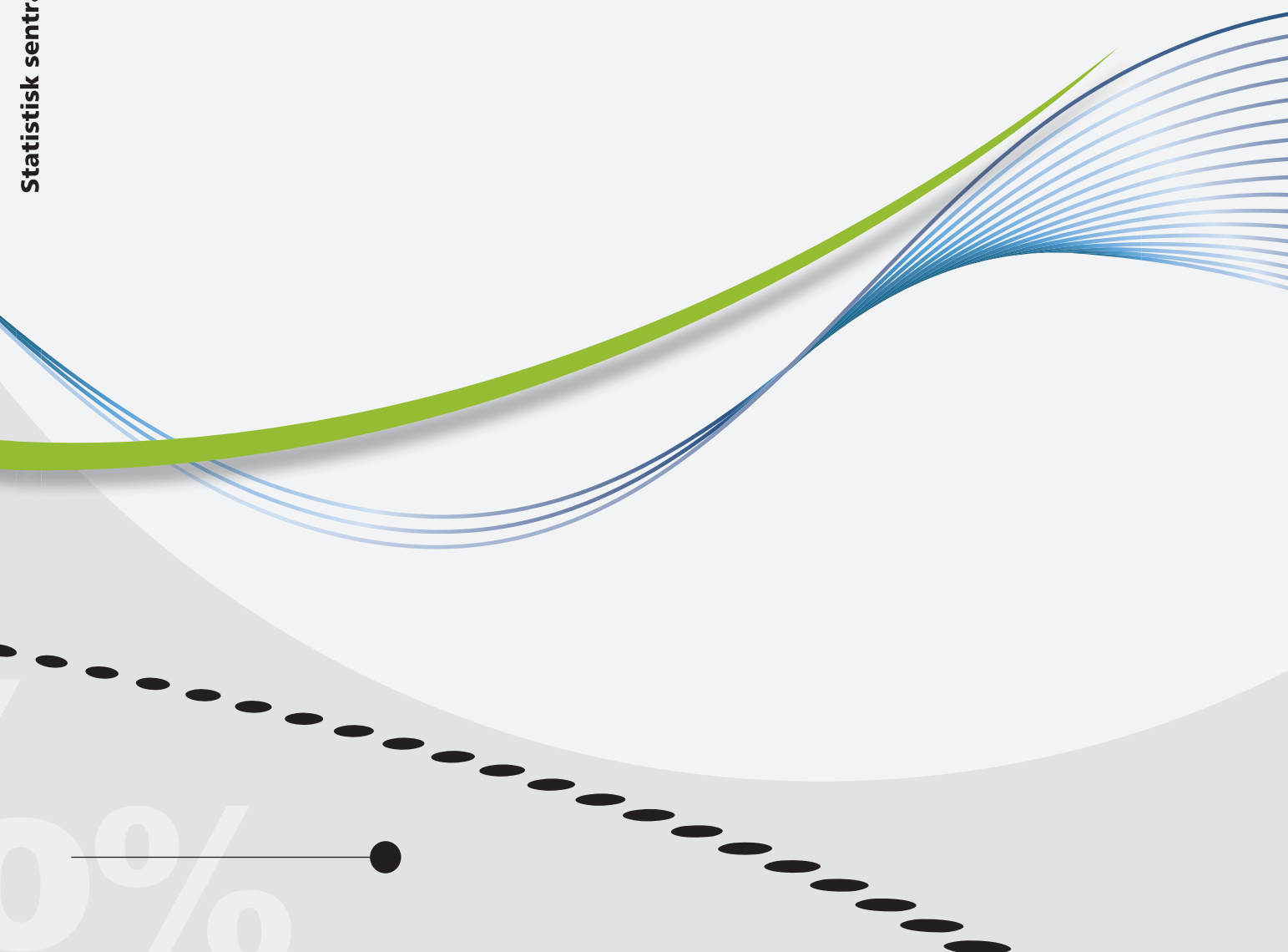




Dinh Quang Pham

Forbedret påskekorrigering for detaljomsetning



Dinh Quang Pham

Forbedret påskekorrigering for detaljomsetning

© Statistisk sentralbyrå	Standardtegn i tabeller	Symbol
Ved bruk av materiale fra denne publikasjonen skal	Tall kan ikke forekomme	.
Statistisk sentralbyrå oppgis som kilde.	Oppgave mangler	..
Publisert januar 2013	Oppgave mangler foreløpig	...
	Tall kan ikke offentliggjøres	:
	Null	-
ISBN 978-82-537-8567-7 (trykt)	Mindre enn 0,5 av den brukte enheten	0
ISBN 978-82-537-8568-4 (elektronisk)	Mindre enn 0,05 av den brukte enheten	0,0
ISSN 1891-5906	Foreløpig tall	*
Emne: 08.03.20 og 08.90	Brudd i den loddrette serien	—
	Brudd i den vannrette serien	
Trykk: Statistisk sentralbyrå	Desimaltegn	,

Forord

Ujusterte tall må korrigeres for effektene av ukedager, påske og pinse for sesongjustering. Tidsserier i detaljhandelomsetningsvolumindeksen (DOI) blir påvirket av påskeeffekten. Siden den innebygde rutinen i X-12-ARIMA for påskekorrigering ikke er tilpasset med norske data, laget Solheim og Dinh (1997) og Dinh (2007) en ny rutine. Det er imidlertid fortsatt ulemper med rutinen men nå kommer en forbedring av metoden. Denne blir beskrevet i dette notatet. Resultatene er blitt presentert for metodegruppen. Jeg vil takke mine kolleger for nyttige kommentarer. En spesiell takk til Jan Bjørnstad, Leiv Solheim og Roger Jensen. Jeg vil også takke Ane Seierstad for korrekturlesing av manuskriptet.

Statistisk sentralbyrå, 7. desember 2012

Hans Henrik Scheel

Sammendrag

Handelen blir redusert i påsken og øker i mange dager før, dvs at handelsmønsteret forskyves. Siden påsken ikke faller på samme dato hvert år, endres handelsmønsteret i mars og april fra år til år. Dette medfører at sesongvariasjoner i mars og april blir ustabile. Uten å korrigere de ujusterte tallene for denne påskeeffekten, før sesongjustering, vil effekten av påske fremdeles finnes i de sesongjusterte tallene. Dette medfører at vi ikke får et riktig bilde av aktiviteten i handelen i mars og april i de sesongjusterte tallene.

Et problem med påskekorrigeringen for tidsserier i detaljhandelomsetningsvolumindeksen (DOI) er at vi må gjette når påskeeffekten starter. Vi legger til grunn at det ikke skjer tidligere enn en uke før påske. Antagelsen gjelder for alle tidsserier. Det viser seg seinere i notatet at denne antagelsen ikke passer for mange tidsserier. En ny metode blir laget for å forbedre denne ulempelen.

Innhold

1	Innledning	5
2	Beskrivelse av tidsserier i DOI	6
3	Beskrivelse av metoden	8
4	Anvendelse av metoden for SNN47.11	9
4.1	Å sjekke om variasjonen i w_1 når forskjellige mengder av datainput blir brukt .	11
4.2	Sammenligning av de estimerte effekter for $w_1 = 5$ og $w_1 = 6$	12
4.3	Sammenligning av sesongjusterte tall i mars og april	13
4.4	Oppsummering av påskekorrigering for SNN47.11	13
5	Anvendelse av metoden for 27 tidsserier i DOI	13
6	Oppsummering	17

Tabeller

1	27 tidsserier i detaljhandelsvolumindeksen	7
2	SNN47.11. Verdiene av σ^{w_1} , AICC og påskeeffekt etter w_1	10
3	SNN47.11. σ^{w_1} og AICC for fire mengder av data.	12
4	SNN47.11. Estimerte verdier ved regARIMA modell for $w_1 = 5$ og 6.	12
5	SNN47.11. Ujusterte og sesongjusterte tall i mars og april med $w_1 = 0, 5, 6$ og 7. .	13
6	w_1 og estimerte verdier av påskeeffektene av 27 tidsserier.	15
7	w_1 dager av 27 tidsserier.	16

