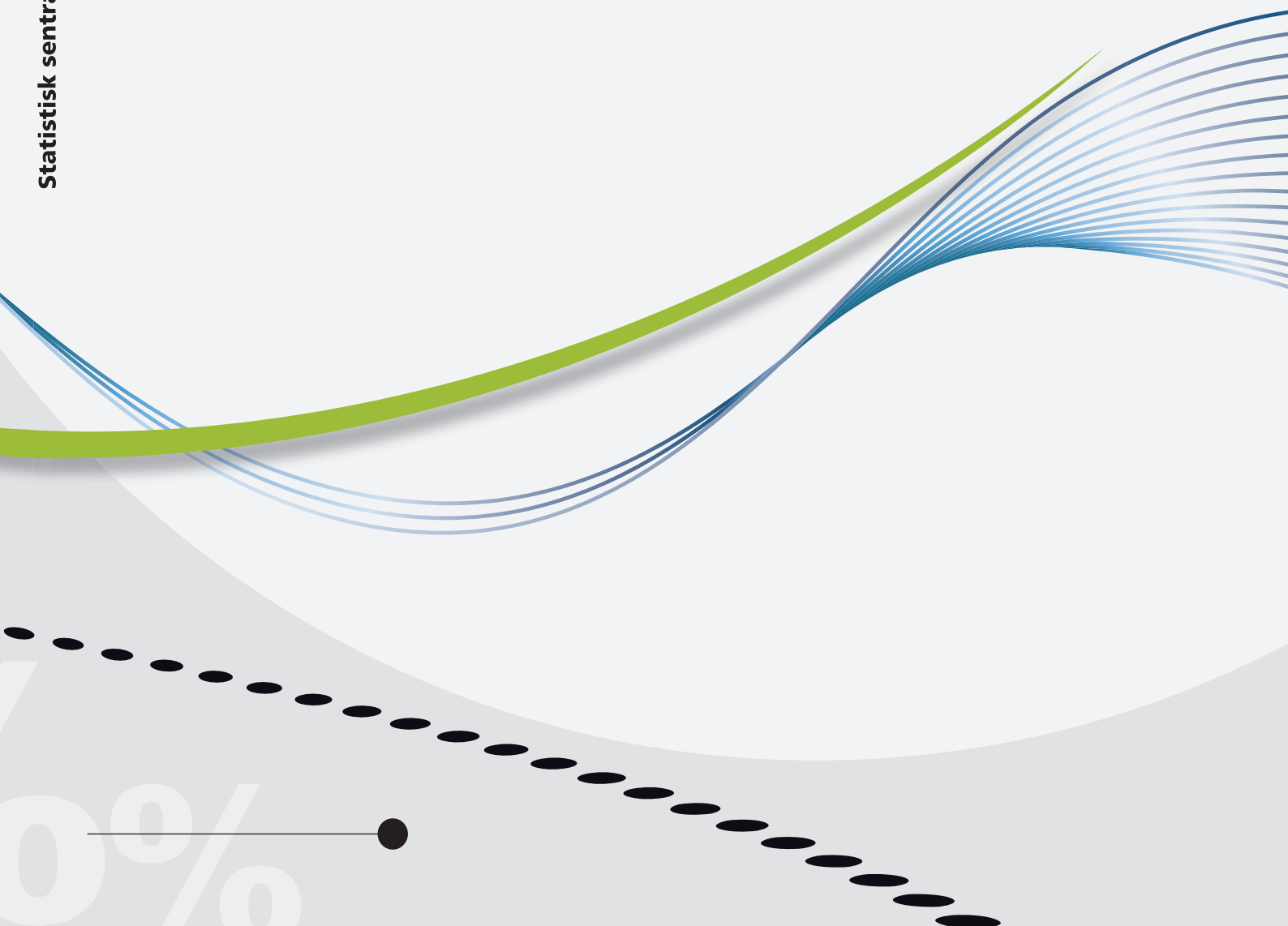


Torunn Heggland

Sesongjustering av statistikken over ledige stillinger



Torunn Heggland

Sesongjustering av statistikken over ledige stillinger

I serien Notater publiseres dokumentasjon, metodebeskrivelser, modellbeskrivelser og standarder.

© Statistisk sentralbyrå
Ved bruk av materiale fra denne publikasjonen skal Statistisk sentralbyrå oppgis som kilde.

Publisert 16. november 2017

ISBN 978-82-537-9623-9 (elektronisk)

Standardtegn i tabeller	Symbol
Tall kan ikke forekomme	.
Oppgave mangler	..
Oppgave mangler foreløpig	...
Tall kan ikke offentligjøres	:
Null	-
Mindre enn 0,5 av den brukte enheten	0
Mindre enn 0,05 av den brukte enheten	0,0
Foreløpig tall	*
Brudd i den loddrette serien	—
Brudd i den vannrette serien	
Desimaltegn	,

Forord

Notatet «Sesongjustering av statistikken over ledige stillinger» gir en oppsummering av utviklingsarbeidet som ble utført i forkant av første publisering (1. kvartal 2017) av sesongjusterte tall og trendfall for denne statistikken. Oppdaterte tall publiseres nå hvert kvartal på <http://www.ssb.no/ledstill>.

Dette er første gang sesongjusteringsverktøyet JDemetra+, som er utarbeidet på vegne av Eurostat og er en anbefalt programvare for sesongjustering, tas i bruk på en statistikk i SSB.

Statistisk sentralbyrå, 13.11.2017

Torstein Bye

Sammendrag

Statistikken over ledige stillinger går tilbake til 1. kvartal 2010, og gir antall ledige stillinger per utgangen av midterste måned i kvartalet, samt andelen ledige stillinger (i prosent). Denne er gitt ved

$$\frac{\text{antall ledige stillinger}}{\text{antall ledige stillinger} + \text{antall ansatte}}$$

Antall ansatte i denne statistikken er ensbetydende med «antall ansatte i virksomheten», slik at en person som er ansatt to steder telles som to ansatte. Statistikken er fordelt etter næringshovedområdene A-S i Standard for næringsgruppering 2007 (SN2007)¹, men der hovedområdene D og E er slått sammen, og Q er delt i tre grupper etter 2-siffer næring. Det er dermed 20 serier for antall ledige stillinger og 20 serier for antall ansatte.

I juni 2017 publiserte vi sesongjusterte serier og trendserier for antall ledige stillinger og for andelen ledige stillinger for første gang. Dette var en utvidelse av statistikken, som frem til dette kun bestod av ujusterte tall. Ved første publiseringstidspunkt hadde vi observasjoner til og med 1. kvartal 2017, og dette notatet er basert på disse dataene.

De 40 tidsseriene som inngår i statistikken er sesongjustert med X-13ARIMA. Vi har benyttet X-13ARIMA-versjonen som ligger i programvaren JDemetra+ (JD+), som er et Windows-grensesnitt for sesongjustering utarbeidet på vegne av Eurostat.

Hovedformålet med notatet er å presentere de valgene som ble tatt i forbindelse med oppstarten av sesongjustering av statistikken. De vurderingene som er gjort på hver serie beskrives, og i den forbindelse blir flere av de diagnostiske testene i JD+ omtalt. Notatet gir også en innføring i sesongjusteringsmetoden, der blant annet begrepene pre-korrigerer, kalender- og virkedagskorrigerer, regARIMA-modellering, trend- og sesongfiltre forklares.

Overgangen til A-ordningen ved 1. kvartal 2015 førte til et brudd i seriene for antall ansatte, og dette diskuteres i notatet. Nivåskiftet som denne overgangen førte til er justert for i de sesongjusterte tallene for ansatte. Dette betyr ikke nødvendigvis at hele bruddet er justert for, da overgangen til A-ordningen også har ført til endringer i sesongmønsteret for noen næringer. Vi har justert for endringer i sesongmønsteret for næringshovedgruppene «A: Jordbruk, skogbruk og fiske» og «P: Undervisning». For enkelte serier vil det ta lengre tid før vi kan konkludere angående endret sesongmønster.

¹ <http://www.ssb.no/klasse/klassifikasjoner/6>

Innhold

Forord	3
Sammendrag	4
1. Innledning	6
1.1. Formål.....	6
1.2. Datakilder.....	6
1.3. Tidsseriene	7
2. Metode	7
2.1. Om metoden	8
2.2. Spesifikasjonsvalg for denne statistikken.....	12
2.3. Håndtering av brudd i seriene	14
3. Sesongjustering av seriene	18
3.1. Vurderinger som er gjort på seriene.....	18
3.2. Ledige stillinger	22
3.3. Ansatte.....	25
3.4. Årlig gjennomgang av seriene.....	29
3.5. Sesongjusterte totaltall.....	29
Referanser	32
Vedlegg A: Tabeller	33

1. Innledning

1.1. Formål

Statistikken over ledige stillinger gir det estimerte antallet ledige stillinger pr utgangen av midterste måned i kvartalet fordelt etter 20 næringsgrupper. Disse er næringshovedgruppene A-S i SN2007, men der D og E er slått sammen, og næringshovedgruppe Q er delt i tre etter 2-siffer næring, og benevnes i notatet som Q86, Q87 og Q88. I tillegg publiseres andelen ledige stillinger (i prosent) som er gitt ved

$$\frac{\text{antall ledige stillinger}}{\text{antall ledige stillinger} + \text{antall ansatte}}$$

Antall ansatte i denne statistikken er ensbetydende med «antall ansatte i virksomheten», slik at en person som er ansatt to steder telles som to ansatte. Andelen ledige stillinger er sammenlignbar mellom land og næringer. Dette i motsetning til antallet ledige stillinger, som selvsagt vil variere avhengig av størrelsen på landet eller næringen.

Tallene rapporteres til Eurostat hvert kvartal på bakgrunn av Europarlaments- og rådsforordning 453/2008. Når serien inneholder minst fire år med data, krever Kommissjonsforordning 1062/2008 at det også rapporteres sesongjusterte tall. Sesongjusterte tall og trendtall gir muligheten til å følge utviklingen i tallet på ledige stillinger fra kvartal til kvartal, i stedet for å sammenligne hvert kvartal med tilsvarende kvartal tidligere år. 4. kvartal 2013 hadde vi i teorien nok data, men sesongjusteringen av en slik kort serie (kortere enn 5 år) kan ofte være ustabil; dvs. at den sesongjusterte serien blir vesentlig endret når serien oppdateres gjennom revisjoner eller ved at nye observasjoner blir lagt til. Dette omtales nærmere i kapittel 7.4 i Office for National Statistics (2007). Sommeren 2016 startet vi å jobbe med å finne ut hvilke serier som egnet seg for sesongjustering, og i juni 2017 publiserte vi sesongjusterte tall for første gang, som en del av publiseringen for 1. kvartal.

Dette notatet tar for seg de metode- og spesifikasjonsvalgene vi tok i denne prosessen, og gir en vurdering av hver enkelt series egnethet for sesongjustering. En gjennomgang av hovedideene i X-13ARIMA metoden blir også gitt.

1.2. Datakilder

For å lage statistikken brukes både en virksomhetsundersøkelse og registerdata.

1.2.1 Utvalgsundersøkelsen om ledige stillinger

Det utføres en kvartalsvis utvalgsundersøkelse der omtrent 8 000 virksomheter deltar. Denne undersøkelsen startet 1. kvartal 2010. Virksomhetene trekkes fra 35 næringsstrata, som igjen er inndelt i fem størrelsesstrata etter antallet ansatte i virksomheten. Skillelinjene for størrelsesstrataene varierer mellom næringsstrataene. For å lette oppgavebyrden for små virksomheter, trekkes ingen i størrelsesstratumet for de minste virksomhetene (1-4 ansatte) ut til å delta, mens alle virksomheter i størrelsesstratumet med de største virksomhetene er med. For de tre midterste størrelsesstrataene er det ulike trekk sannsynligheter for de ulike næringene, der trekk sannsynlighetene hvert år kan bli justert på bakgrunn av variasjonskoeffisientene for det foregående året, og med den hensikt å ende opp med omtrent 8 000 virksomheter i utvalget. Virksomhetene som deltar i undersøkelsen kommer dermed fra $35 \times 4 = 140$ strata, og disse besvarer spørsmålet om hvor mange stillinger som var ledige på et referansetidspunkt ved utgangen av midterste måned i kvartalet. For å blåse tallene opp til populasjonsnivå benyttes en

modell-basert rateestimator innad i hvert stratum, og antall ansatte i virksomhetene er forklaringsvariabel i denne modellen².

1.2.2 Antall ansatte fra Aa-registeret og A-ordningen

Tallene for antall ansatte i virksomhetene ble til og med 4. kvartal 2014 hentet fra Virksomhets- og foretaksregisteret (VoF) i SSB. Dette registeret fikk fram til utgangen av 2014 tall på ansatte fra Aa-registeret (Arbeidstaker- og arbeidsgiverregisteret til NAV), og får fra og med 2015 tall på ansatte fra A-ordningen. Etter overgangen til A-ordningen har ansattetallene i VoF fått et større tidslag enn tidligere, siden vi nå venter en ekstra måned på oppdateringer fra arbeidsgiverne for å få tilstrekkelig kvalitet. Fra og med 1. kvartal 2015 hentes derfor tallene til statistikken direkte fra SSB sin versjon av A-ordningen kalt Leveranseområde 2 (L2), i stedet for at de hentes ut fra VoF, mens andre oppdateringer i VoF angående virksomhetene fremdeles blir benyttet.

1.3. Tidsseriene

Tidsseriene for antall ledige stillinger, videre omtalt som «ledige stillinger», består av 140 delserier, jamfør avsnitt 1.2.1. Dataene aggregeres innad i hvert næringsstratum til de 35 næringsstrataene. Disse 35 gruppene blir slått sammen til de 20 næringsgruppene omtalt under avsnitt 1.1, som statistikken over ledige stillinger for tiden blir publisert etter. Vi har valgt å ta utgangspunkt i disse 20 gruppene også for sesongjusteringen. Vi kunne valgt å betrakte de 35 næringsstrataene hver for seg, men antar at dette kan fange opp mer støy, enn å velge 20 grupper.

