) + \ln X_T + (\theta_{i,0}u_T + \theta_{i,1}u_{t-1} + \theta_{i,2}u_{T-2})
\end{aligned}$$

$$E\left[\ln x_{T-i} \left| \ln X_T\right.\right] = -(1+\ln 4 - \theta_i) + \ln X_T \qquad (46)$$

Ved å erstatte uobserverbare verdier av den eksogene variabelen med observerbare verdier, ser vi fra relasjon (45) at vi får et restledd som følger en MA-prosess. Fra (44) ser vi videre at man er nødt til å kjenne konstantleddet i den marginale prosessen til den eksogene variabelen, se relasjon (4), for å kunne beregne koeffisientene i denne MA-prosessen. Siden vi ikke kjenner denne konstanten, kan vi heller ikke beregne MA-prosessen fra kvartalsrelasjonen. Alternativet blir da å estimere MA-prosessen i årsrelasjonen på årsdataene, jf. kapittel 2.

Vi kan også her beregne koeffisientene for den eksogene variabelen ved hjelp av matriseregning. Den eneste endringen vi trenger å gjøre er å benytte H^*-matrisen definert i (47) i steden for H-matrisen. Koeffisientene i polynomet for den eksogene variabelen i årsrelasjonen beregnes fortsatt ved hjelp av matrisemultiplikasjon i (14).

$$H^*_{4n\times 4n} = I_n \otimes \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad (47)$$

Tilsvarende kan vi beregne koeffisientene i MA-prosessen som skyldes at ikke alle verdiene av den eksogene variabelen er observerbare, det vil si koeffisientene i Γ-polynomet, ved matriseregning. Her benytter vi U^*-matrisen definert i (48) (i steden for U-matrisen) for å beregne koeffisientene i Γ-polynomet ved hjelp av matrisemultiplikasjon i (15).

$$U^*_{4n\times 4n} = I_n \otimes \begin{pmatrix} \theta_{0,0} & \theta_{1,0} & \theta_{2,0} & \theta_{3,0} \\ \theta_{0,1} & \theta_{1,1} & \theta_{2,1} & \theta_{3,1} \\ \theta_{0,2} & \theta_{1,2} & \theta_{2,2} & \theta_{3,2} \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{48}$$

5 Endogen variabel er strømningsvariabel

Hvis den endogene variabelen er en strømningsvariabel benytter vi følgende approksimasjon:

$$\begin{aligned} \ln Y_T &\approx \frac{1}{4}\sum_{j=0}^{3}\ln y_{T-j} + \ln 4 + \overline{\ln y} \\ T &= 4, 8, 12, ..., \tau, \end{aligned} \tag{49}$$

der $Y_T = y_T + y_{T-1} + y_{T-2} + y_{T-3}$ og

$$\overline{\ln y} = \frac{1}{(\tau/4)} \sum_{T\in(4,8,12,...,\tau)} \left(\ln Y_T - \frac{1}{4}\sum_{j=0}^{3}\ln y_{T-j} - \ln 4 \right). \tag{50}$$

Bevis. Vi beregner først en første ordens Taylor-approksimasjon rundt $4y_{T-j}$, $j = 0, 1, 2, 3$: $\ln Y_T \approx \ln(4y_{T-j}) + \sum_{i=0}^{3} \frac{y_{T-i}-y_{T-j}}{Y_T}$. Dette gir 4 alternative approksimasjoner for $\ln Y_T$. Et uveid gjennomsnitt av disse approksimasjonene gir

$$\begin{aligned} \ln Y_T &\approx \sum_{j=0}^{3}\frac{1}{4}\left[\ln 4y_{T-j} + \sum_{i=0}^{3}\frac{y_{T-i}-y_{T-j}}{Y_T}\right] \\ &= \frac{1}{4}\ln 4y_T + \frac{1}{4}\ln 4y_{T-1} + \frac{1}{4}\ln 4y_{T-2} + \frac{1}{4}\ln 4y_{T-3} \\ &= \frac{1}{4}\sum_{j=0}^{3}\ln y_{T-j} + \ln(4), \end{aligned} \tag{51}$$

fordi $\sum_{j=0}^{3}\sum_{i=0}^{3}\frac{y_{T-i}-y_{T-j}}{Y_T} = \frac{1}{Y_T}\sum_{i=0}^{3}(y_{T-i} - y_{T-i}) = 0$. I tillegg legger vi til den forventede approksimasjonsfeilen, gitt ved $\overline{\ln y}$ i relasjon (50). ∎

Ifølge Jensens ulikhet er $\ln Y_T \geq \frac{1}{4}[\ln y_T + \ln y_{T-1} + \ln y_{T-2} + \ln y_{T-3}] + \ln 4$. Dermed må $\overline{\ln y} \geq 0$.