Figurer

1	<i>Illustrasjon av påske i 2012</i>	6
2	<i>Vektene av 27 tidsserier.</i>	7
3	<i>Ujusterte tall av totalen og SNN47.11</i>	8
4	<i>Ujusterte tall av SNN47.11 for 12 måneder</i>	8
5	<i>SNN47.11. σ^{w_1} mot w_1.</i>	11
6	<i>SNN47.11. AICC mot w_1.</i>	11
7	<i>SNN47.11. Korrigerte tall for kalendereffekter for 12 måneder for $w_1 = 6$.</i>	11
8	<i>SNN47.11. Ujusterte og sesongjusterte tall i mars og april med $w_1 = 0, 5, 6$ og 7. .</i>	14
9	<i>SNN47.11. σ^{w_1} og AICC etter w_1 med forskjellige data mengder.</i>	18

1 Innledning

Vi sesongjusterer tidsserier i SSB med X-12-ARIMA fra US Census. I dette programmet er det en innebygd rutine for påskekorrigering. Rutinen blir brukt i DOI (detaljhandelomsetningsvolumindeksen) for å korrigere påskeeffekter. Det viser seg at rutinen ikke fungerer bra for norske data, siden X-12-ARIMA er basert på kalenderen i USA, der påske bare er en søndag. Det blir fortsatt påskeeffekt igjen i sesongjusterte tall. For å få en bedre påskekorrigering for norske data laget Leiv Solheim og Dinh en ny metode som er beskrevet i notat 97/73. Påskekorrigering blir utført med den nye metoden, men sesongjustering gjøres fortsatt med X-12-ARIMA. Metoden er basert på antagelsen om at påske påvirker data i tre uker; en uke før påske, påskeuken, og en uke etter påske. Antagelsen gjelder for alle tidsserier. Dette ser urimelig ut, siden tidsserier kan ha forskjellige egenskaper. Det finnes varer som vi handler mer enn vanlig av dagene før påske og det finnes andre varer som ikke er helt viktige for påske. Men vi behandler alle tidsserier helt likt. Vi har forsøkt å forbedre dette med en ny metode, som er beskrevet i notat 2007/43, der tre uker blir erstattet med tre perioder: (i) en periode med w_1 dager før påske; (ii) påskedagene med Skjærtorsdag, Langfredag, påskeaften, 1.påskedag og 2.påskedag (5 dager); (iii) en periode etter påske med w_2 dager. Metoden er mer fleksibel og fungerer mye bedre enn den som vi laget i notat 97/73. Hittil har den vært brukt i DOI, varekomsumprisindeksen (VKI) og produksjonsindeksen (PI). Det er også en ulempe ved denne metoden. Vi må kjenne verdiene av w_1 og w_2 på forhånd for å lage regressorer av effekter av perioder før og etter påske.

Leseren henvises til disse notatene for å lese mer om metodene. Vi oppsummerer her påskekorrigeringen med noen korte setninger. Påsken er en religiøs høytid i Norge, og er det vi kaller bevegelige helligdager. Julen er ikke en bevegelig helligdag siden den faller på samme dato hvert år. Påskesøndag, som vi kaller 1.påskedag kan variere fra 22.mars (tidligst) til 25.april (seinest). US Census har laget en tabell med frekvenser av datoene for påske fra 1600-2099 på internett. Påske påvirker handelen av varer. Vi handler ofte inn varer for påsken mange dager før påske, og handler mye mindre i påskedagene. Påsken forskyver handelen fra påskedagene til dagene før påske. Påskesøndag i 2010 var 4.april, dvs hele påsken fra Skjærtorsdag til andre påskedag var i første uke i april. Dermed handlet vi varer mer enn vanlig i mars for påske i april. Men vi gjorde ikke det i mars 2011 siden påskesøndag var så sent som 24.april. Omsetningen til en matbutikk i mars 2010 økte på grunn av påsken i april mens det ikke ble noen ekstra økning i mars 2011. Dermed ble tallet i mars 2010 justert ned for å kunne sammenligne aktiviteten av handelen i mars 2010 og 2011. Selv om påske og jul er høytider, behandler vi jul som sesongvariasjoner, siden julen er på samme dato hvert år. Vi korrigerer ujusterte tall i mars og april for påskeeffekt, men ikke desembertall for effekten av jul.

Vi bruker en regARIMA modell i X-12-ARIMA (regressions model with ARIMA errors) for å estimere effekter (se ligningen (27) i notat 2007/43 for beskrivelsen). I estimering lager vi regressorer for effektene av påske i perioden før og etter påske og påskedagene, ved å beregne andelene av w_1 og w_2 dager og de 5 påskedagene som faller i mars og april (se ligningene (18)-(21) i notat 2007/43). Siden vi ikke har en metode for å estimere w_1 og w_2 fra data gjetter vi verdier for dem. Vi får dermed ikke alltid rimelige verdier av w_1 og w_2 for alle tidsserier. Vi har valgt $w_1 = 7$ og $w_2 = 0$ for alle tidsserier i DOI siden vi antar at handelen for påske ikke er tidligere enn en uke før påske og at handelen blir normal igjen etter påske. Det ser ut som at $w_1 = 7$ ikke passer for SNN47.52 (som er butikkhandel med jernvarer, fargevarer og glass) og SNN47.53 (som er butikkhandel med tapeter, gulvtepper og gardiner), siden dette ikke er nyttige varer til påske. Testen (som vi skal se seinere i neste avsnitt) viser at handelen med disse varene før påske er helt normal. Dermed, når vi velger $w_1 = 7$, blir ujusterte tall i mars og april korrigert for endringen (økning eller nedgang) av handelen i 7 dager før påske. Dette medfører at mønsteret av korrigererte tall i mars og april ser annerledes ut enn det virkelige mønsteret. Formålet med å korrigere tall er å skape et nytt mønster av tall lik som er det virkelige mønsteret.

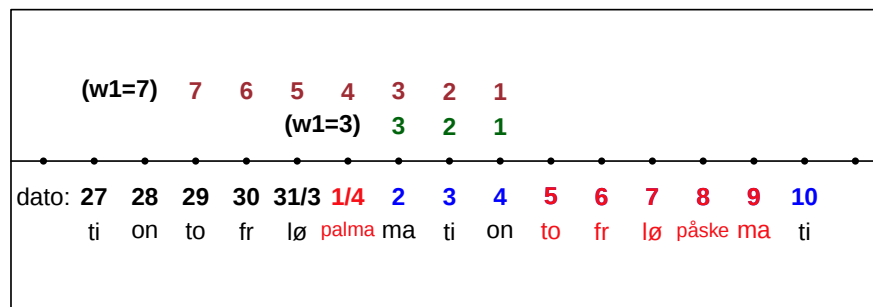
Vi tar et konkret eksempel for å forklare hvordan verdiene av w_1 og w_2 påvirker påskekorrigering. I figur 1 har vi plottet kalenderen i mars og april 2012. Påskedagene varierer fom Skjærtorsdag (5.april) tom 2.påskedag (9.april). Vi antar at vi handler varer mer enn vanlig for påske i tre dager: mandag, tirsdag og onsdag i påskeuken, og at handelen blir normal igjen etter påske. Da blir $w_1 = 3$ og $w_2 = 0$. I 2012 blir de tre dagene i april. Dette betyr at vi handler varer i april for påske i april. Det blir dermed ingen korrigering for påskeeffekt i 2012 for $w_1 = 3$. Hvis vi antar at handelen for påske er i 7 dager før påske, blir denne uken i både mars og april (torsdag 29, fredag 30 og lørdag 31.mars, og mandag 1, tirsdag og onsdag 3.april). Dette betyr at vi handler varer mer enn vanlig i tre dager: tirsdag, fredag og lørdag i mars for påske i april. Dermed blir tall i mars og april korrigert for de tre dagene. Forklaringen gjelder for alle årene.

Vi ser at ved å velge en verdi av w_1 eller w_2 som er langt fra den sanne verdien, blir effektene galt estimert. I begge tilfeller vil det medføre at mønsteret av de korrigerte tallene blir annerledes fra det virkelige mønsteret. Vi får sesongjusterte tall som ikke beskriver aktiviteten i handelen i mars og april helt riktig.

Vi beregner ujusterte tall for totalen, SNN47, ved å ta veiet gjennomsnitt av ujusterte tall av tids-serien på tresifret (SNN47.2 og SNN47.3) og firesifret nivå (SNN47.11, SNN47.19,...). Vi gjør det på samme måte for korrigerte tall med påskeeffekter og sesongjusterte tall. Dermed når vi velger w_1 eller w_2 som ikke er rimelige for mange tidsserier, påvirker dette kvaliteten på den korrigerte tall totalen.