Kravet fra Eurostat innebærer å levere sesongjusterte tall for totalen fordelt etter kun 10 næringsgrupper. De 20 gruppene vi sesongjusterer, lar seg aggregere opp til disse 10 gruppene. Vi vurderte direkte sesongjustering av hver av de 10 gruppene, men fant ulikt sesongmønster for noen av delseriene som utgjør en gruppe, og valgte dermed indirekte sesongjustering, der hver av de 20 gruppene er sesongjustert direkte. Også totalen er indirekte sesongjustert, ved at de 20 sesongjusterte seriene er summert. Seriene over «ansatte» behandles på tilsvarende måte.

2. Metode

X-13ARIMA-SEATS igjennom JDemetra+

Praksisen i Statistisk sentralbyrå (SSB) har vært å bruke den ikke-parametriske, filterbaserte sesongjusteringsmetoden X-12ARIMA, der ARIMA betegner en type generelle tidsseriemodeller for framskrivning. De siste årene har en videreutvikling av denne metoden; X-13ARIMA-SEATS³, blitt utviklet av U.S. Census Bureau i samarbeid med Bank of Spain.

X-13ARIMA-SEATS gir muligheten til å velge mellom en bearbeidet versjon av TRAMO/SEATS og X-13ARIMA. Begge disse metodene bruker elementer fra TRAMO⁴ for seleksjon av ARIMA-modell, mens X-12-ARIMA ikke bruker dette. I X-12-ARIMA og X-13ARIMA brukes ARIMA-modeller kun i pre-korrigeringen og for å forlenge serien. I TRAMO/SEATS brukes derimot ARIMA-modeller også for dekomponering av serien, og omtales derfor som en modellbasert metode.

² Se «Undersøkelsen om ledige stillinger», dokumentasjon https://www.ssb.no/a/publikasjoner/pdf/notat_201225/notat_201225.pdf

³ X-13ARIMA-SEATS Reference Manual Version 1.1 <http://www.census.gov/srd/www/x13as/>

⁴ TRAMO er forkortelse for "Time Series Regression with ARIMA Noise, Missing Observations and Outliers"

Vi har valgt å ta i bruk X-13ARIMA, igjennom den åpne programvaren JDemetra+⁵ (JD+), og følger dermed metodepraksisen i SSB. JD+ er et brukervennlig Windows grensesnitt for X-13ARIMA-SEATS, utviklet av Deutsche Bundesbank og National Bank of Belgium på vegne av Eurostat. Ifølge punkt 3.1 i European Statistical System (ESS) sine retningslinjer for sesongjustering er både X-13ARIMA og TRAMO/SEATS anbefalt av Eurostat, men det presiseres at man bør holde seg til en av metodene (ESS, 2015).

Teorien i kapittel 2 og 3 i dette notatet bygger hovedsakelig på Office for National Statistics (2007) og Grudkowska (2017).

2.1. Om metoden

X-13ARIMA består av to steg:

1. Pre-korrigerende for outliers og kalendereffekter samt forlenging av serien ved en regARIMA-modell.
2. Oppspalting av serien i ulike komponenter, derav en sesongkomponent, for så å kunne fjerne denne.

Selve sesongjusteringen foregår altså i det andre steget, og vi vil se nærmere på dette steget først siden sesongjustering er hovedfokuset vårt, mens steg 1 er en klargjøring av seriene. Utgangspunktet for steg 2 er dermed en pre-korrigert serie O_t^* .

Steg 2

Sesongjusteringen i X-13ARIMA bygger på X-11-metoden⁶, utviklet av US Census Bureau på 1960-tallet, der hovedforskjellen mellom de to metodene er at X11 går direkte til steg 2 samt de konsekvenser dette medfører. Dette kommer vi noe inn på senere i delkapittelet.

X-11 er en dekomponeringsrutine, der en serie O_t dekomponeres i tre delserier eller komponenter; trend, sesongkomponent og irregulær komponent:

- Trenden T_t er den langsiktige utviklingen av serien, og den viser dermed ikke variasjonene innad i året.
- Sesongkomponenten S_t viser variasjonene innad i året som gjentar seg omtrent likt år etter år.
- Den irregulære komponenten I_t er «resten»,

slik at de tre komponentene kan multipliseres for å få de observerte verdiene:

$$O_t = T_t \times S_t \times I_t, \quad (2.1.1)$$

evt. legges sammen: $O_t = T_t + S_t + I_t$.

Siden vi i dette steget (steg 2) betrakter den pre-korrigerte serien, vil vi gjøre oppspaltingen $O_t^* = T_t^* \times S_t^* \times I_t^*$ (evt. $O_t^* = T_t^* + S_t^* + I_t^*$), der * markerer at komponenten er pre-korrigert. I avsnittet «I hvilken komponent havner de ulike effektene» kommer vi tilbake til hvor de pre-korrigerte effektene havner etter oppspaltingen for å oppnå relasjonen i (2.1.1).

Beregning av førsteutkast på trenden

⁵ JDemetra+ Reference Manual Version 2.2.

https://ec.europa.eu/eurostat/cros/system/files/jdemetra_reference_manual_version_2.2_0.pdf

⁶ Innføring i metoden gis i "Seasonal Adjustment with the X-11 Method", Ladiray og Quenneville, 2001

Dekomponeringsprosessen starter med beregning av trend, da vi trenger en serie uten trend for å kunne beregne sesongkomponenten. For hvert tidspunkt tas et vektet gjennomsnitt av observasjonene for året som omslutter tidspunktet, der hver av de to endeobservasjonene vektet $\frac{1}{2}$, mens hver av de andre observasjonene har vekt 1. Trendestimatet for 1. kvartal 2011 er dermed gitt ved

$$T_{1k11}^* = \frac{1}{4} \left(\frac{1}{2} O_{3k10}^* + O_{4k10}^* + O_{1k11}^* + O_{2k11}^* + \frac{1}{2} O_{3k11}^* \right).$$

Bruk av dette glidende gjennomsnittet på alle observasjoner i serien vil både jevne ut systematisk sesong og redusere støy. Det glidende gjennomsnittet kalles et trendfilter, siden trenden filtreres ut. Beregning av trendestimatet for 1. og 2. kvartal 2010, og for de to siste kvartalene er kun mulig dersom serien forlenges i begge retninger. Hvordan dette gjøres skal vi se nærmere på i steg 1 nedenfor. X11-metoden bruker derimot usymmetriske filtre for endepunktene, i stedet for å framskrive og tilbakeskrive serien. Vi kan nå dele hver observasjon i serien på sin respektive trendfaktor (evt. trekke fra trenden, dersom modellen er additiv), og står da igjen med en serie bestående av sesong og irregulærhet. Denne serien kalles SI komponenten.

Trendestimatet vi her har omtalt er kun ment som et førsteutkast, da filteret som er brukt bare kan gjenskape en lineær trend, mens trenden i virkeligheten kan anta krumninger.

Beregning av sesongkomponenten

Fra SI komponenten beregnes førsteutkastet på sesongkomponenten. Dette gjøres ved å glatte ut støyen, dvs. irregulærhetene, og oppnås gjennom å ta gjennomsnittet over samme kvartaler i ulike år. Et mulig estimat på sesongfaktoren for 2. kvartal 2014 er dermed gitt ved

$$S_{2k14}^* = \frac{1}{3} (SI_{2k13}^* + SI_{2k14}^* + SI_{2k15}^*).$$

Et tilsvarende gjennomsnitt beregnes for alle tidspunkt i serien. Det benyttes da et 3×1 ledds glidende gjennomsnitt, som vi omtaler som et sesongfilter. Et eksempel på et lengre sesongfilter er 3×3 ledd filteret gitt ved

$$S_{2k14}^* = \frac{1}{3} \left(\frac{1}{3} (SI_{2k12}^* + SI_{2k13}^* + SI_{2k14}^*) + \frac{1}{3} (SI_{2k13}^* + SI_{2k14}^* + SI_{2k15}^*) + \frac{1}{3} (SI_{2k14}^* + SI_{2k15}^* + SI_{2k16}^*) \right).$$

Et langt filter vil gi lavere vekt til den enkelte SI-faktor, og er dermed hensiktsmessig ved store irregulærheter. Sesongfilteret kan velges automatisk av JD+ på bakgrunn av den relative mengden irregulærhet i serien. Ved å dele (evt. trekke fra) hver observasjon O_t^* på sin respektive sesongfaktor S_t^* , får vi en sesongjustert serie $SA_t^{*(1)}$, der (1) viser til at dette er et førsteutkast.

En iterativ prosess

Prosessen fortsetter med å beregne trend på nytt, nå ved bruk av Henderson filteret⁷, på den sesongjusterte serien $SA_t^{*(1)}$. Dette filteret kan også filtrere ut en krummet trend, men krever at serien er fri for sesongvariasjoner. Dette er grunnen til at det lineære trendfilteret ble benyttet først, for da å kunne beregne førsteutkastet på sesongkomponenten og dermed lage førsteutkastet på den sesongjusterte serien. Deretter beregnes sesongkomponenten $S_t^{*(2)}$ fra den nye SI-komponenten, før et endelig trendestimat $T_t^{*(3)}$ beregnes på bakgrunn av $SA_t^{*(2)}$. Hele den til nå omtalte prosessen gjennomføres tre ganger, der målet er å

⁷ <http://www.seasonaladjustment.com/henderson> gir en gjennomgang av Henderson filteret

identifisere observasjoner som skiller seg ut, og modifisere/vekte ned disse, slik at best mulige estimater på de ulike komponentene kan hentes ut.

Steg 1

Steg 1 i X-13ARIMA består som sagt av pre-korrigeringsprosedur for outliere og kalendereffekter samt forlenging av serien ved hjelp av regARIMA-modellering. Vi vil her se på hva pre-korrigeringen innebærer og gi en kort introduksjon til ARIMA-modellen som brukes til å forlenge serien.

Pre-korrigeringsprosedur for effekt av outliere

Filtrene som brukes i sesongjusteringen i steg 2 er ulike gjennomsnitt, og de påvirkes dermed lett av outliere. Brukes et langt trend- eller sesongfilter (gjennomsnitt tas over mange observasjoner) vil en enkeltoutlier ha mindre innvirkning enn om filteret er kortere. Den iterative sesongjusteringsprosessen sørger for at outliere vektas ned, men dersom serien vår har outliere som kan forklares, f.eks. med en streik, ekstreme værforhold eller endringer i datakildene, ønsker vi å justere for disse før vi starter oppspaltingen, for at kvaliteten på sesongjusteringen skal bli best mulig. Med outliere mener vi ekstreme enkeltobservasjoner (additive outlier AO), nivåskift (level shift LS) der seriens nivå endres permanent, midlertidige nivåskift (transitory change TC) der nivået gradvis går tilbake til nivået før en spesifikk hendelse inntraff, og sesongbrudd (seasonal outlier SO) der sesongmønsteret endres for ett eller flere kvartal fra og med et spesifikt år. For de involverte tidspunktene ønsker vi å finne hva observasjonene ville ha vært om hendelsen ikke hadde inntruffet, dvs uten outliereffekten. Vi vil altså estimere outliereffekten, for å kunne fjerne denne fra serien.

Pre-korrigeringsprosedur for kalendereffekter

Særlig for månedsserier kan sammensetningen av ukedager ha betydning for nivået på serien. Observasjonen for en januar måned med fire torsdager kan f.eks. ligge lavere enn observasjonen for januar året etter som viser seg å ha fem torsdager. Effekten på seriens nivå av antall torsdager er et eksempel på en **virkedagseffekt**. Dette er en effekt som kun skyldes kalenderrelaterte variasjoner, og vi renser serien for denne før sesongjustering. Virkedagskorrigering er som oftest ikke relevant for kvartalsserier siden vi da betrakter tre måneder sammen, og det da vil være små variasjoner i sammensetningen av ukedager fra kvartal til kvartal.

Påske er en bevegelig høytid som faller i mars, april eller i overgangen mellom de to månedene. Noen serier påvirkes av påske, f.eks. ved at produksjonen avtar, ved at etterspørselen etter arbeidskraft øker, eller ved at omsetningen øker i forkant og etterkant, og er lavere under selve høytiden. Variasjoner i serien knyttet til at påskens plassering endrer seg har ikke noe med reelle endringer i trenden eller sesongmønsteret å gjøre. Derfor fjerner vi **effekten av bevegelige helligdager** (gjelder også Kristi himmelfartsdag, pinse og evt. andre bevegelige helligdager) før sesongjusteringen, dersom en slik effekt er tilstede. Slike effekter kan være tilstede både i måneds- og kvartalsserier.

Forlengelse av serien

For å kunne ta i bruk trend- og sesongfiltrene forlenger vi serien ved hjelp av en ARIMA-modell. Dette er en modell som fanger opp strukturer i tidsserier, ved at differanser og korrelasjoner mellom observasjoner blir betraktet. Modellen beskrives ved $(p, d, q)(P, D, Q)$, der de store bokstavene tar for seg sesongaspektet. p angir antall foregående observasjoner som er med på å generere O_t^* , q angir antall foregående feilledd ($\varepsilon_s^* = |\hat{O}_s^* - O_s^*$, for $s < t$; altså absolutt differanse mellom verdien som modellen gir oss og pre-korrigert verdi) som er med på å generere O_t^* , og d angir antall differanser som blir tatt ($d = 1$ innebærer at vi betrakter serien $V_t^* = O_t^* - O_{t-1}^*$, mens $d = 2$ innebærer at vi betrakter serien

$W_t^* = V_t^* - V_{t-1}^*$). Denne introduksjonen er basert på Rein (2001) som gir en god innføring i ARIMA-modeller. Den ARIMA-modellen som passer best til de observerte verdiene blir valgt, og dette valget gjøres for vårt tilfelle av JD+ ved en automatisk prosedyre som er bygd på TRAMO. Modellen brukes til å predikere fremtidige observasjoner, i tillegg til å lage tilbakeskrivninger av serien.