I praksis innebærer overnevnte approksimasjon at vi benytter (logaritmen av) et geometrisk gjennomsnitt ($\frac{1}{4}\sum_{j=0}^{3}\ln y_{T-j}$) i steden for (logaritmen av) et aritmetisk gjennomsnitt ($\ln\left(\frac{Y_T}{4}\right)$), hvor vi korrigerer for den gjennomsnittlige approksimasjonsfeilen ($\overline{\ln y}$).[5]

---

[5] Den prosentvise feilen ved å benytte tilnærmingen i (49) kan tilnærmet skrives som

$$100 \cdot \left(\ln Y_T - \frac{1}{4}\sum_{j=0}^{3}\ln y_{T-j} - \ln 4 - \overline{\ln y} \right). \tag{52}$$

For å få et inntrykk av hvor stor feil vi gjør ved å benytte denne tilnærmingen, kan vi ta utgangspunkt i den datagenererende prosessen vi presenterer i kapittel 7. Dataene for utviklingen i den endogene variabelen er gjengitt i vedlegg 2. Hvis vi setter inn for $y_{T-j}, j = 0, 1, 2, 3$ og $Y_T = y_T + y_{T-1} + y_{T-2} + y_{T-3}$ i (52), finner vi at den største prosentvise feilen er på 0,02 prosent. Den gjennomsnittlige (beregnet ved aritmetisk gjennomsnitt av absolutte prosentavvik) er under 0,004 prosent. Til sammenligning er standardavviket i restleddet i den datagenererende prosessen 1 prosent. Approksiamsjonsfeilen vi gjør med relasjon (49) synes derfor å være liten.

Hvis den endogene variabelen er en gjennomsnittsvariabel kan vi benytte relasjon (49), med den endringen av vi trekker fra $\ln 4$ på hver side. Dette følger av et tilsvarende bevis som over, der et uveid gjennomsnitt av approksimasjonene $\ln\left(\frac{Y_T}{4}\right) \approx \ln(y_{T-j}) + \sum_{i=0}^{3} \frac{y_{T-i}-y_{T-j}}{Y_T}$ benyttes.

For å nyttegjøre oss av overstående approksimasjon må vi multiplisere kvartalsrelasjonen med et modifisert C-polynom. C-polynomet som vi må multiplisere med nå er:

$$
\begin{aligned}
C_{3p+3}(L) &= \frac{1}{4}\frac{1-L^4}{1-L}\prod_{j=1}^{p}\frac{1-\phi_j^4L^4}{1-\phi_jL} \\
&= \frac{1}{4}\left(1+L+L^2+L^3\right)\prod_{j=1}^{p}\left(1+\phi_jL+\phi_j^2L^2+\phi_j^3L^3\right) \qquad (53)
\end{aligned}
$$

der $\phi_1, \phi_2, ..., \phi_p$ fortsatt er de inverse av røttene i A-polynomet. Det modifiserte C-polynomet får altså tre flere lag enn det lagpolynomet vi benytter når den endogene variabelen er en beholdningsvariabel. Lag-polynomet har den egenskap at koeffisientene som står foran hvert av de fire kvartalene i et år vil være like etter at A-polynomet er multiplisert med C-polynomet.

Også her kan det være skjulte periodisiteter, jf. Teorem 1. Samme metode som for beholdningsvariable kan da benyttes.

Siden C-polynomet har tre ekstra lag (i forhold til tilfellet der den endogene variabelen var en beholdningsvariabel), vil B^*-polynomet ha tre ekstra lag. For å finne lagpolynomet til den eksogene variabelen multipliserer vi fortsatt B^*-polynomet med $H_{4n\times 4n}$- eller $H^*_{4n\times 4n}$-matrisen (avhengig av om den eksogene variabelen er en beholdnings- eller strømningsvariabel), der nå $n = 1 + \left\{\frac{3p+3+q}{4}\right\}^+$.

MA-prosessen beregnes som i kapittel 2.

5.1 Eksempel 3

Vi kan illustrere metoden ved hjelp av eksempel der A-polynomet har ett lag.

$$A_1(L)\ln y_t = (1-\alpha L)\ln y_t \qquad (54)$$

Den inverse av roten i dette lag-polynomet er α, og C-polynomet blir dermed

$$C_6(L) = \frac{1}{4}\begin{pmatrix} 1+(1+\alpha)L+(1+\alpha+\alpha^2)L^2+(1+\alpha+\alpha^2+\alpha^3)L^3 \\ +(\alpha+\alpha^2+\alpha^3)L^4+(\alpha^2+\alpha^3)L^5+\alpha^3L^6 \end{pmatrix}. \qquad (55)$$

Beregner så produktet av de to polynomene:

$$
\begin{aligned}
A_7^*(L) &= A_1(L)C_6(L) \\
&= (1-\alpha L)\frac{1}{4} \qquad (56) \\
&\quad \times\begin{pmatrix} 1+(1+\alpha)L+(1+\alpha+\alpha^2)L^2+(1+\alpha+\alpha^2+\alpha^3)L^3 \\ +(\alpha+\alpha^2+\alpha^3)L^4+(\alpha^2+\alpha^3)L^5+\alpha^3L^6 \end{pmatrix} \\
&= \frac{1}{4}\left(1(1+L+L^2+L^3)-\alpha^4(1+L+L^2+L^3)L^4\right)
\end{aligned}
$$

Ved å benytte approksimasjonen i relasjon (49), kan vi skrive

$$A_7^*(L)\ln y_T \approx (1-\alpha^4 L^4)\ln Y_T - (1-\alpha^4)\ln 4. \tag{57}$$

6 Konstantledd og MA-prosessen

Vi har tidligere i notatet vist at det kan være problemer med å beregne konstantleddet og koeffisientene i MA-prosessen i årsrelasjonen fra kvartalsrelasjonen. Hvis den eksogene variabelen er en strømningsvariabel, må man kjenne konstantleddet i den marginale prosessen til den eksogene variabelen for å beregne konstanten og MA-prosessen i årsrelasjonen. I tillegg må man kjenne forholdet mellom variansen i den marginale prosessen til den eksogene variabelen og variansen i kvartalsrelasjonen for å beregne MA-prosessen i årsrelasjonen. Et alternativ er da å estimere konstanten og MA-prosessen på årsdataene. Dette kan gjøres ved enten å benytte betinget Maximum Likelihood eller exact Maximum Likelihood. Ved (betinget) Maximum Likelihood settes den ubetingede forventningen til restleddene i årene før estimeringsperiodens begynnelse lik null. Ved exact Maximum Likelihood pålegger kovariansen i restleddsprosessen å være den samme for restleddene før estimeringsperioden som for restleddene i estimeringsperioden.