I dette notatet presenterer vi en enkel metode for å finne verdien til w_1 for tidsserier i DOI, uten å estimere den fra data ved en modell. Siden vi antar at handelen blir normal igjen etter påske, blir $w_2 = 0$. I eksemplet velger vi en tidsserie blant 27 tidsserier i DOI for å illustrere metoden. Vi skal bruke figurer til å forklare.

Figur 1: Illustrasjon av påske i 2012



2 Beskrivelse av tidsserier i DOI

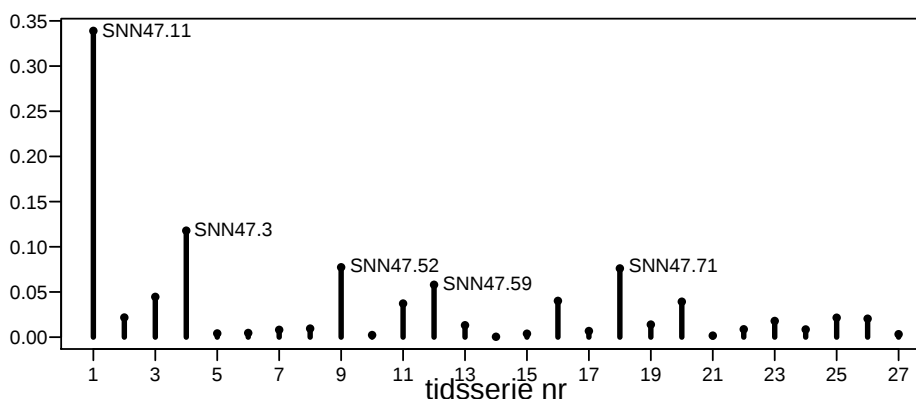
Det er 27 tidsserier på tre og firesifret nivå som skal sesongjusteres hver måned. Generelt sett er det trend og sesongvarisjoner i de fleste tidsseriene. Vi beregner totalen av to sifret (SNN47) ved å ta veiet gjennomsnitt av tre- og firesifret nivå. Forklaringene og vektene er beskrevet i tabell 1. Vektene er plottet i figur 2. Vi ser at SNN47.11 som er butikkhandel med vareutvalg med hovedvekt på nærings- og nytelsesmidler, har den største vekten på 0.34, som er en tredjedel av summen. Dermed bidrar SNN47.11 mest i totalen. Vi skal bruke denne tidsserien for å teste metoden som skal presenteres i neste avsnitt.

I figur 3 plottes vi ujusterte tall av totalen og SNN47.11. Dette viser en stigning av trender hvert år.

Tabell 1: 27 tidsserier i detaljhandelsvolumindeksen

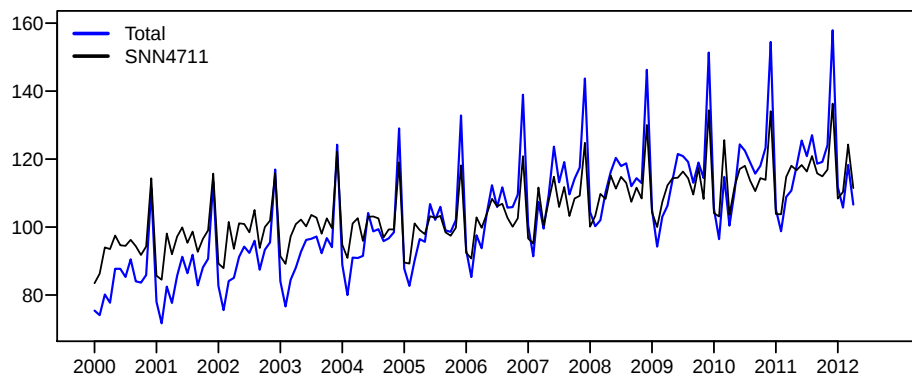
nr	SNN	Vekt	Butikkhandel med
1	47.11	0.338933	bredt vareutvalg med hovedvekt på nærings- og nytelsesmidler
2	47.19	0.021731	bredt vareutvalg ellers
3	47.2	0.044591	nærings- og nytelsesmidler i spesialforretninger
4	47.3	0.117857	Detaljhandel med drivstoff til motorvogner
5	47.41	0.004120	datamaskiner og utstyr til datamaskiner
6	47.42	0.004637	telekommunikasjonsutstyr
7	47.43	0.008093	audio- og videoutstyr
8	47.51	0.009498	tekstiler og utstyrsvarer
9	47.52	0.077352	jernvarer, fargevarer og glass
10	47.53	0.002342	tapeter, gulvtepper og gardiner
11	47.54	0.037212	elektriske husholdningsapparater
12	47.59	0.057969	møbler, belysningsutstyr og andre innredningsartikler
13	47.61	0.013194	bøker
14	47.62	0.000439	aviser og papirvarer
15	47.63	0.003821	innspillinger av musikk og video
16	47.64	0.040139	sportsutstyr og fritidsbåter
17	47.65	0.006765	spill og leker
18	47.71	0.076049	klær
19	47.72	0.013937	skotøy og lærvarer
20	47.73	0.039230	apotekvarer
21	47.74	0.001701	medisinske og ortopediske artikler
22	47.75	0.008715	kosmetikk og toalettartikler
23	47.76	0.017910	blomster og planter, kjæledyr og fôrvarer til kjæledyr
24	47.77	0.008523	ur, gull- og sølvvarer
25	47.78	0.021526	Annen butikkhandel med andre nye varer i spesialforretninger
26	47.91	0.020431	Postordrehandel og handel via internett
27	47.99	0.003272	Detaljhandel utenom utsalgssted ellers

Figur 2: Vektene av 27 tidsserier.

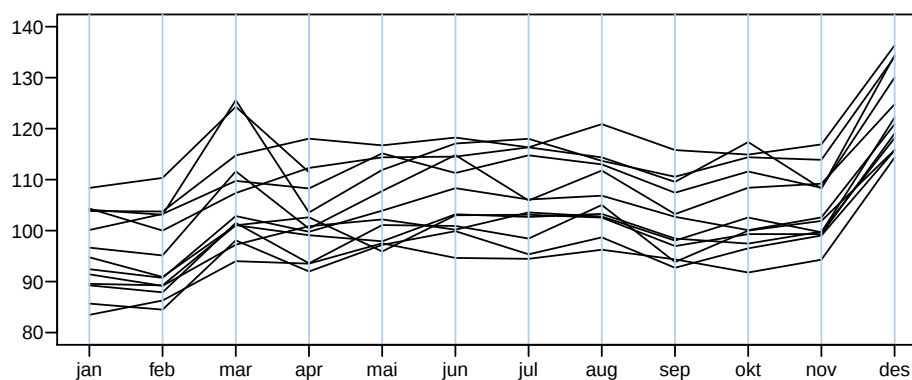


Mønstrene av to tidsserier er ganske like med lave tall i januar, februar og veldige høye tall i desember. I figur 4 er ujusterte tall av SNN47.11 for tolv måneder for 2000-2012 plottet oppå hverandre. På

Figur 3: Ujusterte tall av totalen og SNN47.11



Figur 4: Ujusterte tall av SNN47.11 for 12 måneder



denne måten blir det lettere å se sesongvariasjoner i tidsserien og endringer av ujusterte tall i samme tidspunkt fra år til år enn i figur 3. Det viser et sesongmønster, men det er store variasjoner, slik at det i tillegg må være andre faktorer, som effektene av ukedager eller påske, også har påvirket ujusterte tall. Denne påstanden sjekkes i neste avsnitt av analysen.

3 Beskrivelse av metoden

Vi lager en metode for å finne verdier for w_1 og w_2 . Siden handelen av varer blir normal igjen etter påske, er $w_2 = 0$. Vi har testet nullhypotesen om at $w_2 = 0$ for tidsserier i DOI. Det viser seg at nullhypotesen ikke blir forkastet på 5% nivå. Vi lar $w_2 = 0$ for alle tidsserier i DOI. Siden w_1 ikke blir estimert fra data ved en modell, lager vi et kriterium for å finne en verdi for denne parameteren.

Kriterium. Verdien skal velges slik at avvik mellom korrigerte tall for mars og april blir mest mulig like.

Metode. Vi antar at w_1 kan få verdiene: 0, 1..., W . Når $w_1 = 0$ er det ingen dag før påske (dvs, før Skjærtorsdag) som vi handler varer mer enn vanlig. Når $w_1 = W$ er det en periode med W dager før påske som vi handler varer mer enn vanlig. Mao, påskehandelen skal ikke skje tidligere enn W dager før påske.