RegARIMA-modellering og valg av dekomponeringsmodell

Pre-korrigeringen og forlengelsen av serien skjer simultant ved at vi i tillegg til ARIMA-modellen samtidig også tar i bruk en regresjonsmodell for å beregne effektene vi vil pre-korrigere for. Modellen omtales derfor som en regARIMA-modell.

Vi ser på et eksempel: For ledige stillinger innen bygg- og anleggsvirksomhet ser vi en topp for 2. kvartal 2015 (heltrukken linje i figur 2.1). Dersom vi kjente til en hendelse som inntraff på dette tidspunktet ville vi ha pre-korrigert for effekten av en additiv outlier. Dette gjør vi her for illustrasjon. Vi inkluderer da regresjonsvariabelen

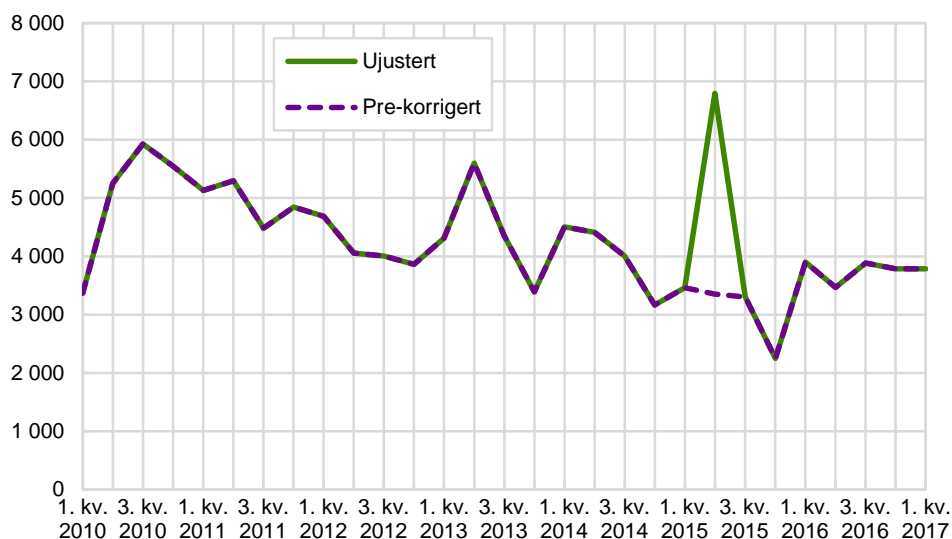
$$x_{1t} = \begin{cases} 1 & \text{for } t = 2K15 \\ 0 & \text{ellers} \end{cases}$$

i modellen vi vil bruke, og antar at det ikke er andre effekter å pre-korrigere for. Modellen vår er da

$$O_t = \beta_1 x_{1t} + O_t^*, \quad (2.1.2)$$

der O_t^* (den pre-korrigerte serien) følger en ARIMA-modell. Tilpasning av modellen vil si å finne verdien på β_1 og den ARIMA modellen $(p, d, q)(P, D, Q)$ som ut fra de tilgjengelige observasjonene gjør at høyre og venstre side av likhetstegnet er mest mulig like totalt sett for alle tidspunkt t vi har observasjoner for. O_t^* er nå en serie som er renset for effekten av den additive outlieren (stiplet linje i figur 2.1), og den er identisk med O_t for alle tidspunkt t bortsett fra for $t = 2K15$. O_{2K15}^* er dermed et estimat på hva observasjonen ville ha vært dersom det ikke hadde skjedd noe spesielt dette kvartalet.

Figur 2.1 Antall ledige stillinger innen "F: Bygg- og anleggsvirksomhet"



Kilde: Statistisk sentralbyrå.

Modellen i (2.1.2) er gyldig når dekomponeringen er additiv. I JD+ blir enten multiplikativ eller additiv dekomponering valgt, mens andre sesongjusteringsverktøy også kan ha andre dekomponeringsmodeller. En multiplikativ modell blir valgt når variasjonen ikke er konstant over tid. ARIMA-modellen bygger på en antagelse om at variasjonen er konstant, og konstant variasjon oppnås ved å ta logaritmen av hver observasjon. Altså ved å gjøre en log-transformasjon. JD+ gjør tester angående variasjonen for å vurdere hvilken dekomponering som bør velges. I det multiplikative tilfellet blir modellen fra (2.1.2) endret til

$$\ln O_t = \beta_1 x_{1t} + \ln O_t^*,$$

og vi har dermed at

$$O_t = \exp(\beta_1 x_{1t} + \ln O_t^*) = \exp(\beta_1 x_{1t}) \times \exp(\ln O_t^*) = \exp(\beta_1 x_{1t}) \times O_t^*.$$

For eksempelet med ledige stillinger innen bygg og anlegg ble en multiplikativ modell valgt. JD+ gir $\beta_1 = 0,706$ og dermed er $O_{2K15}^* = 3\,352,45$, siden vi har at $O_{2K15} = 6\,792,46$. Effekten av den additive outlieren er altså beregnet til å være $\exp(0,706) = 2,026$.

Ved r effekter å pre-korrigere for, utvides regresjonsdelen i modellen:

$$\text{Additiv modell: } O_t = \sum_{i=1}^r \beta_i x_{it} + O_t^*$$

$$\text{Multiplikativ modell: } O_t = \exp\left(\sum_{i=1}^r \beta_i x_{it}\right) \times O_t^*$$

I hvilken komponent havner de ulike effektene?

Etter at sesongjusteringen i steg 2 er gjennomført på den pre-korrigerede serien O_t^* , legges effekten av

- nivåskift til trendkomponenten T^* , slik at vi får T ,
- sesongoutliere og kalender til sesongkomponenten S^* , slik at vi får S ,
- midlertidige nivåskift og additive outliere til den irregulære komponenten I^* , slik at vi får I .

2.2. Spesifikasjonsvalg for denne statistikken

Log-transformasjon

Vi lar JD+ avgjøre om det er nødvendig å log-transformere serien.

Dekomponeringsrutine/modell

I JD+ er det bare to muligheter: Additiv eller multiplikativ dekomponering. Dersom serien er log-transformert velger JD+ multiplikativ dekomponering, hvis den ikke er transformert blir additiv dekomponering valgt.

Pre-korrigerer for virkedager og kalender

Virkedagskorrigerer er ikke relevant for serien over ledige stillinger siden vi estimerer beholdningen av ledige stillinger på et gitt tidspunkt i kvartalet. Dersom vi hadde sett på tilgangen av ledige stillinger i løpet av kvartalet, kunne sammensetningen av ukedager derimot hatt en liten betydning. Heller ikke antall ansatte vil være påvirket av antall virkedager i kvartalet.

Vi anser også påske som irrelevant for de to seriene. Virksomheter med høy påskeaktivitet, som skisentre og høyfjellshoteller, har som oftest høysesong over flere måneder i vinterhalvåret, og vi antar dermed at påskens plassering ikke vil ha betydning for utlysning av stillinger innen næringene som disse virksomhetene tilhører. Når det gjelder antall ansatte er disse tallene hentet ut for februar og mai,

og vi har dermed unngått mars og april, som kan være påvirket av påske for noen næringer.

Pre-korrigerer for outliere

JD+ har en innebygd algoritme for å se etter outliere, og defaultgrensen for å oppfatte observasjoner som outliere er streng (t-verdi på omtrent 4). Koeffisienten må altså være veldig signifikant for at observasjoner skal slå ut som outliere. For hver av de 40 seriene har vi sjekket hvilke outliere JD+ finner automatisk, og vurdert hver av disse. Anbefalt ESS-praksis knyttet til outliere (punkt 2.7 i ESS, 2015) er at man skal ha en forklaring for outlieren, dersom en observasjon skal regnes som dette. Vi har ikke gått i dybden på enhver outlier, men har inkludert regresjonsvariable for enkelte outliere som vi mener lar seg forklare. Disse har da blitt pre-spesifisert i regARIMA-modellen, og vi har deretter stengt muligheten for å la JD+ se etter flere outliere. I kapittel 2.3 kommer vi inn på håndtering av brudd i «ansatte» knyttet til overgangen til A-ordningen.

Valg av ARIMA-modell

Beste ARIMA modell $(p, d, q)(P, D, Q)$ velges automatisk (TRAMO-basert) av JD+ blant alle muligheter under restriksjonene $p, d, q \leq 3$ og $P, D, Q \leq 1$. Modellen som da velges, holdes fast i ett år i tråd med beste praksis i ESS-retningslinjene for sesongjustering (punkt 4.2). Dersom vi ikke hadde låst modellen kunne vi risikere å få forskjellige modeller fra kvartal til kvartal, og dermed mer revisjon i den sesongjusterte serien enn ved å låse den.

Valg av trend- og sesongfilter

Vi lar JD+ velge trend- og sesongfilter automatisk, og de samme filterne blir da valgt for alle kvartal. Et kort filter bruker bare observasjoner som tidsmessig er i nærheten av en spesifikk observasjon for å beregne sesongkomponenten eller trendtallet for denne observasjonen, mens et langt filter også bruker observasjoner lengre borte og blir «glattere». Ifølge punkt 3.6 i ESS-retningslinjene for sesongjustering er beste praksis når det gjelder sesongfiltere å «gå aktivt inn for å skaffe seg informasjon om periode-spesifikke årsaker for endring i sesongmønsteret». Og bruke dette til å «vurdere bruk av periode-spesifikke sesongfiltere, i det minste for justering av viktige makroøkonomiske aggregater». For sesongfiltere kan man altså velge ulikt filter for ulike kvartal. Vi så litt på dette, men endte opp med å ikke velge ulike filtre, da vi ikke hadde noen klare grunner for å gjøre dette.

Ifølge punkt 16.4.3 i Guide to Seasonal Adjustment with X-12-ARIMA (Office for National Statistics (2007)) fra det britiske statistikkbyrået er det i de fleste tilfeller tilstrekkelig å bruke filteret som velges automatisk. For ledige stillinger innen bergverksdrift og utvinning har vi valgt å bruke et 3×3 filter for alle kvartal, i stedet for 3×5 filteret som ble valgt automatisk. Årsaken til dette var for raskere å fange opp endringer i sesongmønsteret, da dette mønsteret ser ut til å endre seg ved overgangen 2014-2015.

Revisjonsrutine

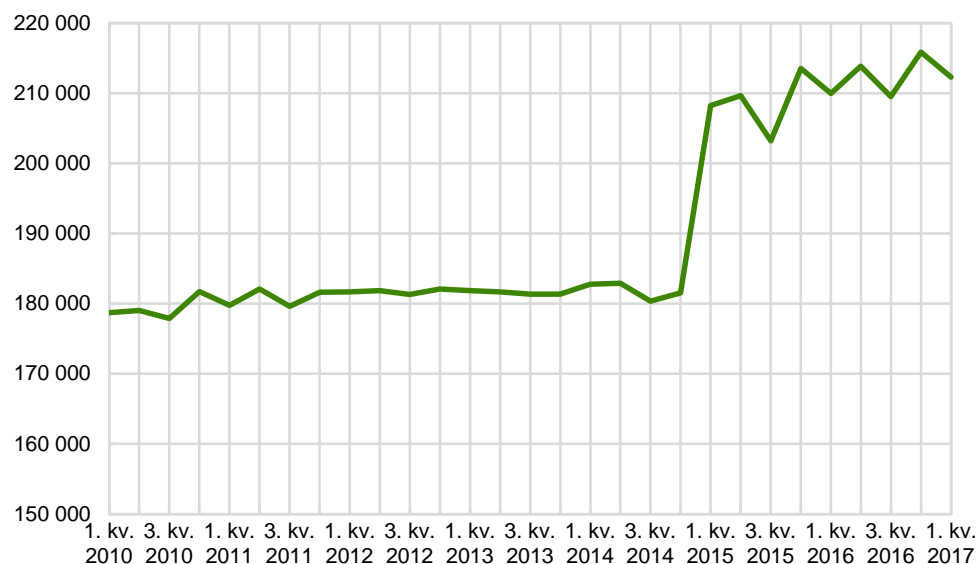
Vi har valgt delvis løpende korrigerer (partial concurrent adjustment) som revisjonsrutine. Dette innebærer at filterne (sesong og trend) og modellen; dekomponeringsmodell og regARIMA-modell (hvilke effekter det skal pre-korrigeres for og hvilken ARIMA modell som velges) identifiseres ved en årlig gjennomgang av sesongjusteringen, og deretter holdes fast i løpet av året. Alle parameterne i regARIMA-modellen re-estimeres derimot hver gang nye data kommer til. Dette fører til at revisjoner i de sesongjusterte tallene og i trendtallene fordeles utover året.

2.3. Håndtering av brudd i seriene

Brudd i «ansatte» ved innføring av A-ordningen i 2015

Ved overgangen til A-ordningen 1. kvartal 2015 ser vi et tydelig nivåskift i «ansatte» for 9 av de 20 næringsgruppene vi publiserer for. For disse næringsgruppene finner JD+ nivåskiftet automatisk. Et eksempel er «Q88: Sosiale omsorgstjenester uten botilbud» der den ujusterte serien er vist i figur 2.2.

Figur 2.2 Antall ansatte innen "Q88: Sosiale omsorgstjenester uten botilbud"



Kilde: Statistisk sentralbyrå.