Et problem med en MA-prosess i restleddet er at det er vanskelig å forholde seg til restleddsprosessen ved prediksjoner når restleddet ikke settes lik null. Man kan sette restleddet ulik null ved modellprediksjoner for å "overstyre" modellen av forskjellige grunner. Det er da en stor fordel å ha et unikt restledd, og vi skal derfor her vise hvordan vi kan utnytte kjennskapen til MA-prosessen i restleddet samtidig som vi eliminerer problemet med å forholde seg til en MA-prosess i restleddet fremover i tid. Vi tar utgangspunkt i relasjon (58), som er en relasjon på år. For å unngå forvirring med lagoperatoren benytter vi også her L^4 som årslag-operator, og vi forenkler ved å benytte bare ett årslag. Metoden kan uten problemer benyttes ved flere lag.

$$(1-\alpha_4 L^4)\ln y_T = \alpha_0 + (\beta_0 + \beta_4 L^4)\ln x_T + (1+c_1 L^4)\varepsilon_T \tag{58}$$

La oss anta at vi kjenner verdiene til den endogene variabelen frem til og med periode τ (f.eks. 4. kvartal 1999), mens vi også kjenner fremtidige verdien til den eksogene variabelen. Vi ønsker å benytte relasjonen til å predikere verdien til den endogene variabelen i periode $\tau+4$ (f.eks. 4. kvartal 2000). Fra relasjon (59) ser vi at den eneste feilkilden er restleddet i periode $\tau+4$. Restleddet i periode τ er derimot kjent og benyttes i prediksjonen av den endogene variabelen i periode $\tau+4$.

$$\begin{aligned} & \ln y_{\tau+4} - E\left[\ln y_{\tau+4} \middle| \{\ln y_t\}_{t=4,8,..,\tau}, \{\ln x_\tau\}_{t=4,8,..}, \{\varepsilon_\tau\}_{t=4,8,..,\tau}\right] \\ = \ & \ln y_{\tau+4} - (\alpha_0 + \alpha_4 \ln y_\tau + \beta_0 \ln x_{\tau+4} + \beta_4 \ln x_\tau + c_1\varepsilon_\tau) \\ = \ & \varepsilon_{\tau+4}, \end{aligned} \tag{59}$$

der $\{z_t\}_{t=4,8,...,\tau} = \{z_4, z_8, ..., z_\tau\}$.

Hvis vi vil predikere verdien av den endogene variabelen i periode $\tau+8$, ser vi fra relasjon (60) at vi ikke benytter noen historiske restledd direkte.

$$\begin{aligned} & E\left[\ln y_{\tau+8} \middle| \{\ln y_t\}_{t=4,8,..,\tau}, \{\ln x_\tau\}_{t=4,8,..}, \{\varepsilon_\tau\}_{t=4,8,..,\tau}\right] \\ = \; & \alpha_0 + \alpha_4 L^4 E\left[\ln y_{\tau+4} \middle| \{\ln y_t\}_{t=4,8,..,\tau}, \{\ln x_\tau\}_{t=4,8,..}, \{\varepsilon_\tau\}_{t=4,8,..,\tau}\right] \\ & +\beta_0 x_{\tau+8} + \beta_4 x_{\tau+4} \end{aligned} \tag{60}$$

I en MA(4)-prosess i restleddet kan vi bare nyttegjøre oss denne restleddsprosessen ved predikering ett år (4 kvartaler) frem i tid. Tilsvarende kan vi med en MA($4q$)-prosess benytte oss av denne bare ved predikeringer til og med q år frem i tid. Vi skriver derfor om relasjon (58) til

$$(1 - \alpha_4 L^4) \ln y_T = \alpha_0 + (\beta_0 + \beta_4 L^4) x_T + (1 + c_1 D L^4)\varepsilon_T, \tag{61}$$

der D er en dummy som er lik 1 frem til og med periode $\tau + 4$, og null deretter. Denne dummyen tillater dermed at vi benytter kjennskapen om restleddet fra periode τ ved predikeringen av den endogene variabelen i perioden τ+4, samtidig som prediksjonen av den endogene variabelen i periode $\tau + 8$ ikke blir påvirket av at restleddet i periode $\tau + 4$ kan være satt ulik null.