Metoden går ut på å gjenta en prosess for å estimere effekter og lage korrigerte tall for en gitt w_1 . Når w_1 varierer fra 0 til W får vi $(W + 1)$ tidsserier av korrigerte tall. Vi plukker ut den tidsserien

som er mest stabile i mars og april og den tilsvarende verdien av w_1 .

Detaljert beskrivelse av metoden. Vi starter med $w_1 = 0$. Siden vi ikke bruker den innebygde rutinen for påskekorrigering i X-12-ARIMA, lager vi selv regressorer for effekter. Vi estimerer ikke bare påskeeffekter, men også effektene av pinse og ukedager med hensyn til røde dager som 1.januar, 1.mai og 17.mai. Det er en regressor for påske, en for pinse og seks for ukedager. La $K_{mar,i}^{w_1=0}$ og $K_{apr,i}^{w_1=0}$ være de korrigerte verdiene i mars og april år i , der $i = 1, \dots, I$, der I er antall år i tidsserien som har både mars og april. Merk at det ikke er nødvendig å beregne $K_{mar,i}^{w_1=0}$ og $K_{apr,i}^{w_1=0}$ fra modellen. Tallene kan hentes fra tabell B.1 i X-12-ARIMA. Forskjellen mellom korrigerte tall i mars og april i år i , er

$$\Delta K_i^{w_1=0} = K_{mar,i}^{w_1=0} - K_{apr,i}^{w_1=0}$$

For alle årene har vi

$$\Delta K_1^{w_1=0}, \Delta K_2^{w_1=0}, \dots, \Delta K_I^{w_1=0}$$

Vi måler spredningen av disse tallene med standaravviket

$$\sigma^{w_1=0} = \text{sd}(\Delta K_1^{w_1=0}, \Delta K_2^{w_1=0}, \dots, \Delta K_I^{w_1=0})$$

Når $\sigma^{w_1=0}$ er lav, er mønsteret av korrigerte tall fra mars til april stabilt.

Prosessen gjentas med $w_1 = 1, \dots, W$. Vi får en rekke med størrelsene $(\sigma^{w_1=0}, \sigma^{w_1=1}, \dots, \sigma^{w_1=W})$. Vi plukker ut den laveste verdien av σ^{w_1} med en tilsvarende verdi av w_1 . Det er denne verdien av w_1 som vi vil finne.

4 Anvendelse av metoden for SNN47.11

Vi velger denne tidsserien som et eksempel for anvendelse av metoden siden det er den viktigste tidsserien i DOI. I neste avsnitt bruker vi metoden for å finne w_1 for alle tidsserier, men vi lister bare ut resultater. SNN47.11 er butikkhandel med bredt vareutvalg med hovedvekt på nærings- og nyttelemidler. Handelen blir påvirket av påske med en veldig høy økning før påske og kraftig nedgang i påskedagene. Vi skal finne hvor mange dager før påske som vi handler varer for påske. Tallene er observert fra januar 2000 til april 2012.

Vi begrenser verdien av w_1 ved å anta at handelen av varer for påske ikke er tidligere enn to uker før påske (mao, før Skjærtorsdag). Dermed blir w_1 i intervallet $[0-15]$. Det ser rimelig ut med denne antagelsen siden hele påsken varer i fem dager, men vi kan handle nyttige varer på kort tid på påskeaften. Det er ikke nødvendig å handle varer mange uker før.

Vi vil ikke korrigere påskeeffekter alene fra ujusterte tall, siden tidsserien også blir påvirket kraftig av ukedageffekter (det er effekter som skyldes ulikt antall ukedager i måned). Hvis vi gjør det slik blir det effekter av ukedager igjen i korrigerte tall, noe som gjør at det blir vanskelig å si noe nøyaktig om variasjoner av korrigerte tall i mars og april. Det er heller ikke mulig å se et klart mønster av korrigerte tall i mars og april. Vi beregner påskeeffekten sammen med effektene av pinse og ukedager med hensyn til røde dager som 1.januar, 1.mai og 17.mai. Vi har laget et program i R for å lage regressorer av disse effektene, og det blir brukt i modellen nedenfor

$$O_t = \lambda LY_t + \sum_{j=1}^6 \beta_j X_{jt} + \sum_i \alpha^i E_t^i + \sum_k \gamma_k Y_{kt} + z_t \quad (1)$$

I ligningen er X_{jt} , E_t^i og Y_{kt} regressorer av henholdsvis ukedager, bevegelige helligdager og ekstremverdier. LY_t er regressor for skuddår. Restleddet z_t beskrives ved en ARIMA modell,

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D z_t = \theta(B)\Theta(B^s)a_t \quad (2)$$

Tabell 2: *SNN47.11. Verdiene av σ^{w_1} , AICC og påskeeffekt etter w_1 .*

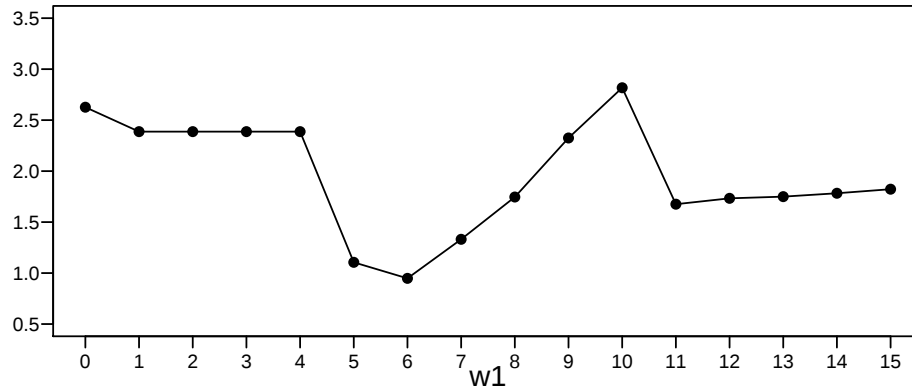
w_1	σ^{w_1}	AICC	før påske		påske	
			$\hat{\alpha}^{fpk}$	t -verdi	$\hat{\alpha}^{pk}$	t -verdi
0	2.6265	541.11	-	-	0.0232	4.25
1	2.3870	538.04	0.1021	12.97	-0.0838	-8.86
2	2.3870	538.04	0.1021	12.97	-0.0838	-8.86
3	2.3870	538.04	0.1021	12.97	-0.0838	-8.86
4	2.3870	538.04	0.1021	12.97	-0.0838	-8.86
5	1.1052	514.19	0.1099	15.81	-0.0880	-10.78
6	0.9479	511.66	0.1074	15.73	-0.0831	-10.47
7	1.3310	516.58	0.1029	14.92	-0.0771	-9.66
8	1.7458	513.56	0.0990	14.25	-0.0713	-9.02
9	2.3248	537.45	0.0962	12.40	-0.0645	-7.55
10	2.8175	549.12	0.0914	11.11	-0.0580	-6.52
11	1.6753	531.11	0.0469	5.18	-0.0154	-1.71
12	1.7329	532.23	0.0441	5.04	-0.0116	-1.35
13	1.7495	531.72	0.0442	5.12	-0.0099	-1.21
14	1.7826	531.74	0.0435	5.12	-0.0078	-1.00
15	1.8224	532.09	0.0423	5.08	-0.0058	-0.76

der a_t er hvit støy prosess.