For enkelte av de andre næringsgruppene er nivåskiftet signifikant på 5 % nivå, men vi finner det kun når vi ber programmet spesifikt om å se etter dette nivåskiftet (dette dreier seg om den tidligere omtalte strenge grensen for når observasjoner skal regnes som outlierer).

Det ble vurdert om vi skal inkludere dette nivåskiftet a) kun i de 9 seriene, b) for alle seriene der nivåskiftet er signifikant på omtrent 5 % nivå ($|t\text{-verdi}| \geq 2$), c) for alle serier som har $|t\text{-verdi}| \geq 1$ (valgt et tall lavere enn 2 for å redusere litt på «kravet» om signifikans, siden vi kjenner en årsak til brudd), d) for alle seriene. Inkludering av nivåskift for et tidspunkt innebærer at observasjonene før bruddet i den pre-korrigerte serien er flyttet til det nye nivået som har blitt etablert. I regARIMA-modellen bruker vi da regresjonsvariabelen

$$LS_{1K15} = \begin{cases} -1 & \text{for } t < 1K15 \\ 0 & \text{ellers.} \end{cases}$$

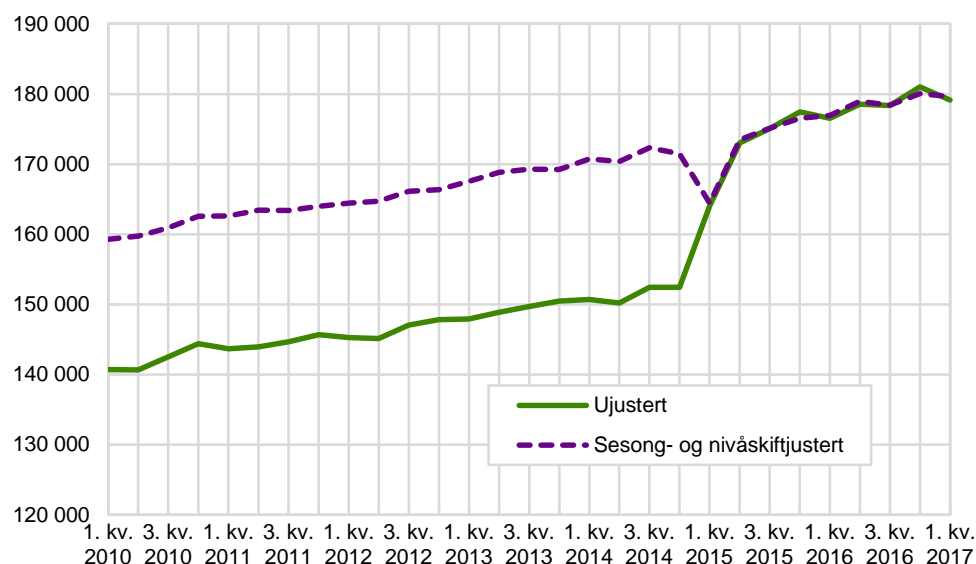
Spesifisering av et nivåskift eller en annen outlier har i seg selv bare innvirkning på selve oppspaltingen. Den pre-korrigerte serien uten nivåskift spaltes i T^* , S^* og I^* , og når den iterative sesongjusteringsprosessen er ferdig legges effekten av nivåskiftet til i trendkomponenten. Nivåskiftet vil dermed være synlig i den sesongjusterte serien. Men vi kan også velge å lage sesongjusterte serier som er justert for nivåskiftet, noe vi kommer tilbake til lenger ned.

Vi sesongjusterer 20 delserier for «ansatte», og utleder deretter sesongjustert total ved sum. Det som er usignifikante nivåskift for flere delserier/næringer, kan fort summere seg opp til ett signifikant nivåskift for totaltallene. Vi har derfor vurdert både alternativ a) og b) som uaktuelle. Når vi kjenner endringer som har påvirket serien vår, er det viktig å spesifisere dette i modellen. Vi ser at innføringen av A-

ordningen gir et hopp for enkelte av næringsgruppene og vet dermed at den har en effekt. Men betrakter vi alle seriene, er hoppet stort for noen serier og nær null for andre. Siden innføringen av A-ordningen gjaldt alle næringer og siden vi med sikkerhet vet at det hadde en effekt for noen, har vi løsrevet oss fra tanken om at et brudd må være signifikant for å inkludere nivåskift. Derfor har vi valgt å gå for å innføre nivåskift for alle «ansatte» delseriene, det vil si alternativ d).

For å få en helhetlig serie som i minst mulig grad er påvirket av overgangen til ny datakilde, har vi i tillegg valgt å la de sesongjusterte seriene være justert for dette nivåskiftet i 2015. Dette innebærer at effekten av nivåskiftet ikke legges til i trenden etter oppspaltingen. Ved å justere for nivåskiftet har vi ikke nødvendigvis justert for hele bruddet ved overgangen til A-ordningen. Et eksempel er «ansatte» innen «O: Offentlig administrasjon» der ansatte-tallet for 1. kvartal 2015 befinner seg mellom nivået for periodene før overgangen til A-ordningen og nivået f.o.m. 2. kvartal 2015, se figur 2.3. Dette har vi tolket som «oppstartsproblemer» for A-ordningen, og har pre-spesifisert en additiv outlier for 1. kvartal 2015. Denne har vi imidlertid ikke justert den sesongjusterte serien for, og den er dermed synlig i de sesongjusterte tallene.

Figur 2.3 Antall ansatte innen "O: Offentlig administrasjon"



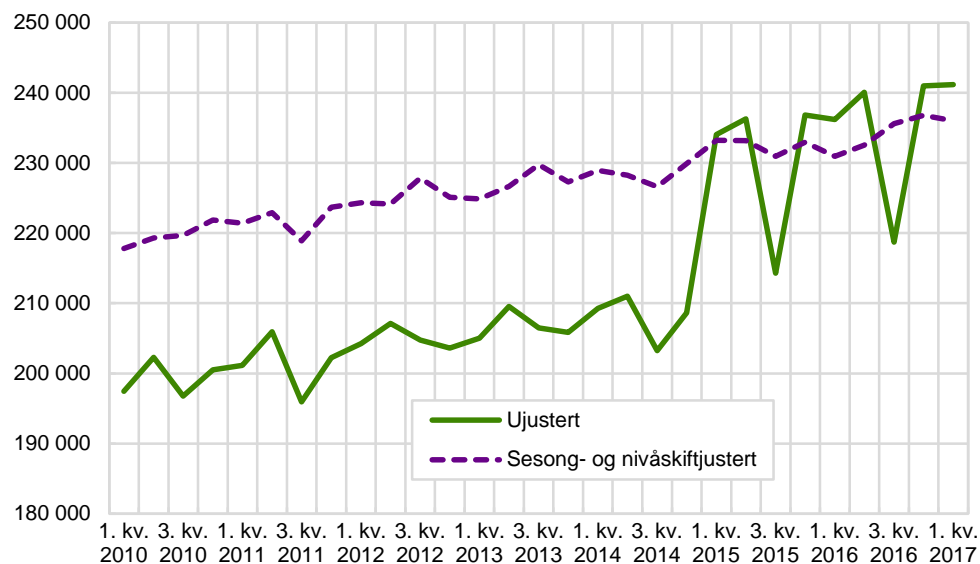
Kilde: Statistisk sentralbyrå.

For noen av seriene har vi pre-spesifisert sesongbrudd (omtales som sesongoutliere), siden disse er signifikante og sesongmønsteret ser ut til plutselig å ha endret seg markant. Alle disse er knyttet til kvartaler i 2015, og kan dermed kobles til endret sesongmønster når vi fra 2015 har et nytt datagrunnlag. En av disse seriene er «ansatte» innen «P: Undervisning» som har en sesongoutlier for 3. kvartal 2015, se figur 2.4. For den ujusterte serien ser vi tydelig at det for 3. kvartal i 2015 og 2016 er større fall i nivået enn for tidligere år. Dette knytter vi til endringer i innrapporteringen fra arbeidsgiverne som viser seg å slå særlig ut for 3. kvartal for ansatte innen undervisningssektoren. Dersom vi ikke pre-korrigerer for sesongoutlieren, får vi sesongkomponenten i figur 2.5 som viser et særlig bevegelig sesongmønster. Ved å inkludere sesongoutlieren

$$SO_{3K15} = \begin{cases} 0 & t \geq 3K15 \\ 1 & t < 3K15 \cap t \in 3K \\ -\frac{1}{3} & \text{ellers} \end{cases}$$

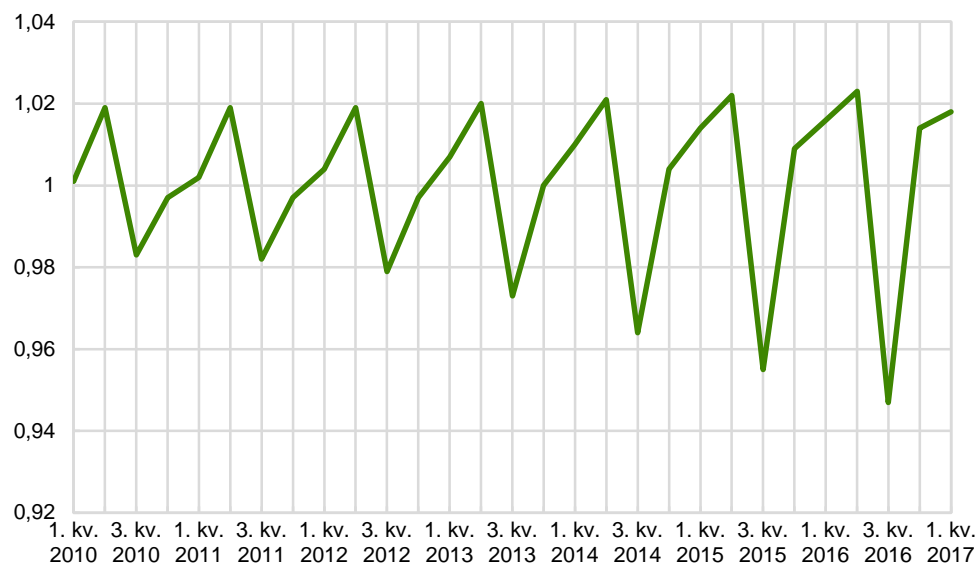
får vi sesongkomponenten i figur 2.6 som har et mer stabilt mønster. Pre-korrigerer for sesongoutlieren fører til at endringene i innrapporteringsrutinene ikke forstyrrer beregningen av sesongkomponenten. For den sesongjusterte serien i figur 2.4 er sesongoutlieren pre-spesifisert. I figur 2.5 og 2.6 ser vi at y-aksene er ulike. Dette skyldes at multiplikativ modell ble valgt når vi ikke pre-korrigerer for den signifikante sesongoutlieren, mens additiv modell ble valgt når vi inkluderte den.

Figur 2.4 Antall ansatte innen "P: Undervisning"



Kilde: Statistisk sentralbyrå.

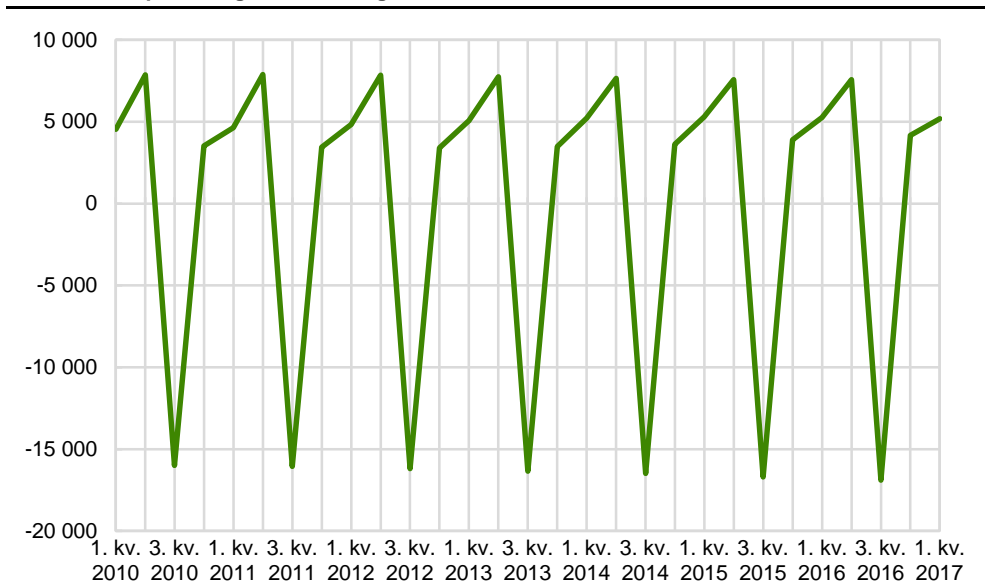
Figur 2.5 Beregnet sesongkomponent for antall ansatte innen "P: Undervisning" når vi ikke har pre-korrigert for sesongoutlier 3. kvartal 2015



Kilde: Statistisk sentralbyrå.

Også for noen av seriene der vi ikke har pre-spesifisert sesongoutliere ser vi tendenser til endringer i sesongmønsteret fra 2015. Dette påpekes for de aktuelle seriene i gjennomgangen i kapittel 3.3, og disse seriene bør vurderes nærmere når seriene har blitt lengre, for å bedre kvaliteten på sesongjusteringen. Dette vil også gi et bedre grunnlag for å lage en sesongjustert serie som er fullstendig justert for bruddet ved overgangen til A-ordningen.

Figur 2.6 Beregnet sesongkomponent for antall ansatte innen "P: Undervisning" når vi har pre-korrigert for sesongoutlier 3. kvartal 2015



Kilde: Statistisk sentralbyrå.

Brudd i «ledige stillinger» ved innføring av A-ordningen?