7 Monte Carlo simulering

For å teste metoden gjennomfører vi en Monte Carlo simulering. Vi tar utgangspunkt i følgende datagenererende prosess (DGP) på kvartal:

$$\begin{aligned}
\ln y_t &= 0,5 \cdot \ln y_{t-1} + 0,3 \cdot \ln x_t + 0,2 \ln x_{t-1} + \varepsilon_t \\
\ln x_t &= \ln x_{t-1} + u_t \qquad (62)
\end{aligned}$$

der $E(\varepsilon_t) = 0$, $E(u_t) = 0$, $E(\varepsilon_t \cdot \varepsilon_s) = \sigma^2$, $E(u_t \cdot u_s) = \sigma_x^2$ og $E(\varepsilon_t \cdot u_t) = 0$, $\forall s, t$ hvor $s \neq t$. Ved hjelp av denne DGP-en simulerer vi verdier for y og x i perioden 1968Q1-1999Q4, der vi har forutsatt $\sigma = \sigma_x = 1\%$ og $y_{1968Q1} = x_{1968Q1} = 1$. Dataene i DGP-en er gjengitt i vedlegg 2. I (63) har vi estimert relasjonen over perioden 1970Q1-1999Q4, og vi ser at den estimerte relasjonen i liten grad avviker fra DGP-en.

$$\begin{aligned}
\ln y_t = &\underset{(0{,}002)}{-0,0018} + \underset{(0{,}072)}{0,4980} \cdot \ln y_{t-1} \qquad (63) \\
&+\underset{(0{,}108)}{0,3391} \cdot \ln x_t + \underset{(0{,}124)}{0,1637} \ln x_{t-1} + \varepsilon_t
\end{aligned}$$

$\mathrm{R}^2 = 87,66\%$	$\sigma = 1,12\%$
$\mathrm{DW} = 2,00$	$\mathrm{RSS} = 0,0145$
AR 1-5 $= \underset{[0{,}9301]}{0,26734}$	LinRes $= \underset{[0{,}9875]}{\mathrm{NA}}$

For å kunne sammenlikne årsrelasjonene med kvartalsrelasjonen har vi gjengitt en del tester. R^2, σ, og DW (Durbin-Watson) rapporterer henholdsvis forklaringskraft, standardavvik i restleddet og 1. ordens autokorrelasjon i restleddet. RSS rapporterer summen av kvadrerte restledd, og AR 1-5 tester autokorrelasjon opptil femte orden. I testen LinRes tester vi ved hjelp av en Wald-test om koeffisientene i den estimerte relasjonen (unntatt konstantleddet) er lik koeffisientene i DGP-en.

Vi skal nå regne om DGP-en ved fire alternativer; (i) både endogen og eksogen variabel er beholdningsvariable, (ii) endogen variabel er beholdningsvariabel mens eksogen variabel er strømningsvariabel, (iii) endogen variabel er strømningsvariabel mens eksogen variabel er beholdningsvariabel, og (iv) begge variabelen er strømningsvariable.

Den inverse av roten til lagpolynomet for den endogene variabelen er $0,5$. I de to første tilfellene er den endogene variabelen en beholdningssvariabel, slik at lagpolynomet vi skal multiplisere med er $C(L) = (1 + \frac{1}{2}L + \frac{1}{4}L^2 + \frac{1}{8}L^3)$, jf. (2). Når eksogen variabel er beholdningsvariabel erstatter vi uobserverbare verdier med et veid gjennomsnitt av observerbare verdier, jf. (9). Vi ender da opp med følgende relasjon:

$$\ln y_T = 0,0625 \ln y_{T-4} + 0,671875 \ln x_T + 0,265625 \ln x_{T-4} + \epsilon_T \qquad (64)$$

der uttrykket for ϵ_T er beregnet i eksempelet i kapittel 2.1. Det er ingen autokorrelasjon på år i dette restleddet.

Når den eksogene variabelen er strømningsvariabel erstatter vi uobserverbare verdier av den eksogene variabelen med første observerbare etter, se relasjon (45). Vi ender da opp med relasjonen

$$\ln y_T = 0,0625 \ln y_{T-4} + 0,9125 \ln X_T + 0,025 \ln X_{T-4} + \epsilon_T \tag{65}$$

der $X_T = x_T + x_{T-1} + x_{t-2} + x_{T-3}$ og ϵ_T følger en MA(3) prosess slik at det ikke er noen autokorrelasjon på år.

Når den endogene variabelen er en strømningsvariabel multipliserer vi med $C(L) = \frac{1}{4}(1+1\frac{1}{2}L+1\frac{3}{4}L^2+1\frac{7}{8}L^3+\frac{7}{8}L^4+\frac{3}{8}L^5+\frac{1}{8}L^6)$, jf. eksempel 5.1. Venstre side blir da:

$$\begin{aligned} & \frac{1}{4}(1+1\frac{1}{2}L+1\frac{3}{4}L^2+1\frac{7}{8}L^3+\frac{7}{8}L^4+\frac{3}{8}L^5+\frac{1}{8}L^6)(1-0,5L)\ln y_T \\ = & \frac{1}{4}(1-0,0625L^4)(1+L+L^2+L^3)\ln y_T \end{aligned} \tag{66}$$

Benytter videre at $\ln Y_T \approx \frac{1}{4}[\ln y_T + \ln y_{T-1} + \ln y_{T-2} + \ln y_{T-3}] + \ln 4 + \overline{\ln y}$, der $\overline{\ln y}$ er gitt ved relasjon (50).

$$\begin{aligned} & \frac{1}{4}(1-0,0625L^4)(1+L+L^2+L^3)\ln y_T \\ \approx & (1-0,0625L^4)\ln Y_T + \text{constant} \end{aligned} \tag{67}$$