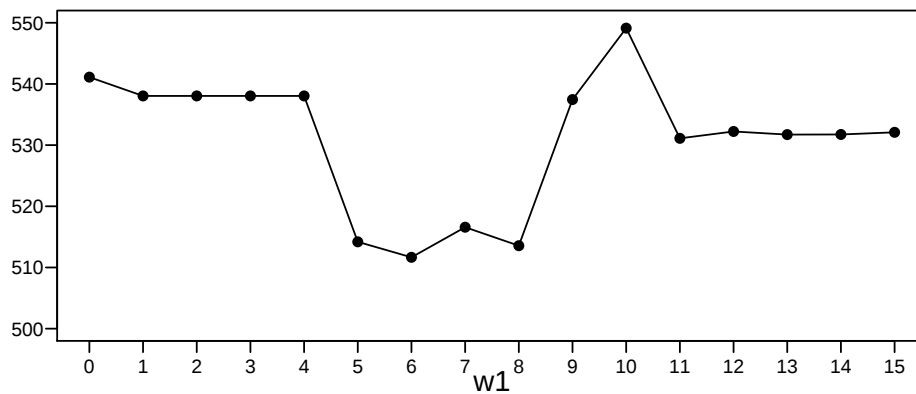
I tabell 2 lister vi ut verdiene av σ^{w_1} og AICC (Akaike's information criterion corrected, for å velge en ARIMA modell) for w_1 fra 0 til 15 sammen med de estimerte verdiene av effektene $\hat{\alpha}^{fpk}$ for handelen før påske og $\hat{\alpha}^{pk}$ for handelen i påskedagene. Den samme ARIMA(0,1,1)(0,1,1) blir valgt for alle w_1 . Vi får følgende kommentarer

1. σ^{w_1} er lavest med $w_1 = 6$. Dermed blir denne verdien valgt for påskekorrigering, som betyr at vi handler varer for påske ikke tidligere enn seks dager før påske. AICC er også lavest med $w_1 = 6$. Merk at vi ikke får verdier for $\hat{\alpha}^{fpk}$ for $w_1 = 0$, siden det er effekten av $w_1 > 0$.
2. Effekten av påske, $\hat{\alpha}^{pk}$, blir positiv for $w_1 = 0$. Da blir det en økning på handelen i påske. Dette ser urimelig ut siden de fleste butikker er stengt i denne perioden. Handelen blir kraftig redusert. $\hat{\alpha}^{pk}$ må være negativ. Dermed blir $w_1 = 0$ ikke en tilpassende verdi. Vi plotter i figur 5 og 6 σ^{w_1} og AICC mot w_1 . De utvikler seg i samme retning.
3. For de andre verdiene av w_1 med $1 \leq w_1 \leq 10$ blir det en høy positiv effekt for handelen før påske og en høy negativ effekt i påske. Dette ser rimelig ut. Dette viser en stor økning og en kraftig nedgang på handelen før og i påske.
4. For w_1 med $11 \leq w_1 \leq 15$ får vi fremdeles en positiv effekt før påske og en negativ effekt i påske, men de blir mye mindre og har en tendens til å gå mot 0. Vi vil begrense w_1 med $1 \leq w_1 \leq 10$ for neste kjøring når nye data tilføyes.
5. Korrigerte tall for $w_1 = 6$ for tolv måneder blir plottet i figur 7. I sammenligning med figur 4, som er ujusterte tall for tolv måneder er mønsteret av de korrigerte tallene mye mer stabilt.
6. AICC fungerer godt slik som σ^{w_1} ved å finne en verdi av w_1 for påskekorrigering.

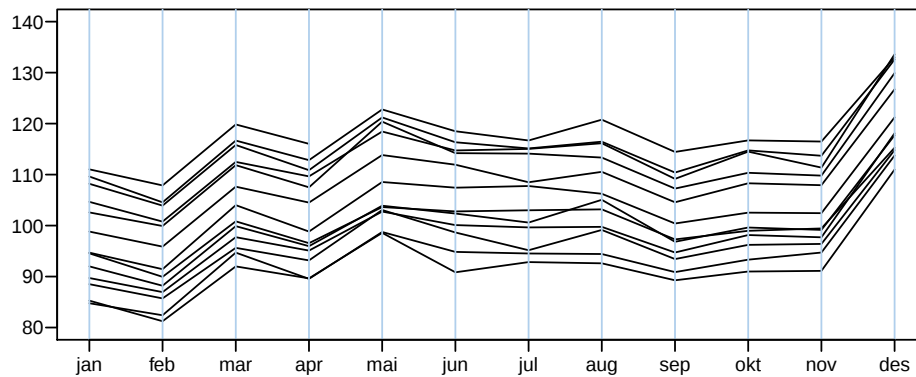
Figur 5: *SNN47.11. σ^{w_1} mot w_1 .*



Figur 6: *SNN47.11. AICC mot w_1 .*



Figur 7: *SNN47.11. Korrigerte tall for kalendereffekter for 12 måneder for $w_1 = 6$.*



4.1 Å sjekke om variasjonen i w_1 når forskjellige mengder av datainput blir brukt

De estimerte verdiene i tabell 2 er basert på datainput fra januar 2000 til april 2012. Da er σ^{w_1} lavest med $w_1 = 6$. Vi vil sjekke at verdien av w_1 ikke blir endret mye med nye mengder av datainput. Vi tester dette med fire nye mengder [januar.2000-desember.2008], [januar.2000-desember.2009], [januar.2000-desember.2010] og [januar.2000-desember.2011].

Tabell 3: *SNN47.11. σ^{w_1} og AICC for fire mengder av data.*

w_1	2000-2008		2000-2009		2000-2010		2000-2011	
	σ^{w_1}	AICC	σ^{w_1}	AICC	σ^{w_1}	AICC	σ^{w_1}	AICC
0	2.4594	367.78	2.4201	411.29	2.2457	445.39	2.2326	486.77
1	2.2468	354.59	2.0061	399.44	1.8886	432.06	1.8693	471.86
2	2.2468	354.59	2.0061	399.44	1.8886	432.06	1.8693	471.86
3	2.2468	354.59	2.0061	399.44	1.8886	432.06	1.8693	471.86
4	2.2468	354.59	2.0061	399.44	1.8886	432.06	1.8693	471.86
5	1.0272	346.07	0.9500	387.63	1.0530	429.81	1.0147	456.76
6	1.1078	337.72	1.0539	379.18	1.0257	418.63	0.9899	456.41
7	1.2942	349.65	1.2214	390.82	1.3413	422.85	1.2912	460.48
8	1.4165	351.98	1.2973	392.96	1.7422	428.70	1.6922	466.62
9	1.5472	354.36	1.5379	395.75	2.2954	439.35	2.2602	478.23
10	1.6450	356.18	1.7246	398.18	1.9032	449.83	2.7875	490.23

I tabell 3, lister vi ut, σ^{w_1} og AICC for fire mengder med begrensningen $1 \leq w_1 \leq 10$. Ved å basere på σ^{w_1} er $w_1 = 5$ valgt for de to første mengdene og $w_1 = 6$ for de to siste. Ved å basere på AICC er $w_1 = 6$ valgt. Vi ser at w_1 varierer mellom 5 og 6. Disse to verdiene er gode kandidater for påskekorrigering for SNN47.11.

Vi plotter σ^{w_1} og AICC for fire mengder av datainput i figur 9 for å se deres variasjoner etter w_1 . Merk at AICC er mye større enn σ^{w_1} . For å kunne plote dem i samme figurer justerer vi skalaen av AICC slik som skalaen av σ^{w_1} . Størrelsen av AICC blir endret, men mønsteret er akkurat som før. Vi ser at mønstrene deres er ganske like.

4.2 Sammenligning av de estimerte effekter for $w_1 = 5$ og $w_1 = 6$

Vi korrigerer ujusterte tall for effektene av skuddår, ukedager, påske og pinse når vi velger $w_1 = 5$ og 6, med ligning (1). Resultatene er vist i tabell 4.

Tabell 4: *SNN47.11. Estimerte verdier ved regARIMA modell for $w_1 = 5$ og 6.*

Parameter		$w_1 = 5$			$w_1 = 6$		
		Estimerte	std.avvik	t-verdi	Estimerte	std.avvik	t-verdi
skuddår	$\hat{\lambda}$	0.0326	0.0059	5.51	0.0323	0.0059	5.44
mandag	$\hat{\beta}_1$	-0.0034	0.0020	-1.66	-0.0026	0.0020	-1.26
tirsdag	$\hat{\beta}_2$	0.0044	0.0020	2.19	0.0040	0.0020	1.96
onsdag	$\hat{\beta}_3$	-0.0014	0.0020	-0.67	-0.0006	0.0020	-0.27
torsdag	$\hat{\beta}_4$	0.0026	0.0021	1.20	0.0017	0.0021	0.78
fredag	$\hat{\beta}_5$	0.0108	0.0020	5.29	0.0109	0.0020	5.31
lørdag	$\hat{\beta}_6$	0.0136	0.0021	6.45	0.0123	0.0021	5.79
før påske	$\hat{\alpha}^{fpk}$	0.1099	0.0069	15.81	0.1074	0.0068	15.73
påske	$\hat{\alpha}^{pk}$	-0.0880	0.0081	-10.78	-0.0831	0.0079	-10.47
pinse	$\hat{\alpha}^{pi}$	-0.0064	0.0037	-1.70	-0.0070	0.0037	-1.84

Effektene som er signifikante på 5% nivå, er skuddår $\hat{\lambda}$, ukedagene med tirsdag $\hat{\beta}_2$, fredag $\hat{\beta}_5$, lørdag $\hat{\beta}_6$, påske med før påske $\hat{\alpha}^{fpk}$ og påskedagene $\hat{\alpha}^{pk}$, pinse $\hat{\alpha}^{pi}$. De estimerte verdiene som er beregnet

med $w_1 = 5$ og 6 er ganske like, Vi forventer at de korrigerte og sesongjusterte tallene skal ligge nær hverandre. Vi sjekker dette i neste avsnitt.