Antall ansatte brukes som den eneste forklaringsvariabelen i estimeringen av ledige stillinger. Det kunne da være naturlig at også serien «ledige stillinger» justeres for nivåskift ved overgangen til A-ordningen. Ved å pre-spesifisere nivåskiftet, dvs. å legge inn regresjonsvariabel for dette i reARIMA-modellen, finner JD+ signifikante nivåskift på 5 prosent nivå i 7 av seriene. Den automatiske prosedyren for å se etter outlierer finner derimot bare signifikant nivåskift på 5 prosent nivå i serien «N: Forretningsmessig tjenesteyting». Dette kommer av at mange av seriene er ganske irregulære, slik at det som gir et signifikant nivåskift ved pre-spesifisering tolkes annerledes når den automatiske prosedyren for å se etter outlierer benyttes. Vi har valgt å ikke legge inn noe nivåskift for «ledige stillinger».

Nivåskift i «ansatte» for næringsgruppene «B: Bergverksdrift og utvinning» og «C: Industri» også i 2011

JD+ finner nivåskift for 1. kvartal 2011 i de to seriene «B: Bergverksdrift og utvinning» og «C: Industri» over «ansatte». Vi sjekket oversikter laget for Kvartalsvis nasjonalregnskap (KNR) over virksomheter som byttet næring på dette tidspunktet, og tallene gikk tilnærmet opp i opp for disse to næringsgruppene. Det var en plutselig vekst innen bergverksdrift og utvinning, og en plutselig nedgang innen industri. Vi pre-korrigerer for nivåskiftene, men fortsetter å la hoppet være synlig i de sesongjusterte tallene og i trenden, da dette er reelle endringer i tallet på ansatte innen disse næringene.

3. Sesongjustering av seriene

I 3.1 tar vi en gjennomgang av de ulike vurderingene som er gjort angående kvaliteten på sesongjusteringen, mens resultatene for de 20 delseriene for «ledige stillinger» og for «ansatte» gis i respektive 3.2 og 3.3.

3.1. Vurderinger som er gjort på seriene

Et viktig aspekt ved brukervennligheten til JD+ er bruken av fargekoder for å avhjelpe tolkningen av resultater, se figur 3.1. Grønt indikerer bra, rødt indikerer dårlig, og gult indikerer usikker kvalitet.

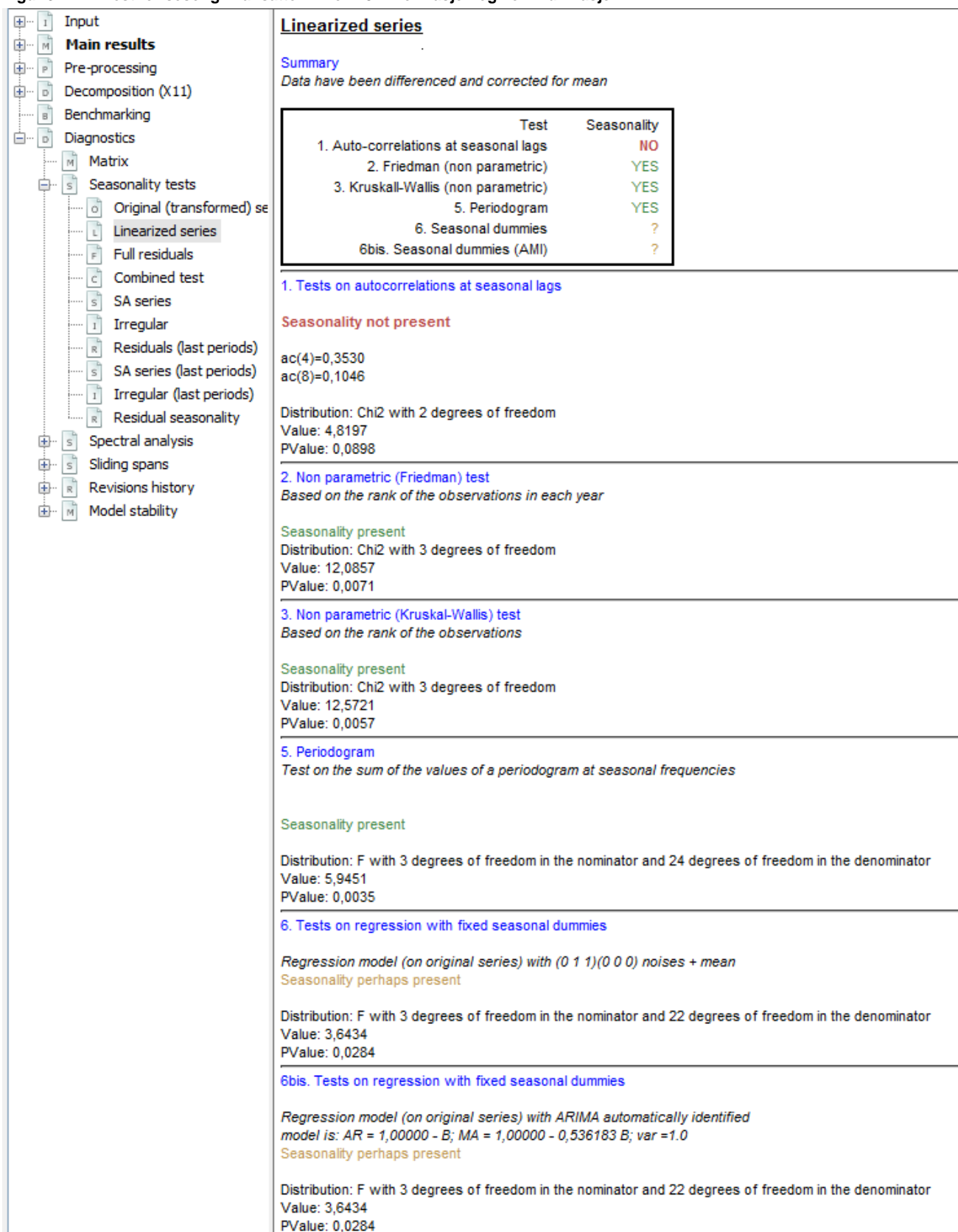
Er det sesongeffekter i serien?

En forutsetning som må være oppfylt for å sesongjustere, er at serien har systematiske sesongeffekter/ variasjoner. X-13ARIMA i JD+ inneholder seks tester for sesong under stien «Diagnostics» → «Seasonality tests». Nullhypotesen i de ulike testene er at det ikke er sesong i dataene. Hypotesen forkastes på 1 prosent nivå, altså dersom p-verdien er lavere enn 0,01. JD+ gir da resultatet «YES» og «Seasonality present», der altså spørsmålet om det er sesongeffekter i serien blir positivt besvart. Dersom p-verdien ligger i intervallet [0,01, 0,05] får vi som resultat at serien kanskje har en sesongeffekt, og JD+ gir oss da «?» og «Seasonality perhaps present». For høyere p-verdier får vi resultatet «NO» og «Seasonality not present». Figur 3.1 viser testresultatet for «ansatte» for delserien «J: Informasjon og kommunikasjon».

Den første testen regnes for å være den viktigste, og vi har fokusert mest på den i tilfeller der testene gir ulike resultater. Denne tester for autokorrelasjon ved sesonglag, altså om det er en lineær sammenheng mellom de samme periodene for de ulike årene. Nullhypotesen er dermed at det ikke er noen lineær sammenheng med samme periode tidligere år. De seks testene utføres både på originalserien og på den lineariserte serien. Dersom serien er pre-korrigert vil vi betrakte den lineariserte serien når vi sjekker for sesong. Her er deterministiske effekter (kalendereffekter og effektene av outliere) fjernet, og det vil være på grunnlag av denne serien at sesongkomponenten beregnes.

I noen tilfeller der testene gir negativt eller usikkert resultat, kan det å la være å sesongjustere gi gjenværende sesong i serien: Dette ser vi under «Diagnostics» → «Seasonality tests» → «Residual seasonality». Her må man vurdere hvert enkelt tilfelle, men som hovedregel har vi valgt å sesongjustere i slike tilfeller. I tvilstilfeller har vi også sett på om den valgte ARIMA-modellen har en sesongdel, altså om en eller flere av P, D, Q i $(p, d, q)(P, D, Q)$ er forskjellig fra null.

Figur 3.1 Test for sesong i «ansatte» innen «J: Informasjon og kommunikasjon»



Kilde: Statistisk sentralbyrå.

Kvalitetsmål (M-målene)

JD+ gir 11 kvalitetsmål på sesongjusteringen M1-M11, samt oppsummeringsmålene Q og (Q - M2) som er vektete gjennomsnitt av de 11 M-målene. Disse finner vi under «Decomposition» → «Quality measures» → «Summary». De 13 kvalitetsmålene har verdier i intervallet [0, 3], der en

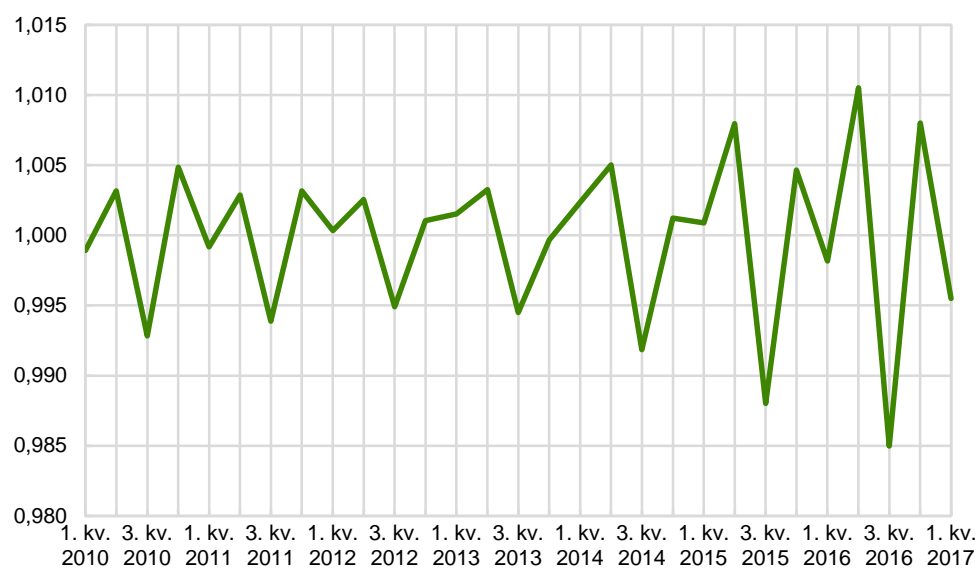
tilfredsstillende verdi er lavere enn 1. De viktigste kvalitetsmålene er M7 og Q, og vi har hovedsakelig fokusert på disse under sesongjusteringen.

M7 er et mål på mengden stabil sesong (tilnærmet konstant sesongmønster) tilstede relativt til mengden bevegelig sesong (sesongmønster som utvikler seg over tid). M1 er også viktig, og viser hvor stor den irregulære komponenten er sammenlignet med sesongkomponenten. Høy M1 impliserer at den irregulære komponenten er stor og at det derfor kan være vanskelig å estimere sesongkomponenten nøyaktig. Høye verdier for M8 og M9 forteller at en serie kan være problematisk å sesongjustere fordi sesongmønsteret endrer seg for raskt, mens M10 og M11 er de samme målene, men kun for de tre siste årene. M7, og flere av de andre målene blir ikke beregnet når vi ikke har med sesongkomponenten.

Er sesongeffektene identifiserbare?

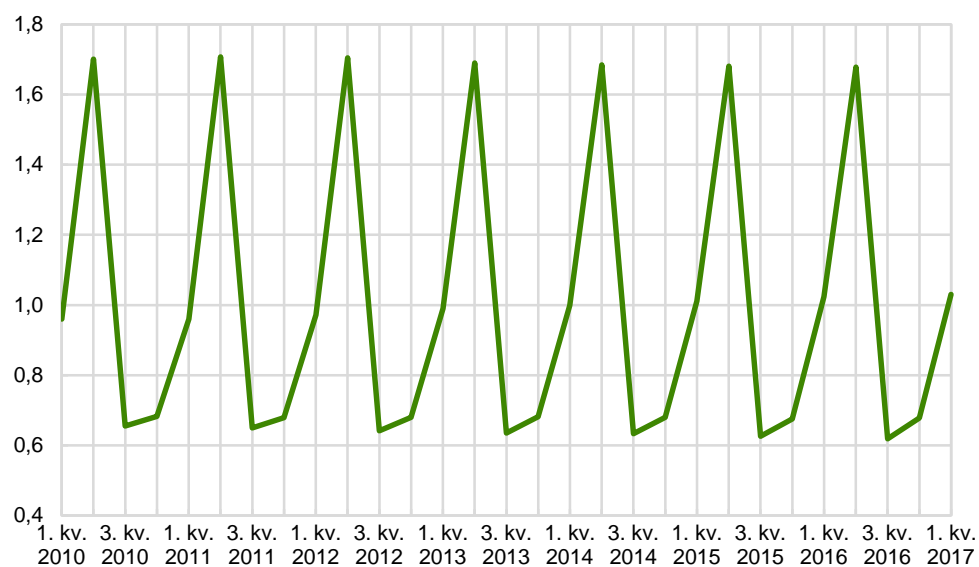
M7 ser altså på hvor mye bevegelig sesong det er i en serie, relativt til mengden stabil sesong. Dersom sesongen beveger seg mye er det vanskelig å beregne en sesongkomponent med god kvalitet. I figur 3.2 ser vi den beregnede sesongkomponenten for «ansatte» innen «Q88: Sosiale omsorgstjenester uten botilbud». M7 er 1,17 for denne serien, og det er ingen identifiserbar sesong i serien på grunn av stor grad av bevegelse i sesongmønsteret, noe vi ser tydelig fra grafen. Kvaliteten på sesongjusteringen er usikker.