For den eksogene variabelen får vi:

$$\begin{aligned} & \frac{1}{4}(1+1\frac{1}{2}L+1\frac{3}{4}L^2+1\frac{7}{8}L^3+\frac{7}{8}L^4+\frac{3}{8}L^5+\frac{1}{8}L^6)(0,3+0,2L)\ln x_T \\ = & 0,075 \ln x_T + 0,1625 \ln x_{T-1} + 0,20625 \ln x_{T-2} \\ & +0,22813 \ln x_{T-3} + 0,15938 \ln x_{T-4} + 0,071875 \ln x_{T-5} \\ & +0,028125 \ln x_{T-6} + 0,00625 \ln x_{T-7} \end{aligned} \tag{68}$$

I tilfellet der den eksogene variabelen er en beholdningsvariabel, dvs. tilfelle (iii), tar vi utgangspunkt i lagpolynomet for den eksogene variabelen etter at vi har multiplisert med C-polynomet.

$$\begin{pmatrix} 0,075+0,1625L+0,20625L^2+0,22813L^3+0,15938L^4 \\ +0,071875L^5+0,028125L^6+0,00625L^7 \end{pmatrix} \ln x_T \tag{69}$$

Ved å sette inn for de uobserverbare verdiene i (69) ved hjelp av relasjon (9) i Teorem 2, ender vi opp med

$$\begin{aligned} & 0,35703 \ln x_T + 0,54376 \ln x_{T-4} + 0,036719 \ln x_{T-8} \\ & -0,28203u_T - 0,11953u_{T-1} + 0,086718u_{T-2} + 0,31485u_{T-3} \\ & -0,069531u_{T-4} + 0,0023438u_{T-5} + 0,030469u_{T-6} + 0,036719u_{T-7}. \end{aligned} \tag{70}$$

Vi har dermed funnet både lag-polynomet til den eksogene variabelen og MA-prosessen i restleddet som skyldes at vi bare benytter observerbare verdier av den eksogene variabelen. I tillegg har vi en MA-prosess fra kvartalsrelasjonen. Vi skal nå omregne disse to MA-prosessene til en MA-prosess på år. Når vi tar hensyn til at variansen i kvartalsrelasjonen og den marginale prosessen er like, er årsvariansen og årskovariansen

$$\begin{aligned} Var(V_T) &= \sigma^2 \cdot 0,87947 \\ Cov(V_T, V_{T-4}) &= \sigma^2 \cdot 0,13705, \end{aligned} \tag{71}$$

jf. uttrykkene i (18) og (19). Vi benytter så relasjon (34) til å finne koeffisienten i MA-prosessen på år:

$$c_4^* = \frac{0,87947 \pm \sqrt{0,87947^2 - 4 \cdot 0,13705^2}}{2 \cdot 0,13705} = \begin{cases} 6,2573 \\ 0,15981 \end{cases} \tag{72}$$

Vi har dermed kommet frem til følgende relasjon

$$\begin{aligned} \ln Y_T = \; & 0,0625 \ln Y_{T-4} \\ & +0,35703125 \ln x_T + 0,54375 \ln x_{T-4} + 0,03671875 \ln x_{T-8} \\ & +\epsilon_T + 0,15981 \epsilon_{T-4} \end{aligned} \tag{73}$$

På tilsvarende måte finner vi relasjonen når også den eksogene variabelen er en strømningsvariabel. Vi tar igjen utgangspunkt i lagpolynomet for den eksogene variabelen etter at vi har multiplisert med C-polynomet, jf. relasjon (68). Setter inn fra (45), når vi ser bort fra konstantleddet, og husker at $\delta = 0 \Rightarrow \Lambda = 1$, som ifølge relasjon (44) impliserer

$$\theta_{ij} = \begin{cases} \frac{3-j}{4} & \text{for } i \leq j = 0,1,2,3 \\ -\frac{j+1}{4} & \text{for } i > j = 0,1,2. \end{cases} \tag{74}$$

Etter at vi har satt inn for de uobserverbare verdiene av den eksogene variabelen ender vi opp med (75).

$$\begin{aligned} & 0,67188 \ln X_T + 0,26563 \ln X_{T-4} \\ & -0,09297 u_T - 0,09844 u_{T-1} - 0,06016 u_{T-2} \\ & +0,092973 u_{T-4} + 0,0\,9844 u_{T-5} + 0,060158 u_{T-6} \end{aligned} \tag{75}$$

Vi har nå funnet både lagpolynomet til den eksogene variabelen og MA-prosessen som skyldes at uobserverbare verdier av den eksogene variabelen er erstattet med observerbare. Beregner varians og kovarians, samt koeffisientene i MA-prosessen:

$$\begin{aligned} Var(V_T) &= \sigma^2 \cdot 0,71578 \\ Cov(V_T, V_{T-4}) &= \sigma^2 \cdot 0,081563 \end{aligned} \tag{76}$$

$$c_4^* = \frac{0,71578 \pm \sqrt{0,71578^2 - 4 \cdot 0,081563^2}}{2 \cdot 0,081563} = \begin{cases} 8,6603 \\ 0,11547 \end{cases} \tag{77}$$

Dette gir følgende årsrelasjon hvis den eksogene variabelen er en strømningsvariabel

$$\begin{aligned} \ln Y_T &= 0,0625 \ln Y_{T-4} \\ &\quad +0,671875 \ln X_T + 0,265625 \ln X_{T-4} \\ &\quad +\epsilon_T + 0,11547\epsilon_{T-4} \end{aligned} \tag{78}$$

I tabell 1 og 2 har vi gjengitt årsrelasjonene beregnet fra kvartalsrelasjonene. I tillegg har vi estimert årsrelasjonene ved hjelp av programmene PcGive 9.1 og Microfit 4.0, se Hendry og Doornik (1996) og Pesaran og Pesaran (1997). Alle estimeringene er gjort fra og med 1970.