4.3 Sammenligning av sesongjusterte tall i mars og april

I sammenligningen tar vi med $w_1 = 7$ siden det er verdien som vi har brukt for påskekorrigering fram til nå. Vi ser at sesongjusterte tall med $w_1 = 5, 6$ og 7 er nær hverandre, men de skiller seg fra $w_1 = 0$ i 2007, 2010 og 2012, der påske er i tidlig april, se figur 8. Vi får en stor endring av ujusterte tall i mars og april 2010, $O_{apr.2010} - O_{mar.2010} = 103.55 - 125.58 = -22.03$. Endringen av sesongjusterte tall med $w_1 = 0$ blir enda større, $A_{apr.2010}^{w_1=0} - A_{mar.2010}^{w_1=0} = 103.88 - 127.22 = -23.33$. Det blir mye mindre med 0.72, -0.25 og -1.39 for $w_1 = 5, 6$ og 7, henholdsvis.

Tabell 5: *SNN47.11. Ujusterte og sesongjusterte tall i mars og april med $w_1 = 0, 5, 6$ og 7.*

år	mars					april				
	ujus.	$w_1 = 0$	$w_1 = 5$	$w_1 = 6$	$w_1 = 7$	ujus.	$w_1 = 0$	$w_1 = 5$	$w_1 = 6$	$w_1 = 7$
2000	94.00	94.18	94.46	94.75	94.80	93.52	95.52	96.09	96.01	96.03
2001	98.09	96.64	97.11	97.57	97.76	91.98	95.77	96.05	95.78	95.67
2002	101.51	101.93	99.43	99.56	99.59	93.59	95.36	98.52	98.51	98.48
2003	97.23	100.17	100.44	100.60	100.60	100.86	101.60	101.88	101.66	101.51
2004	101.04	102.41	102.62	102.74	102.73	102.61	102.56	102.60	102.38	102.24
2005	101.06	100.02	99.73	99.80	99.82	99.10	99.58	100.17	100.12	100.02
2006	102.85	102.93	103.03	103.39	103.56	99.77	102.61	102.36	101.91	101.56
2007	111.63	109.78	107.79	106.89	106.33	100.39	105.38	107.03	107.84	108.27
2008	109.75	111.04	110.80	110.83	110.77	108.27	111.62	111.55	111.32	111.16
2009	107.38	111.68	111.96	112.09	112.12	112.29	114.11	113.28	112.91	112.71
2010	125.58	127.22	114.46	114.86	115.38	103.55	103.88	115.18	114.60	113.98
2011	114.75	115.43	116.04	116.35	116.55	118.03	116.94	116.45	116.14	116.00
2012	124.29	122.02	120.28	119.19	118.56	111.46	117.22	118.84	119.73	120.32

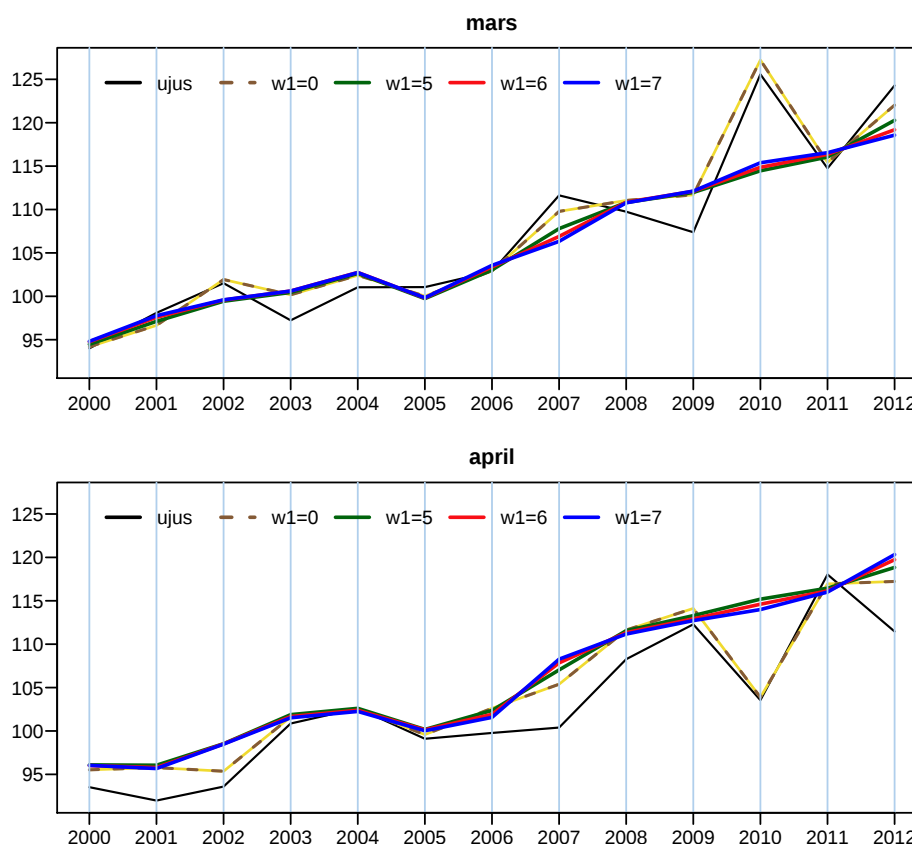
4.4 Oppsummering av påskekorrigering for SNN47.11

SNN47.11 er butikkhandel med bredt vareutvalg med hovedvekt på nærings- og nytelsesmidler. Tidsserien blir påvirket av påske. Handelen blir redusert mye i påskedagene og kraftig økt i mange dager før påske, med 6 dager før påske som det beste anslaget. Etter korrigering får vi tydelige og stabile sesongvariasjoner. Sesongjusterte tall i mars og april 2012 er $A_{mar.2012} = 119.19$ og $A_{apr.2012} = 119.73$. Tallene som vi publiserte i april 2012 er 118.56 og 120.32 som er basert på antagelsen at handelen blir økt i 7 dager før påske.

5 Anvendelse av metoden for 27 tidsserier i DOI

Metoden i avsnitt 3 blir brukt for alle tidsserier. Resultatene blir listet ut i tabell 6. La $\hat{\alpha}^{fpk}$ og $\hat{\alpha}^{pk}$ være effekten av påske på handelen før påske og effekten av handelen i påskedagene, henholdsvis. $\hat{\alpha}^{pk}$ er negative for alle tidsserier. Dette viser at handelen blir redusert i påskedagene. Vi begrenser w_1 med verdiene 0, ..., 10. Tidsseriene som har en liten økning eller nedgang på handelen før påske, er

Figur 8: SNN47.11. Ujusterte og sesongjusterte tall i mars og april med $w_1 = 0, 5, 6$ og 7 .