Figur 3.2 Sesongkomponent for antall ansatte innen "Q88: Sosiale omsorgstjenester uten botilbud"



Kilde: Statistisk sentralbyrå.

I figur 3.3 ser vi den beregnede sesongkomponenten for «ledige stillinger» innen «P: Undervisning». M7 er her 0,14, og det er en identifiserbar sesong i serien. Kvaliteten på sesongjusteringen er god.

Figur 3.3 Sesongkomponent for antall ledige stillinger innen "P: Undervisning"

Kilde: Statistisk sentralbyrå.

Påstandene over om identifiserbarhet for sesongmønsteret er resultater fra «Combined seasonality test». Denne testen finner vi under «Diagnostics» → «Seasonality tests» → «Combined test» og den har tre mulige utfall:

- «Identifiable seasonality present»,
- «Identifiable seasonality probably not present»,
- «Identifiable seasonality not present».

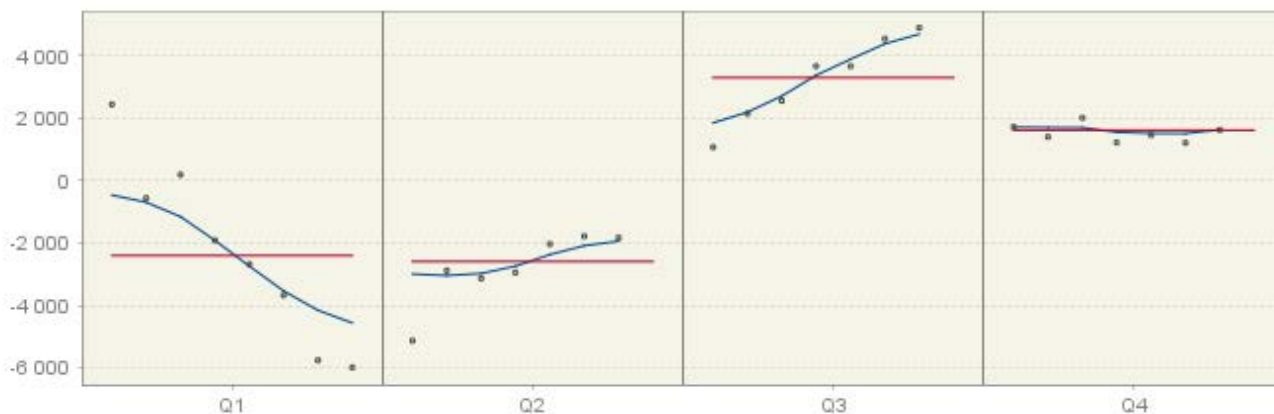
Det er i orden å sesongjustere dersom resultatet av testen er et av de to første utfallene, så sant andre tester også tilsier sesongjustering, mens sesongjustering bør unngås ved det siste utfallet. Testen vurderer identifiserbarhet for sesongen, altså om det er sesong til stede og om den ikke beveger seg for mye. Testen er nært knyttet til M7.

SI ratios

Vi finner «SI ratios» på «Main results» fanen i JD+. Figur 3.4 viser SI ratios for «ansatte» innen «N: Forretningsmessig tjenesteyting». Den blå kurven er den estimerte sesongkomponenten (oppdelt etter kvartal), mens de svarte punktene utgjør sesong + irregulær komponent (SI) og kalles SI ratios. Den røde linja viser gjennomsnittet av SI ratioene. Dersom de blå grafene tilnærmet sammenfaller med sine respektive gjennomsnitt har vi et sesongmønster med høy grad av stabilitet og med liten bevegelse. Sesongkomponenten forventes å være stabil, men det er også naturlig at sesongmønsteret endres noe over tid. Dersom SI ratioene ligger langt fra den estimerte sesongkomponenten indikerer dette at tidsserien er dominert av den irregulære komponenten, og sesongkomponenten kan da endre seg vesentlig når nye observasjoner kommer til, som igjen vil gi vesentlige revisjoner i de sesongjusterte tallene.

SI ratios kan brukes til å vurdere lengden på sesongfilteret. Dersom SI ratioene ikke følger den estimerte sesongen sier punkt 16.4.3 i Office for National Statistics (2007) at det kan være passende å bruke et kortere sesongfilter enn det som velges automatisk.

Figur 3.4 SI ratios for «ansatte» innen «N: Forretningsmessig tjenesteyting»



Kilde: Statistisk sentralbyrå.

Samlet kvalitet

JD+ gir en samlet kvalitetsindikator for sesongjusteringen (god – usikker - alvorlig). Denne innbefatter bl.a. tester for gjenværende sesong og M-målene.

Annen diagnostikk

Det er mye diagnostikk i JD+ som kan vurderes for de ulike seriene. F.eks. sliding spans (krever rundt 8år med data) og revisjonshistorie (velge den modellen som gir minst revisjon, men vi velger uansett beste modell, og stoler på at denne da blant annet vil oppfylle minst mulig revisjon), og denne diagnostikken har blitt sjekket ut til en viss grad. I og med at det her dreier seg om hele 40 delserier har vi i dette notatet kun vurdert den diagnostikken som allerede er nevnt i avsnittene over.

3.2. Ledige stillinger

Vi vil nå gi en oppsummering av de vurderingene som ble gjort på delseriene for «ledige stillinger» på bakgrunn av diagnostikken fra JD+.

A: Jordbruk, skogbruk og fiske

Test 1 for sesong i rådataene gir negativt resultat (p-verdi på 0,20), mens enkelte av de andre testene er usikre. Det blir muligens noe gjenværende sesong ved ikke å sesongjustere, og vi velger derfor å sesongjustere serien. Kvaliteten er god nå, men kan bli usikker ved senere kvartaler. Oppsummeringsmålet Q er 0,92, og dersom det senere overstiger 1, kan vi få usikker kvalitet på sesongjusteringen. Vi ser topper i serien for 2. kvartal i 2010 og fra og med 2013. 2. kvartal i 2011 og 2012 bidrar derfor, sammen med flere andre observasjoner, til å forstyrre den beregnede sesongkomponenten. Combined seasonality test gir usikkert resultat, noe som indikerer at det er OK å sesongjustere. Serien sesongjusteres.

B: Bergverksdrift og utvinning

Alle testene for sesong i rådataene gir negativt resultat. Test 1 har p-verdi 0,33. Ved ikke å sesongjustere blir kvaliteten god, og vi finner ingen gjenværende sesong i serien. Også sesongjustering gir god kvalitet i dette tilfellet, selv om både $M7(=1,33)$ og $Q(=1,06)$ er større enn 1. For forrige serie kommenterte vi at $Q>1$ kan gi usikker kvalitet. Her ser vi at vi får god kvalitet selv om $Q>1$. Det er som tidligere nevnt flere elementer som inngår i kvalitetstesten enn de vi går inn på i denne gjennomgangen.

Når vi kun betrakter de fem første årene, altså t.o.m. 4. kvartal 2014, ser vi en tydelig sesongeffekt for 2. kvartal. Ved å ikke sesongjustere serien t.o.m. 2014 får vi muligens gjenværende sesong i serien, men også nå er p-verdi for test 1 for rådataene høyere enn 0,05 (p-verdi=0,14). Ved sesongjustering av serien t.o.m. 2014 får vi $M7$ på kun 0,66, mens $M1$ er på hele 1,83, så den irregulære

komponenten er stor. Combined seasonality test gir usikkert resultat, noe som indikerer at det er OK å sesongjustere. Samlet sett blir det god kvalitet i begge tilfeller.

Denne serien er et tvilstilfelle, og det kunne være riktig å ikke sesongjustere. Sesongjustering ble likevel valgt, på grunn av mønsteret vi ser til og med 2014, og siden kvaliteten er tilfredsstillende. Et 3×5 sesongfilter ble valgt av JD+, men vi valgte å endre dette til et 3×3 filter for raskere å fange opp endringer i sesongmønsteret som ser ut til å endre seg ved overgangen 2014-2015. Serien sesongjusteres.

C: Industri

Det er klare sesongvariasjoner i serien; alle testene for sesong i den ujusterte serien er positive. Test 1 har en p-verdi på 0,005. M7 er 0,37 og Q er 0,90, der det særlig er M1 på 2,16 som trekker opp Q. Det er dermed en del irregulareteter, men sesongen er klart identifiserbar. Serien sesongjusteres.

D+E: Kraft og vannforsyning

Test 1 for sesong i rådataene gir negativt resultat (p-verdi=0,69), mens enkelte av de andre er usikre. ARIMA-modellen har ingen sesongdel. Ved ikke å sesongjustere får vi usikker kvalitet og muligens gjenværende sesong i serien. Også ved å sesongjustere blir kvaliteten usikker. Vi får da $Q=1,19$ og $M1=3$, så den irregulære komponenten er stor. I tillegg er særlig $M10 (=1,86)$ og $M11 (=1,77)$ høye. Disse to målene vurderer sesongkomponenten for de siste tre årene; størrelsen på fluktuasjoner og den gjennomsnittlige lineære bevegelsen. M7 er derimot bare 0,82, men det er usikkerhet rundt identifiserbarhet for sesongmønsteret. Når vi tar alle disse momentene i betraktning har vi valgt å ikke sesongjustere serien. Serien sesongjusteres ikke.

F: Bygge- og anleggsvirksomhet

Alle testene for sesong i rådataene gir negativt resultat, der test 1 har p-verdi på 0,18. Ved ikke å sesongjustere får vi muligens noe gjenværende sesong når vi betrakter hele serien, men kvaliteten er god. Ved å sesongjustere får vi M7 på 1,25 og Q på 1,79, ikke-identifiserbar sesong, men god kvalitet. For denne serien har vi at 2. kvartal 2015 skiller seg ut som et kvartal med særlig mange ledige stillinger, omtrent 6 800, mot drøye 3 000 i 1. og 3. kvartal samme år. Vi har ikke undersøkt dette nærmere. Dersom dette kvartalet skulle anses for å være en ekstremverdi ville vi fått M7 på 0,89 og Q på 1,24, men fremdeles ikke-identifiserbar sesong, og nå blir kvaliteten usikker. Vi har valgt å ikke sesongjustere denne serien, da det ikke er tydelige indikasjoner på gjenværende sesong ved å la være å sesongjustere. Serien sesongjusteres ikke.

G: Varehandel, reparasjon av motorvogner

Test 1 for sesong i rådataene gir negativt resultat (p-verdi=0,14), mens tre av de andre er positive. Ved ikke å sesongjustere får vi god kvalitet, men testen for gjenværende sesong for hele serien er positiv på 1 prosent nivå. Dette indikerer at det er sesongvariasjoner i serien. Ved sesongjustering er ikke sesongmønsteret identifiserbart, men testene gir at stabil sesong er tilstede på 1 % nivå og at det ikke er bevegelig sesong på 20 % nivå, samt at både M7 og Q er klart lavere enn 1 og at samlet kvalitet er god. På bakgrunn av dette har vi valgt å sesongjustere. Serien sesongjusteres.

H: Transport og lagring

Test 1 for sesong i rådataene gir usikkert resultat (p-verdi=0,03), mens tre av testene er positive. Ved å ikke sesongjustere får vi gjenværende sesong i serien, men ikke for de tre siste årene. Sesongjustering gir usikkerhet rundt

identifiserbarhet for sesongmønsteret, men alle M-målene og Q er lavere enn 1. Serien sesongjusteres.

I: Overnattings- og serveringsvirksomhet

4. kvartal ligger lavest for hvert år, og alle testene for sesong i rådataene er positive. Sesongmønsteret er identifiserbart, og bare M5 (=1,02) blant kvalitetsmålene er høyere enn 1. Serien sesongjusteres.

J: Informasjon og kommunikasjon

Test 1 for sesong i rådataene gir usikkert resultat, men p-verdien er bare 0,02. To av testene er positive. Både M7 og Q er klart lavere enn 1 ved sesongjustering, mens M8-M11 er omtrent 1,3. Sesongmønsteret er identifiserbart. Serien sesongjusteres.

K: Finansiering og forsikringsvirksomhet

Test 1 for sesong i rådataene gir negativt resultat (p-verdi=0,08), mens tre av de andre testene er positive. Ved å ikke sesongjustere blir resultatet av testen for gjenværende sesong på 5 prosent nivå positiv, men kvaliteten er god. Vi får også god kvalitet ved å sesongjustere, og både M7 og Q er da lavere enn 1. Det er identifiserbar sesong tilstede. Når vi lar JD+ se etter outliere finner programmet et nivåskift for 1. kvartal 2012; dette har vi ikke tatt hensyn til. Serien sesongjusteres.

L: Omsetting og drift av fast eiendom

Alle testene for sesong i rådataene gir negativt resultat (p-verdi \approx 1 for test 1). Ved ikke å sesongjustere blir kvaliteten god, og det er ikke statistisk bevis for gjenværende sesong. Serien sesongjusteres ikke.

M: Teknisk tjenesteyting

Alle testene for sesong i rådataene gir negativt resultat (p-verdi=0,75 for test 1). Når vi lar være å sesongjustere er kvaliteten usikker, blant annet på grunn av noe avhengighet i regARIMA residualene, men det er ikke statistisk bevis for gjenværende sesong i serien. Når vi i tillegg lar JD+ se etter outliere finner programmet et midlertidig nivåskift 1. kvartal 2016. Dette kommer av at nivået på serien er lavt, med et snitt på rundt 3 700 for hele tidsperioden, og at serien er ganske stabil for perioden 2012-2014, synker noe i 2015, og deretter gjør et hopp oppover i 1. kvartal 2016. Dette tolker vi som tilfeldige variasjoner, og har ikke undersøkt det noe nærmere. Inkludering av outlieren fører til god kvalitet, men den er ikke tatt med. Serien sesongjusteres ikke.