Fra tabellene ser vi at koeffisientverdiene i de bergnede relasjonene i liten grad avviker fra de estimerte verdiene. Verdiene i parentes er standardavvik, mens verdiene i hakeparentes er p-verdier. I LinRes testes ved hjelp av en Wald-test hypotesen om koeffisientene i den estimerte relasjonen er lik de beregnede koeffisientene. I alle fire tilfellene er vi langt fra en forkastning av at de beregnede relasjonene er korrekte.

Tabell 1: Sammenligning av beregnede og estimerte årsrelasjoner

$\ln y_t$	(i) y og x		(ii) y og X	
	Beregnet	Estimert	Beregnet	Estimert
Constant	$-0,0013$	$-0,0002$	$-1,3023$	$-1,3256$
	(0,003)	(0,006)	(0,003)	(0,311)
$\ln y_{t-1}$	$0,0625$	$0,0365$	$0,0625$	$0,0067$
	(NA)	(0,190)	(NA)	(0,209)
$\ln x_t / \ln X_t$	$0,6719$	$0,7573$	$0,9125$	$0,9268$
	(NA)	(0,174)	(NA)	(0,189)
$\ln x_{t-1} / \ln X_{t-1}$	$0,2656$	$0,2264$	$0,0250$	$0,0262$
	(NA)	(0,242)	(NA)	(0,243)
R^2	78,81%	79,01%	78,69%	78,88%
σ	1,44%	1,51%	1,44%	1,52%
DW	1,90	1,92	1,95	1,93
RSS	0,0060	0,0059	0,0060	0,0059
AR 1-1	$0,0277$	$0,2407$	$0,0036$	$0,0006$
	[0,87]	[0,63]	[0,95]	[0,98]
LinRes	$0,2503$ $[0,97]$		$0,2410$ $[0,97]$	

Tabell 2: Sammenligning av beregnede og estimerte årsrelasjoner

$\ln Y_t$	(iii) Y og x		(iv) Y og X	
	Beregnet	Estimert	Beregnet	Estimert
Constant	NA (NA)	1.2811 (0,459)	NA (NA)	$-0,0341$ (0,086)
$\ln Y_{t-1}$	$0,0625$ (NA)	0.0760 (0.331)	0.0625 (NA)	0.0687 (0.243)
$\ln x_t / \ln X_t$	0.3570 (NA)	0.4055 (0.105)	0.6719 (NA)	0.7085 (0.102)
$\ln x_{t-1} / \ln X_{t-1}$	0.5437 (NA)	0.6369 (0.212)	0.2657 (NA)	0.2458 (0.2543)
$\ln x_{t-2} / \ln X_{t-2}$	0.03672 (NA)	-0.0743 (0.203)		
$\hat{\varepsilon}_{t-1}$	0.1598 (NA)	0.3145 (0.287)	0.1155 (NA)	0.2117 (0.263)
R^2	NA	90.99%	NA	92.38%
σ	NA	0.99%	NA	0.89%
DW	NA	1.94	NA	1.98
RSS	NA	0.0024	NA	0.0020
AR 1-1	NA	NA	NA	NA
LinRes	2.6083 [0.76]		0.5192 [0.97]	

8 Konklusjon

Vi har i dette notatet vist hvordan vi kan beregne årsrelasjoner fra kvartalsrelasjoner. Metoden avhenger av om variablene er beholdnings- eller strømningsvariable. Hvis en eller flere av variablene er strømningsvariable må en approksimasjon benyttes.

Metoden beskrevet i denne rapporten er benyttet av Statistisk sentralbyrå til å beregne enkelte av relasjonene i årsmodellen MODAG på bakgrunn av tilsvarende relasjoner i kvartalsmodellen KVARTS. I Hungnes (2000) vises det at skiftberegninger på de beregnede MODAG-relasjonene blir sammenfallende med skiftberegninger på KVARTS-relasjonene. Beregnede referansebaner avviker derimot noe.

Problemet med tidsaggregering av strømningsvariable har likhetstrekk med teorien for aggregering av relasjoner i mikro til makro når sammenhengene beskrives ved en loglineær funksjon. I videre arbeid vil det derfor være interessant å undersøke om det er resultater i teorien for mikro/makro-aggregering som kan benyttes ved tidsaggregering for at approksimasjonene skal bli bedre.

En alternativ fremgangsmåte er å skrive hele systemet av relasjoner som et VAR-system. Dette kan spesielt være effektivt hvis den eksogene variabelen ikke er sterk eksogen. Men også i tilfellet der den eksogene variabelen er strengt eksogen kan en slik tilnærming være bedre, siden man i større grad kan utnytte informasjonen i hele systemet av relasjoner. Flere konsekvenser av tidsaggregering innenfor et VAR-system er vist i Marcellino (1999).

Referanser

Amemiya, T. og R. Y. Wu (1972), "The Effect on Aggregation and Prediction in the Autoregressive Model", *Journal of the American Statistical Assosiation*, **339**, pp. 628-32.