- SNN47.42, butikkhandel med telekommunikasjonsutstyr. Vi får $w_1 = 7$, $\hat{\alpha}^{fpk} = -0.0597$ med t -verdi -1.15 og $\hat{\alpha}^{pk} = -0.0282$ med t -verdi -0.64 . Tallene viser negative effekter av handelen før påske og i påskedagene. Handelen blir redusert i 7 dager før påske og hele påsken. Tolkningen ser rimelig ut siden det ikke er nyttige varer for påske.
- SNN47.51, butikkhandel med tekstiler og urstørvarsler. Det blir $w_1 = 5$, $\hat{\alpha}^{fpk} = 0.0319$ med t -verdi 1.46 og $\hat{\alpha}^{pk} = -0.0988$ med t -verdi -3.92 . Det viser at handelen har en liten økning før påske og blir redusert i påsken.
- SNN47.52, butikkhandel med jernvarer, fargevarer og glass. Det blir $w_1 = 10$, $\hat{\alpha}^{fpk} = -0.0025$ med t -verdi -0.15 og $\hat{\alpha}^{pk} = -0.1190$ med t -verdi -6.57 . Handelen blir redusert litt i en periode 10 dager før påske og blir kraftig redusert i påske.
- SNN47.53, butikkhandel med tapeter, gulvtepper og gardiner. Vi får $w_1 = 10$, $\hat{\alpha}^{fpk} = 0.0176$ med t -verdi 0.57 og $\hat{\alpha}^{pk} = -0.0725$ og t -verdi -2.18 . Det er nesten ingen økning av handelen før påske, men den blir redusert i påskedagene.
- SNN47.62, butikkhandel med aviser og papirvarer. $w_1 = 10$, $\hat{\alpha}^{fpk} = 0.0165$ med t -verdi 0.72 og $\hat{\alpha}^{pk} = -0.1669$ med t -verdi -6.12 . Handelen øker litt i 10 dager før påske og blir kraftig redusert i påske.
- SNN47.71, butikkhandel med klær. $w_1 = 0$, $\hat{\alpha}^{fpk} = 0$ og $\hat{\alpha}^{pk} = -0.1410$ med t -verdi -12.84 . Det blir ingen økning av handelen før påske, men den blir kraftig redusert i påsken.
- SNN47.72, butikkhandel med skotøy og lærvarer. $w_1 = 1$, $\hat{\alpha}^{fpk} = -0.0398$ med t -verdi -1.13 og $\hat{\alpha}^{pk} = -0.1098$ med t -verdi -2.62 . Handelen blir redusert i påsken og bare en dag før

Tabell 6: w_1 og estimerte verdier av påskeeffektene av 27 tidsserier.

Tidsserie	før påske			påske	
	w_1	$\hat{\alpha}^{fpk}$	t -verdi	$\hat{\alpha}^{pk}$	t -verdi
SNN4711	6	0.1074	15.73	-0.0831	-10.47
SNN4719	9	0.1512	7.83	-0.0991	-4.70
SNN472	8	0.1702	14.78	-0.0903	-7.15
SNN473	6	0.0238	2.45	-0.0130	-1.15
SNN4741	7	0.0704	2.86	-0.1322	-4.52
SNN4742	7	-0.0597	-1.55	-0.0282	-0.64
SNN4743	6	0.0509	2.34	-0.1012	-3.99
SNN4751	5	0.0319	1.46	-0.0988	-3.92
SNN4752	10	-0.0025	-0.15	-0.1190	-6.57
SNN4753	10	0.0176	0.57	-0.0725	-2.18
SNN4754	10	0.0535	2.76	-0.0910	-4.33
SNN4759	1	0.0425	2.95	-0.1525	-8.75
SNN4761	5	0.1151	3.76	-0.2571	-6.92
SNN4762	10	0.0165	0.72	-0.1669	-6.12
SNN4763	5	0.1207	4.54	-0.1446	-4.63
SNN4764	1	0.0703	2.66	-0.0896	-2.82
SNN4765	7	0.0605	1.76	-0.1859	-4.43
SNN4771	0	—	—	-0.1410	-12.84
SNN4772	1	-0.0398	-1.13	-0.1098	-2.62
SNN4773	10	0.0646	8.08	-0.1023	-11.67
SNN4774	7	-0.0093	-0.37	-0.0730	-2.53
SNN4775	0	—	—	-0.0634	-7.50
SNN4776	0	—	—	-0.0382	-2.50
SNN4777	0	—	—	-0.1162	-11.12
SNN4778	8	0.0306	1.64	-0.1119	-5.29
SNN4791	10	0.0026	0.12	-0.1149	-4.63
SNN4799	10	0.1563	2.93	-0.2969	-5.17

påske.

- SNN47.74, butikkhandel med medisinske og ortopediske artikler. $w_1 = 7$, $\hat{\alpha}^{fpk} = -0.0093$ med t -verdi -0.37 og $\hat{\alpha}^{pk} = -0.0730$ med t -verdi -2.53 . Handelen blir som normal før påske og redusert i påskedagene.
- SNN47.75, butikkhandel med kosmetikk og toalettartikler. $w_1 = 0$, $\hat{\alpha}^{fpk} = 0$ og $\hat{\alpha}^{pk} = -0.0634$ med t -verdi -7.50 . Det blir ingen økning av handelen før påske, men den blir kraftig redusert i påske.
- SNN47.76, butikkhandel med blomster og planter, kjæledyr og og fôrvarer til kjæledyr. $w_1 = 0$, $\hat{\alpha}^{fpk} = 0$ og $\hat{\alpha}^{pk} = -0.0382$ med t -verdi -2.50 . Handelen blir som normal fram til påske og kraftig redusert i påskedagene.
- SNN47.77, butikkhandel med ur, gull- og sølvvarer. $w_1 = 0$, $\hat{\alpha}^{fpk} = 0$ og $\hat{\alpha}^{pk} = -0.1162$ med t -verdi -11.12 . Handelen før påske blir normal og kraftig redusert i påske.
- SNN47.78, annen butikkhandel med andre nye varer i spesialforretninger. $w_1 = 8$, $\hat{\alpha}^{fpk} = 0.0306$ med t -verdi 1.64 og $\hat{\alpha}^{pk} = -0.1119$ med t -verdi -5.29 . Handelen øker i 8 dager før

påske og reduseres i påskedagene.

- SNN47.91, postordrehandel og handel via internett. $w_1 = 10$, $\hat{\alpha}^{fpk} = 0.0026$ med t -verdi 0.12 og $\hat{\alpha}^{pk} = -0.1149$ med t -verdi -4.63 . Det er nesten ingen økning av handelen før påske men den blir redusert i påskedagene.

De tidsseriene som har en kraftig økning på handelen før påske og en sterk nedgang i påskedagene, er

- SNN47.11, butikkhandel med bredt vareutvalg med hovedvekt på nærings- og nytelsemidler.
- SNN47.19, butikkhandel med bredt vareutvalg ellers.
- SNN47.2, butikkhandel med nærings- og nytelsesmidler i spesialforretninger.
- SNN47.61, butikkhandel med bøker.
- SNN47.63, butikkhandel med innspillinger av musikk og video.
- SNN47.99, detaljhandel utenom utsalgssted ellers.

De tidsseriene som har mindre økning på handelen før påske, er

- SNN47.3, detaljhandel med drivstoff til motorvogner.
- SNN47.41, butikkhandel med datamaskiner og utstyr til datamaskiner.
- SNN47.43, butikkhandel med audio- og videoutstyr.
- SNN47.54, butikkhandel med elektriske husholdningsapparater.
- SNN47.59, butikkhandel med møbler, belysningsutstyr og andre innretningsartikler.
- SNN47.64, butikkhandel med sportsutstyr og fritidsbåter.
- SNN47.65, butikkhandel med spill og leker. Vi får lav t -verdi, $t = 1.76$.
- SNN47.73, butikkhandel med apotekvarer.

Siden tidsseriene SNN47.52, SNN47.74 og SNN47.91 har veldig lave effekter av handelen før påske, $\hat{\alpha}^{fpk} \approx 0$ vil vi la dem være lik 0. Dermed blir $w_1 = 0$. De endelige verdiene av w_1 for 27 tidsserier er gitt i tabell 7 og brukes for å lage regressorer for påskeskorrigerings.

Tabell 7: w_1 dager av 27 tidsserier.