N: Forretningsmessig tjenesteyting

Test 1 for sesong i rådataene gir usikkert resultat, men p-verdien er bare 0,03. Tre av testene er positive. M7 er 0,48, Q er 0,55 og kun M10 og M11 er høyere enn 1 (1,25). Sesongmønsteret er identifiserbart. Serien sesongjusteres.

O: Offentlig administrasjon

Test 1 for sesong i rådataene gir negativt resultat (p-verdi=0,86), mens to av testene er usikre. Ved ikke å sesongjustere blir kvaliteten usikker, med gjenværende sesong i serien, men ikke for de siste tre årene. Sesongjustering gir god kvalitet, og både M7 og Q er lavere enn 1, mens M1 er hele 1,80. Det er altså en stor del irregularitet i serien. Det er usikkerhet rundt identifiserbarhet for sesongmønsteret, men i slike tilfeller kan det være i orden å sesongjustere. Serien sesongjusteres.

P: Undervisning

Det er klare sesongvariasjoner i serien; alle testene for sesong i rådataene er positive. M7 er her bare 0,14, mens Q er 0,56. Sesongmønsteret er identifiserbart og kvaliteten er god. Serien sesongjusteres.

Q86: Helsetjenester

Alle testene for sesong i rådataene gir negativt resultat (p -verdi=0,93 for test 1). Det blir ingen gjenværende sesong ved å la være å sesongjustere. Serien sesongjusteres ikke.

Q87: Pleie- og omsorgstjenester i institusjon

Test 1 for sesong i rådataene gir negativt resultat (p -verdi=0,31), mens to av de andre er usikre. Den valgte ARIMA-modellen har sesongdel. Kvaliteten blir god både ved å sesongjustere, og ved ikke å gjøre det. Ved sesongjustering er ikke sesongmønsteret identifiserbart. $M7$ er likevel ikke høyere enn 0,98, mens Q er 1,02, der det særlig er $M1$ -målet på 2,63 som trekker opp det samlede kvalitetsmålet. Uten sesongkomponent sier testen for gjenværende sesong at sesongvariasjoner er tilstede, men ikke for de tre siste årene. Siden vi ikke finner sesong for de siste årene, bør vi være varsomme med sesongjustering. Da valget ble tatt valgte vi likevel å sesongjustere serien på grunnlag av sesongen i serien sett under ett. Serien sesongjusteres.

Q88: Sosiale omsorgstjenester uten botilbud

Test 1 for sesong i rådataene gir positivt resultat (p -verdi=0,001), mens én av de andre testene er usikker. Sesongjustering gir usikker kvalitet, med Q rett over 1, mens $M7$ er 0,61. Kvaliteten er god dersom man bare betrakter de fem første årene, altså opp til innføringen av A-ordningen. Denne endringen i datakilde kan ha endret sesongmønsteret noe, da antall ansatte brukes som forklaringsvariabel for ledige stillinger. For denne femårs perioden er det fremdeles usikkerhet rundt identifiserbar sesong. Når vi igjen ser på hele serien og lar være å sesongjustere, er det statistisk bevis for gjenværende sesong i serien, men ikke for de tre siste årene. Kvaliteten blir alvorlig ved ikke å sesongjustere. Dette er dermed nok en serie som har hatt et sesongmønster (klare topper for 2. kvartal), men der dette mønsteret nå kan være noe endret eller ha falt bort. Serien sesongjusteres.

R: Kulturell virksomhet

Alle testene for sesong i rådataene gir negativt resultat (p -verdi=1 for test 1). Kvaliteten blir god ved ikke å sesongjustere, og det er ingen indikasjon på gjenværende sesong. Serien sesongjusteres ikke.

S: Personlig tjenesteyting

Test 1 for sesong i rådataene gir negativt resultat (p -verdi=0,06), mens de andre er usikre, positive eller negative. Dersom vi utelater sesongkomponenten, altså lar være å sesongjustere, kan det være noe gjenværende sesong i serien, også for de tre siste årene, men kvaliteten blir likevel god. Ved å sesongjustere får vi usikker kvalitet knyttet til høy M -diagnostikk, der $M7=0,73$, $M1=3$ og $Q=1,07$. Det er altså stor grad av irregularitet i serien. Vi velger å sesongjustere serien for å unngå gjenværende sesong, og testen for identifiserbarhet for sesongmønsteret gir usikkert resultat, noe som indikerer at det kan være i orden å sesongjustere. Serien sesongjusteres.

3.3. Ansatte

I og med at vi for «ansatte» har korrigert alle delseriene for nivåskift ved innføring av A-ordningen betrakter vi den lineariserte serien når vi ser etter sesongeffekter.

A: Jordbruk, skogbruk og fiske

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive. Fra og med 2015 er det større variasjoner innen hvert kalenderår, enn det som er tilfelle til og med 2014. For denne næringen har altså innføringen av A-ordningen ført til et endret sesongmønster, og vi tar hensyn til dette ved å inkludere sesongoutliere for 1. og 2. kvartal 2015 i regARIMA modellen. Både

SO1k15 og SO2k15 er signifikante på 0,1% nivå. Sesongmønsteret er identifiserbart. Serien sesongjusteres.

B: Bergverksdrift og utvinning

Test 1 for sesong i den lineariserte serien er usikker (p-verdi=0,02), mens de andre er negative. Testene for stabil sesong er negative, og det er ingen identifiserbar sesong i serien. For 1. kvartal 2011 er det et positivt nivåskift. Dette antas å komme hovedsakelig fra en overflytting av virksomheter fra næringsgruppe «C: Industri», og legges inn som et nivåskift i regARIMA modellen: LS1k11. Nivåskiftet knyttet til overgangen til A-ordningen er veldig lite. Serien sesongjusteres ikke.

C: Industri

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive. 1. kvartal 2011 er det et negativt nivåskift. Som nevnt over, antas dette å skyldes en overflytting av virksomheter til næringsgruppe «B: Bergverksdrift og utvinning», og det legges inn et nivåskift i regARIMA-modellen: LS1k11. Nivåskiftet knyttet til overgangen til A-ordningen er veldig lite. Serien sesongjusteres.

D+E: Kraft og vannforsyning

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. Serien sesongjusteres.

F: Bygge- og anleggsvirksomhet

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive. Sesongmønsteret endrer seg noe over tid (bevegelig sesong tilstede på 5 % nivå), men det er likevel klart identifiserbart. Nivåskiftet knyttet til overgangen til A-ordningen er signifikant på 0,1% nivå. Serien sesongjusteres.

G: Varehandel, reparasjon av motorvogner

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive, og sesongmønsteret er identifiserbart. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 0,1 % nivå. Serien sesongjusteres.

H: Transport og lagring

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive, og sesongmønsteret er identifiserbart. Nivåskiftet knyttet til overgangen til A-ordningen er veldig lite. Serien sesongjusteres.

I: Overnattings- og serveringsvirksomhet

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive, og sesongmønsteret er identifiserbart. Nivåskiftet knyttet til overgangen til A-ordningen er veldig lite. Serien sesongjusteres.

J: Informasjon og kommunikasjon

Test 1 for sesong i den lineariserte serien er negativ (p-verdi er 0,09), mens de fem andre er positive eller usikre. Sesongmønsteret er ikke identifiserbart. Dersom vi utelater sesongkomponenten, altså lar være å sesongjustere, kan det være noe gjenværende sesong i serien, muligens også i de tre siste årene, og kvaliteten blir alvorlig. Grunnet gjenværende sesong velger vi å sesongjustere serien, og får da at kvalitetsmålene M7 og Q ligger rett over 1. M1 er 1,3, og samlet kvalitet er usikker, grunnet høye M-mål. Særlig M8-M11 er høye, og vi ser også fra SI-ratio grafene at det er en del bevegelse i sesongkomponenten. Nivåskiftet knyttet til overgangen til A-ordningen er veldig lite. Serien sesongjusteres, men dette er et tvilstilfelle, særlig siden det ikke er anbefalt å sesongjustere når sesongmønsteret ikke er identifiserbart.

K: Finansiering og forsikringsvirksomhet

Testene for sesong i dataene er alle negative, og vi får ingen gjenværende sesong ved å la være å sesongjustere. Denne serien sesongjusteres ikke, men justeres for nivåskiftet ved overgangen til A-ordningen. Dette nivåskiftet er veldig lite, med en t-verdi på bare 0,19. Serien sesongjusteres ikke.

L: Omsetning og drift av fast eiendom

Test 1 for sesong i den lineariserte serien er usikker (p-verdi er 0,05), mens de fem andre er positive. Sesongmønsteret er identifiserbart. Da vi sjekket for outlierer fant JD+ midlertidige skift for 1. kvartal 2014 og for 4. kvartal 2016, samt sesongoutlier for 1. kvartal 2016. Seriens nivå er lavt og den har en svakt voksende trend; i 1. kvartal 2017 var det omtrent 25 000 ansatte i næringen mot snau 23 000 i 1. kvartal 2010. Variasjonene fra kvartal til kvartal er i tillegg små, så de midlertidige skiftene som fanges opp er dermed ikke noe vi trenger å forfølge nærmere. Når det gjelder sesongoutlieren vil vi derimot følge med på eventuelle endringer i sesongmønsteret ved den årlige gjennomgangen av sesongjusteringen, men ikke ta hensyn til denne nå. Både M7 og Q er lavere enn 1, mens M1 er på hele 2,6. Det er dermed en del irregularet i serien. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. Serien sesongjusteres.

M: Teknisk tjenesteyting

Test 1 (med p-verdi på 0,05) og to av de andre testene for sesong i den lineariserte serien er negative, mens de tre gjenværende er usikre. Sesongmønsteret er ikke identifiserbart. Dersom vi utelater sesongkomponenten, kan det derimot være noe gjenværende sesong i serien. De drøye to årene med data etter overgangen til A-ordningen ligger an til å ha et sesongmønster, mens serien til og med 2014 ikke har noe sesongmønster. Siden vi benytter hele serien i beregningen av sesongkomponenten, blir sesongjusteringen dermed en kombinasjon av en periode med sesongmønster og en periode uten. Vi trenger mer enn to år for å starte beregningen av sesongmønster fra 2015. Dette er dermed en serie som vi ønsker å følge særlig med på.

Vi velger å sesongjustere serien siden det ligger an til å være sesong fra og med 2015. M7 er 1,39, M1 er 1,27 og M8-M11 er alle over 2, mens de resterende M-målene er under 1. Q ligger på 1,02. Et alternativ hadde vært å betrakte 2010-2014 for seg, og da la være å sesongjustere denne perioden. For å få sesongjusterte tall fra og med 2015 ville vi da benyttet minst to foregående år, altså 2013 og 2014, for å ha en lang nok serie for sesongjustering. Vi gikk ikke nærmere inn på dette alternativet i denne omgang. Når vi sjekker for outlierer finner JD+ en additiv outlier for 4. kvartal 2015. Betrakter vi serien ser vi også en topp for 4. kvartal 2016. Outlieren ser dermed ut til å være knyttet til det mulige sesongmønsteret fra 2015, og vi velger å ikke ta hensyn til outlieren på nåværende tidspunkt. Nivåskiftet knyttet til overgangen til A-ordningen er lite og ikke signifikant på 5 % nivå. Serien sesongjusteres.

N: Forretningsmessig tjenesteyting

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive. Sesongmønsteret er identifiserbart, M7 er 0,49, Q er 0,45, mens M8-M11 ligger over 1. Når vi sjekker for outlierer finner JD+ et midlertidig skift for 1. kvartal 2010. Dette er altså første observasjon i serien. Ved å sjekke den tidligere omtalte KNR-kilden fant vi at det noe høye tallet for 1. kvartal 2010 ser ut til å stemme. Det var en nedgang i antall ansatte i forkant av 2010. Vi behandler den ikke som et midlertidig skift, og får dessuten god nok kvalitet ved å gjøre det på denne måten. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. Serien sesongjusteres.

O: Offentlig administrasjon

Test 1 for sesong i den lineariserte serien er negativ (p-verdi på 0,13), test 2 er usikker, mens de resterende er positive. Dersom vi utelater sesongkomponenten, kan det være noe gjenværende sesong i serien. Det er derimot ikke statistisk bevis for gjenværende sesong de siste tre årene ved å la være å sesongjustere. Dette peker i retning av sesongjustering for perioden til og med 2014, mens vi muligens ikke bør sesongjustere fra og med 2015. Det vet vi mer om etter hvert som serien blir lengre. Vi har i denne omgang valgt å sesongjustere hele serien, og får da en M7-verdi på 0,57, en Q-verdi på 0,75 og at det er usikkerhet rundt identifiserbarhet for sesongmønsteret.

Vi har pre-spesifisert en additiv outlier for 1. kvartal 2015, da antall ansatte for dette kvartalet skiller seg ut ved å være lavt i forhold til det nye nivået som etableres ved overgangen til A-ordningen. Tanken her er oppstartsproblemer ved denne overgangen. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. [Serien sesongjusteres.](#)

P: Undervisning

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive. JD+ finner en sesongoutlier for 3. kvartal 2015. Dette observeres i serien som kraftige fall i nivået for 3. kvartal i både 2015 og 2016, og sesongjusteringsprogrammet tolker dette som at serien har fått et nytt sesongmønster. En del lærere er ikke ansatt på sommeren, og dette har kommet tydeligere fram gjennom det nye datagrunnlaget i 2015. Ved å inkludere SO3k15 blir sesongkomponenten for hele serien beregnet etter det nyetablerte sesongmønsteret. Vi pre-spesifiserer sesongoutlieren i regARIMA-modellen. Sesongmønsteret er identifiserbart, og kvalitetsmålene er gode. M7 er kun 0,1 og Q er 0,34. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. [Serien sesongjusteres.](#)

Q86: Helsetjenester

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive. Sesongmønsteret er identifiserbart. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. [Serien sesongjusteres.](#)

Q87: Pleie- og omsorgstjenester i institusjon

Det er klare sesongvariasjoner i serien; alle testene for sesong i den lineariserte serien er positive. Sesongmønsteret er identifiserbart. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. [Serien sesongjusteres.](#)

Q88: Sosiale omsorgstjenester uten botilbud

Test 1 for sesong i den lineariserte serien er usikker (p-verdi=0,01), mens de fem andre er positive. Når vi sjekker for outlierer finner JD+ et midlertidig skift for 2. kvartal 2010, et nivåskift for 3. kvartal 2014, samt additive outlierer for 3. og 4. kvartal 2015. Inkludering av disse fører til dårlig kvalitet, grunnet det høye tallet på inkluderte outlierer og at flere av M-målene er godt over 1. Disse outlierene er valgt ut med tanke på beste tilpasning til en regARIMA-modell, og i dette tilfellet fikk den valgte modellen for mange outlierer i regresjonsdelen. Vi ser ingen grunner for å inkludere noen av outlierene, men observerer at et tydeligere sesongmønster ser ut til å etablere seg etter overgangen til A-ordningen. Dette kan antyde bedret kvalitet etter hvert som vi får flere observasjoner.