Brewer, K. R. W. (1973), "Some Consequences of Temporal Aggergation and Systematic Sampling for ARMA and ARMAX Models", *Journal of Econometrics*, pp. 133-54

Hendry, David F. og Jurgen A. Doornik (1996), *Empirical Econometric Modelling Using PcGive for Windows*, London: Timberlake Consulting

Hungnes, Håvard (2000), "Omregning av KVARTS-relasjoner til MODAG-relasjoner", Notat 2000/28, Statistisk sentralbyrå

Kendall, M. G. (1973), *Time Series Models*, Halsted Press: Oxford

Marcellino, Massimiliano (1999), "Some Consequences of Temporal Aggergation in Empirical Analysis", *Journal of Business & Economic Statistics*, pp. 129-36

Peseran, M. Hashem og Bahram Peseran (1997), *Microfit 4.0 An Interactive Econometric Software Package*, Oxford University Press

Starm, Daniel O. og William W. S. Wei (1986), "Temporal Aggregation in the ARIMA Process", *Journal of Time Series Analysis*, pp. 279-92

Telser, L. G. (1967), "Discrete Samples and Moving Sums in a Stationary Stochastic Prosess", *Journal of the American Statistical Assosiation*, pp. 484-99

Tiao, G. C. og William W. S. Wei (1976), "Effect of Temporal Aggregation on the Dynamic Relationship of two Time Series Variables", *Biometrika*, pp. 513-23

Wei, William W. S. (1981), "Effect of Systematic Sampling on ARIMA Models", *Communications in Statistical-Theoretical Mathematics A 10*, pp. 2389-98

Wei, William W.S (1989), *Time series analysis: univariate and multivariate methods*, Redwood City, Calif.: Addison Wesley

Weiss, Andrew A. (1984), "Systematical Sampling and Temporal Aggregation in Time Series Models", *Journal of Econometrics*, pp. 271-81

A Vedlegg A:

A.1 Omregning mellom ADL(p,q)- og EqCM(p,q)-modell

Anta at vi har følgende modell, der y er den endogene variabelen og x er den eksogene variabelen. (Modellen kan utvides til å ha med flere eksogene variable, der alle 'behandles' på samme måte som x.)

$$A_p(L)\ln y_t = \alpha_0 + B_q(L)\ln x_t + \varepsilon_t \tag{79}$$

der $A_p(L) = (1 - \alpha_1 L - \alpha_2 L^2 - ... - \alpha_p L^p)$ og $B_q(L) = (\beta_0 + \beta_1 L + \beta_2 L^2 + ... + \beta_q L^q)$. ADL-modellen ovenfor kan skrives om til en EqCM-modell. Hvis vi legger feiljusteringsleddet på første lag, vil EqCM-modellen bli

$$\begin{aligned} \Delta \ln y_t &= \alpha_0^* + \sum_{i=1}^{p-1} \alpha_j^* \Delta \ln y_{t-i} + \sum_{i=0}^{q-1} \beta_i^* \Delta \ln x_{t-i} \\ &\quad - A(1)\left[\ln y_{t-1} - \frac{B(1)}{A(1)}\ln x_{t-1}\right] + \varepsilon_t, \end{aligned} \tag{80}$$

der

$$\alpha_i^* = \begin{cases} \alpha_0 & \text{for } i = 0 \\ -\sum\limits_{j=i+1}^{p} \alpha_j & \text{for } i = 1,2,3,...,p-1 \end{cases} \tag{81}$$

$$\beta_i^* = \begin{cases} \beta_0 & \text{for } i = 0 \\ -\sum\limits_{j=i+1}^{q} \beta_j & \text{for } i = 1,2,3,...,q-1 \end{cases} \tag{82}$$

$$\begin{aligned} A(1) &= 1 - \sum_{j=1}^{p} \alpha_j \quad \text{(summen av elementene i A-polynomet)} \\ B(1) &= \sum_{j=0}^{q} \beta_j \quad \text{(summen av elementene i B-polynomet).} \end{aligned} \tag{83}$$

Hvis man skal regne om fra en EqCM-modell til en ADL-modell får vi tilsvarende

$$\alpha_i = \begin{cases} \alpha_0^* & \text{for } i = 0 \\ a_1^* + 1 - A(1) & \text{for } i = 1 \\ \alpha_i^* - \alpha_{i-1}^* & \text{for } i = 2,3,...,p-1 \\ -\alpha_{p-1}^* & \text{for } i = p \end{cases} \tag{84}$$

$$\beta_i = \begin{cases} \beta_0^* & \text{for } i = 0 \\ \beta_1^* - \beta_0^* + B(1) & \text{for } i = 1 \\ \beta_i^* - \beta_{i-1}^* & \text{for } i = 2,3,...,q-1 \\ -\beta_{q-1}^* & \text{for } i = q. \end{cases} \tag{85}$$

B Vedlegg B:

B.1 Data i datagenererende prosess

	x	y
1968Q1	1,000000000	1,000000000
1968Q2	1,021664027	1,009718694
1968Q3	0,999492722	1,019892949
1968Q4	1,003321967	1,001677627
1969Q1	1,007210905	1,002379648
1969Q2	1,008077601	0,999590763
1969Q3	0,996927919	1,004850830
1969Q4	0,998425450	1,004159604
1970Q1	1,013971129	1,004912095
1970Q2	1,020999344	1,009427549
1970Q3	1,008005952	1,014876453
1970Q4	1,003203586	1,006767150
1971Q1	1,002276358	1,019608584
1971Q2	0,998868683	1,015048193
1971Q3	0,999211427	0,994257013
1971Q4	0,987882590	0,989934123
1972Q1	0,993567119	0,984969852
1972Q2	1,006516296	1,005385003
1972Q3	1,006765962	1,017852480
1972Q4	1,005205466	1,008591040
1973Q1	0,998508655	1,001603675
1973Q2	1,003401370	0,994779965
1973Q3	0,997508320	1,000881438
1973Q4	0,991383900	1,018358734
1974Q1	0,973429643	0,990435804
1974Q2	0,992256097	0,973135242
1974Q3	0,982510193	0,975600881
1974Q4	0,992939285	0,971789338
1975Q1	0,996996978	0,979894365
1975Q2	1,001896660	0,982874057
1975Q3	0,995112546	0,975650541
1975Q4	0,986874792	0,956411651
1976Q1	0,982602158	0,972927672
1976Q2	0,968938106	0,981741904
1976Q3	0,962679520	0,982002065
1976Q4	0,965108701	0,980129294
1977Q1	0,960987680	0,985605579
1977Q2	0,963903191	0,978611301
1977Q3	0,959759668	0,973259361
1977Q4	0,955202503	0,988293038
1978Q1	0,963724456	0,976042208

1978Q2	0,971729328	0,959230431
1978Q3	0,972343386	0,950627056
1978Q4	0,958227704	0,958961393
1979Q1	0,967067183	0,971799547
1979Q2	0,953548823	0,943490677
1979Q3	0,950501176	0,949134094
1979Q4	0,948102263	0,934038087
1980Q1	0,924108541	0,923353925
1980Q2	0,911335104	0,915616524
1980Q3	0,903579055	0,914242925
1980Q4	0,916983108	0,911784842
1981Q1	0,922941842	0,919943271
1981Q2	0,916348410	0,928740937
1981Q3	0,929408590	0,918205005
1981Q4	0,916833982	0,917276202
1982Q1	0,921681090	0,925049234
1982Q2	0,926823463	0,916923532
1982Q3	0,937229774	0,936417094
1982Q4	0,938832888	0,922916405
1983Q1	0,945585802	0,942192902
1983Q2	0,941956392	0,938171502
1983Q3	0,949102296	0,942290369
1983Q4	0,928423755	0,946438077
1984Q1	0,929042798	0,941787274
1984Q2	0,940881885	0,939689828
1984Q3	0,941531220	0,948375978
1984Q4	0,944124397	0,957900416
1985Q1	0,946707268	0,966875152
1985Q2	0,935740687	0,943964062
1985Q3	0,928265623	0,919509579
1985Q4	0,927575956	0,930640817
1986Q1	0,929432776	0,914037409
1986Q2	0,924471854	0,927323054
1986Q3	0,929682230	0,915298082
1986Q4	0,941511397	0,916571807
1987Q1	0,942715642	0,920317942
1987Q2	0,958898648	0,934878936
1987Q3	0,971801348	0,955023587
1987Q4	0,964254436	0,955961054
1988Q1	0,967625919	0,980013731
1988Q2	0,982287882	0,959169083
1988Q3	0,992056010	0,950346488
1988Q4	0,982395541	0,979278114
1989Q1	0,979388421	0,968371793
1989Q2	0,978919888	0,968665843
1989Q3	0,978697221	0,975876101

1989Q4	0,965289548	0,954115135
1990Q1	0,961980077	0,947062284
1990Q2	0,963642956	0,938747762
1990Q3	0,953809553	0,936890439
1990Q4	0,943433362	0,951429237
1991Q1	0,941610476	0,950431247
1991Q2	0,940026311	0,933153459
1991Q3	0,929102306	0,939131935
1991Q4	0,930336655	0,931205704
1992Q1	0,940727652	0,936957148
1992Q2	0,943421563	0,914587638
1992Q3	0,959118075	0,932297051
1992Q4	0,970604891	0,964322133
1993Q1	0,965728084	0,966417324
1993Q2	0,977196718	0,972422429
1993Q3	0,972188661	0,980621842
1993Q4	0,967247242	0,970757570
1994Q1	0,973377443	0,977907026
1994Q2	0,954343357	0,965353426
1994Q3	0,963017367	0,947207166
1994Q4	0,956954435	0,949131437
1995Q1	0,961736487	0,961713514
1995Q2	0,946091470	0,943151117
1995Q3	0,939053296	0,947079344
1995Q4	0,940333555	0,953812804
1996Q1	0,924810128	0,943858772
1996Q2	0,922893598	0,938943038
1996Q3	0,923599691	0,917128973
1996Q4	0,929331671	0,924673736
1997Q1	0,933269057	0,917859585
1997Q2	0,916082288	0,931883229
1997Q3	0,909803209	0,912447856
1997Q4	0,929784874	0,924415266
1998Q1	0,927494592	0,915977783
1998Q2	0,920592776	0,904080931
1998Q3	0,921954216	0,915214429
1998Q4	0,921891309	0,936735987
1999Q1	0,919957860	0,918731789
1999Q2	0,936785957	0,910798727
1999Q3	0,926826213	0,916518162
1999Q4	0,912887761	0,903789890