SNN	w_1	SNN	w_1	SNN	w_1	SNN	w_1	SNN	w_1	SNN	w_1
4711	6	4742	7	4754	10	4764	1	4774	0	4791	0
4719	9	4743	6	4759	1	4765	7	4775	0	4799	10
472	8	4751	5	4761	5	4771	0	4776	0		
473	6	4752	0	4762	10	4772	1	4777	0		
4741	7	4753	10	4763	5	4773	10	4778	8		

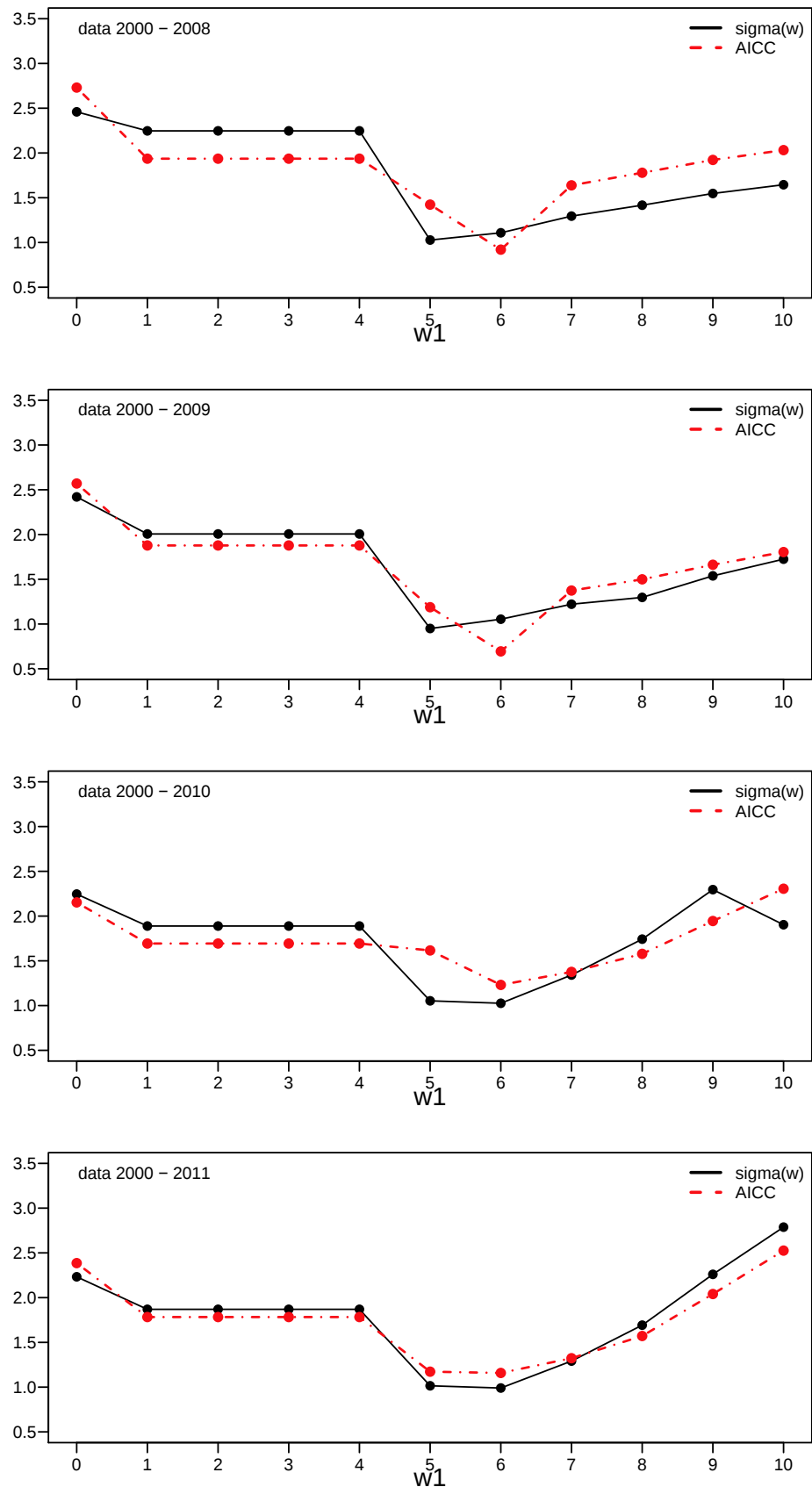
Vi sesongjusterer 27 tidsserier i tabell 1 med w_1 i tabell 7 for påskeskorrigerings. Vi beregner sesongjusterte tall av totalen ved å summere opp sesongjusterte tall av 27 tidsserier med vektorer. Det gir $A_{mar.2012}^{w_1} = 124.25$ og $A_{apr.2012}^{w_1} = 123.60$ mot $A_{mar.2012}^{w_1=7} = 124.09$ og $A_{apr.2012}^{w_1=7} = 123.81$ som er sesongjusterte tall av totalen med $w_1 = 7$ for alle tidsserier. Når vi antar at $w_1 = 0$ blir $A_{mar.2012}^{w_1=0} = 125.83$ og $A_{apr.2012}^{w_1=0} = 121.95$, med et avvik -3.87 .

6 Oppsummering

Vi har tidligere skrevet et notat om påskekorrigering. En forbedring som innføres i dette notatet, er at verdien av w_1 kan finnes ved en enkel metode. Metoden ble anvendt for SNN47.11, den tidsserien som bidrar mest til totalen. Der blir $w_1 = 6$ valgt, og $w_1 = 5$ er også en god kandidat. Vi finner disse verdiene igjen ved forskjellige mengder av datainput. Etter korrigering for kalendereffekter med $w_1 = 6$ for påskekorrigering, blir sesongmønsteret i mars og april ganske stabilt. Metoden ble også brukt til å finne w_1 av 27 andre tidsserier. De er listet ut i tabell 6. Verdiene ser rimelige ut. Det er mange tidsserier med $w_1 = 0$, noe som tyder på at handelen før påske er helt normal fram til påske. Det er også mange tidsserier med $w_1 > 0$ som viser at det er en endring av handelsmønsteret før påske. Når effekten er positiv blir det en oppgang og når effekten er negativ blir det en nedgang av handelen før påske. For eksempel blir handelen med telekommunikasjonsutyr (SNN47.42) redusert ikke bare i påsken, men også i uken før påske.

Vi sesongjusterer 27 tidsserier på 3- og 4-sifret nivå med $w_1 = 0$ og 7 for alle, og w_1 -ene i tabell 7. Sesongjusterte tall av totalen er summen av sesongjusterte tall av tre- og firesifret med vektor. Det viser at avviket mellom sesongjusterte tall av totalen i mars og april 2012 er stort når vi ikke korrigerer for effekten av handelen før påske, men bare i påskedagene (dvs $w_1 = 0$). Ellers blir det mye mindre.

Figur 9: $SNN47.11$. σ^{w_1} og AICC etter w_1 med forskjellige data mengder.



Referanser

- [1] Bell W. R. og Hillmer S. C. (1983): *Modelling Time Series with Calendar Variation*, Journal of the American Statistical Association, September 1983, Vol. 78.
- [2] Bureau of the Census: *X-12 ARIMA Reference Manual, Version 0.2.10, July 26, 2002*.
- [3] Cleveland, W. S. and S. J. Devlin (1980): *Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods*, Journal of the American Statistical Association, 75, pp. 487-496.
- [4] Dagum, E. B., Quenneville, B. and Sutradhar, B. (1992): *Trading-Day Variations Multiple Regression Models with Random Parameters*, International Statistical Review, 60, 57-73.
- [5] Dinh Quang Pham (2007) *Ny metode for påskekorrigering*, Notater 2007/43.
- [6] Findley, D. F. Monsell, B. C., Bell, W. R., Otto, M. C. og Chen B. C. (1998): *New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program*. Journal of Business and Economic Statistics, 16, 127-177.
- [7] Leiv Solheim og Dinh Quang Pham (1997): *Prekorrigering av påskeeffekten for detaljvolumindeksen 1979-1997*. Notater 73/97.
- [8] Pierce, D.A. (1971a), *Least Squares Estimation in the Regression Model With Autoregressive-Moving Average Errors*, Biometrika, 58, 299-312.
- [9] Tsay, R. S. (1984): *Regression Models with Time Series Errors*, Journal of the American Statistical Association, Vol. 79, No. 385 (Mar., 1984), pp. 118-124.
- [10] William S. Cleveland og Susan J. Devlin (1982): *Calendar Effects in Monthly Time Series: Modeling and Adjustment*. Journal of the American Statistical Association, Vol. 77, No. 379 (Sep., 1982), pp. 520-528.
- [11] Young, A.H (1965): *Estimating Trading Day Variation in Monthly Economic Time Series*, Technical Paper 12, U.S. Department of Commerce, Bureau of the Census, Washington, D.C.

B

Returadresse:
Statistisk sentralbyrå
NO-2225 Kongsvinger

Avsender:

Statistisk sentralbyrå

Postadresse:

Postboks 8131 Dep
NO-0033 Oslo

Besøksadresse:

Kongens gate 6, Oslo
Oterveien 23, Kongsvinger

E-post: ssb@ssb.no

Internett: www.ssb.no

Telefon: 62 88 50 00

ISBN 978-82-537-8567-7 (trykt)

ISBN 978-82-537-8568-4 (elektronisk)

ISSN 1891-5906

ISBN 978-82-537-8567-7



9 788253 780290



Statistisk sentralbyrå
Statistics Norway