Sesongmønsteret for hele serien er ikke identifiserbart, da det endrer seg mye over tid. Dersom vi utelater sesongkomponenten, kan det være noe gjenværende sesong i serien. Vi velger å sesongjustere serien selv om kvaliteten er usikker, og vil følge med på utviklingen av sesongmønsteret ved de årlige gjennomgangene.

Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. Serien sesongjusteres.

R: Kulturell virksomhet

Alle testene for sesong i den lineariserte serien er positive, unntatt test 2 som er usikker. Sesongmønsteret er identifiserbart. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. Serien sesongjusteres.

S: Personlig tjenesteyting

Alle testene for sesong i den lineariserte serien er positive. Når vi sjekker for outliers finner JD+ en additiv outlier for 4. kvartal 2015. 4. kvartal viser en topp for flere år, og observasjonen for 2015 er dermed ikke en ekstrem som vi trenger å bry oss om. Sesongmønsteret er identifiserbart og kvalitetsmålene er gode, kun M1 er høyere enn 1. Nivåskiftet knyttet til overgangen til A-ordningen er klart signifikant på 1 % nivå. Serien sesongjusteres.

3.4. Årlig gjennomgang av seriene

For alle serier vil vi ved den årlige gjennomgangen gjøre en sesongjustering fra «scratch», der det testes for om dataene skal log-transformeres, og der beste regARIMA-modell og beste sesong- og trendfilter velges. Dette gjøres ved å endre på spesifikasjonene for hver serie inne i JD+ slik at disse valgene vil bli gjort automatisk. Nærmere beskrivelser av disse spesifikasjonsvalgene gis i intern dokumentasjon. For noen av seriene kan det være aktuelt å sjekke for outliers, men vi vil bare inkludere nye outliers i regresjonsdelen dersom vi har en rimelig forklaring på disse. Man kan enkelt la JD+ se etter outliers automatisk, ved å gjøre et valg om dette i spesifikasjonene. For flere av seriene vil vi, så sant kvaliteten fremdeles er god, ikke foreta oss noe mer enn å låse de sesong- og trendfiltrene som har blitt valgt, mens ARIMA-modell og evt. log-transformering holdes fast via sesongjusteringsprosedyrene i JD+ uten at spesifikasjonene må hardkodes. Låste sesong- og trendfiltre er anbefalt praksis ifølge 2015 versjonen av ESS retningslinjene for sesongjustering (ESS, 2015). Dette var ikke med i forrige versjon av retningslinjene, og er ikke implementert i automatisk delvis løpende korrigering i JD+. Filterlengdene må derfor inntil videre hardkodes for hver tidsserie.

For enkelte av seriene for «ansatte» har vi i gjennomgangen i kapittel 3.3 kommentert at dette er serier vi vil følge særlig med på fordi innføringen av A-ordningen 1. kvartal har ført til at sesongmønsteret har endret seg, ligger an til å utebli eller til å ha oppstått fra og med dette tidspunktet. Her vil vi i første omgang følge med på utviklingen, og vil ta for oss spørsmål som: Er sesongmønsteret identifiserbart? Finner vi gjenværende sesong i de tre siste årene dersom vi ikke sesongjusterer? Er det aktuelt å legge inn sesongoutliere for noen kvartal? Vi kan også teste en avgrensning av serien til de 4-5 siste årene for å fokusere på kvalitet og sesongmønster etter innføringen av A-ordningen. Det er viktig at vi ikke tar forhastede beslutninger når det gjelder å gå fra sesongjustering til ikke-sesongjustering av en serie, og motsatt.

Om noen år vil vi også vurdere å avgrense serien for sesongjustering for å unngå å ha data særlig for «ansatte» både før og etter innføring av A-ordningen, og kun oppdatere statistikkbanken med tall etter innføring av A-ordningen. Fra og med 1. kvartal 2019 vil vi ha fire hele år etter innføringen av A-ordningen, som er minimum lengde for sesongjustering av en serie.

3.5. Sesongjusterte totaltall

Også totaltallene for antall ledige stillinger og antall ansatte sesongjusteres. Dette skjer som nevnt indirekte, ved at de 20 næringsgruppene summeres. Ved

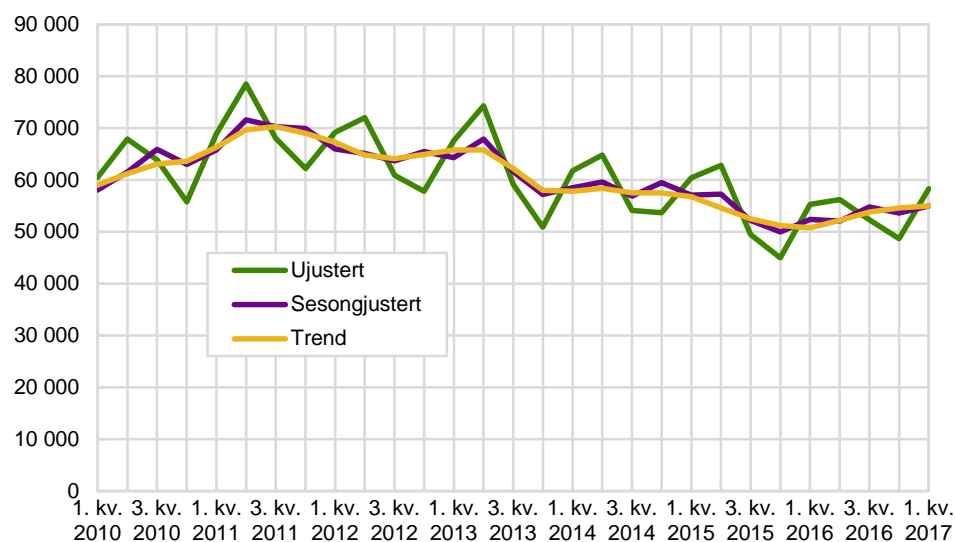
publisering ligger hovedfokuset på antall ledige stillinger, der vi nå fokuserer på det sesongjusterte tallet, mens prosentdelen ledige stillinger hovedsakelig er av interesse for internasjonal sammenligning. Ved strukturendringer i arbeidsmarkedet som vi nok kommer til å fortsette å se mer av framover, kan også prosentdelen ledige stillinger fordelt etter næring bli av større interesse da den er med på å nyansere tallet på ledige stillinger innen hver næringsgruppe.

I det store og det hele er det utviklingen over tid som er av størst interesse ved en statistikk som denne, og da er trenden den viktigste serien. Når vi følger serien fra kvartal til kvartal ligger fokuset på den sesongjusterte serien.

Totaltall for antall ledige stillinger

I figur 3.5 ser vi avrundede tall til nærmeste hundre for totalen av ledige stillinger, ujustert, sesongjustert og trend. Her ser vi tydelig at trenden er en «glattere» versjon av serien enn den sesongjusterte serien, som i tillegg til trenden også inneholder den irregulære komponenten.

Figur 3.5 Antall ledige stillinger totalt. Tallene er avrundet til nærmeste hundre

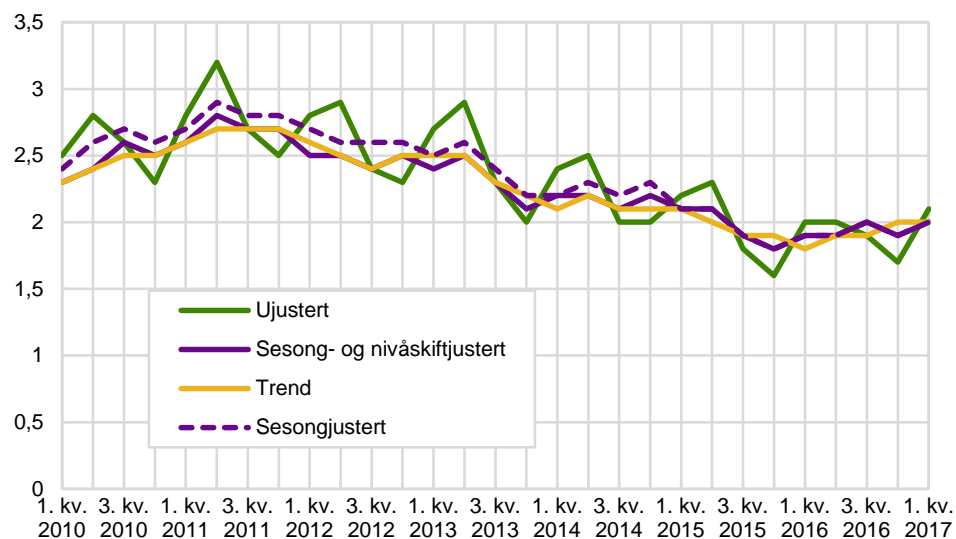


Kilde: Statistisk sentralbyrå.

Totaltall for prosentdelen ledige stillinger

Figur 3.6 viser avrundede tall til én desimal for prosentdelen ledige stillinger. Til og med 2014 ligger den ujusterte serien på et høyere nivå enn den sesong- og «brudd» justerte serien og trenden. Dette skyldes at den ujusterte serien for «ansatte» ikke er korrigert for bruddet ved innføringen av A-ordningen 1. kvartal 2015. Den stiplede linjen viser den sesongjusterte serien dersom vi ikke hadde korrigert den sesongjusterte «ansatte»-serien for nivåskiftet ved innføringen av A-ordningen, men der vi kun hadde pre-korrigert for nivåskiftet for selve oppspaltingen av serien. Vi ser at forskjellen mellom de to sesongjusterte seriene er markant, og at det for sammenligning over tid er vesentlig å justere for bruddet.

Figur 3.6 Prosentdelen ledige stillinger totalt



Kilde: Statistisk sentralbyrå.

Referanser

ESS (2015). *ESS guidelines on seasonal adjustment*, Edition 2015, Eurostat (<http://ec.europa.eu/eurostat/documents/3859598/6830795/KS-GQ-15-001-EN-N.pdf>)

Grudkowska, S. (2017). *JDemetra+ Reference Manual Version 2.2*. (https://ec.europa.eu/eurostat/cros/system/files/jdemetra_reference_manual_version_2.2_0.pdf)

Office for National Statistics (2007). *Guide to Seasonal Adjustment with X-12-ARIMA DRAFT* (<http://www.ssb.no/a/kortnavn/sykefratot/guide-to-seasonal-adjustment.pdf>)

Rein, K. (2001). Om muligheten for å predikere norsk inflasjon ved hjelp av ARIMA-modeller (<http://www.uib.no/filearchive/kjell-arild-rein.pdf>)

Vedlegg A: Tabeller

Tabell A1.1 og A1.2 gir informasjon om valgte regARIMA-modell, sesongfilter og Henderson trendfilter, om sesongmønsteret er identifiserbart, verdiene på M7 og Q, og samlet kvalitet, for respektivt «ledige stillinger» og «ansatte».

Tabell A1.1 Ledige stillinger

	ARIMA modell	Ses.filter	Hend filter	Identifiserbar sesong?	M7	Q	Kvalitet
A	(1,0,0)(0,1,1)	3x5	5	Usikkert	0,639	1,011	God
B	(0,1,0)(1,0,0)	3x3 (valgt)	5	Nei	1,328	1,057	God
C	(0,1,2)(1,0,0)	3x9	7	Ja	0,368	0,897	God
D+E	(1,0,1)(0,0,0)	-	7	-	-	1,035	Usikker
F	(0,0,0)(1,1,0)	-	5	-	-	1,203	God
G	(0,0,0)(0,1,1)	3x5	5	Nei	0,834	0,783	God
H	(1,0,1)(0,0,0)	3x5	5	Usikkert	0,867	0,687	God
I	(0,0,0)(0,1,1)	3x5	5	Ja	0,421	0,571	God
J	(1,0,0)(1,0,0)	3x5	5	Ja	0,619	0,590	God
K	(1,0,0)(0,0,1)	3x3	5	Ja	0,651	0,883	God
L	(1,0,0)(0,0,0)	-	5	-	-	0,663	God
M	(0,0,0)(0,0,0)	-	7	-	-	1,326	Usikker
N	(0,1,0)(0,1,1)	3x5	5	Ja	0,481	0,552	God
O	(0,1,2)(0,0,0)	3x9	5	Usikkert	0,684	0,750	God
P	(0,0,0)(0,1,1)	3x9	7	Ja	0,144	0,563	God
Q86	(0,0,0)(1,1,0)	-	5	-	-	1,000	God
Q87	(1,0,0)(0,1,1)	3x9	5	Nei	0,982	1,023	God
Q88	(1,0,1)(0,0,1)	3x5	7	Usikkert	0,609	1,075	Usikker
R	(0,1,1)(0,0,0)	-	7	-	-	1,053	God
S	(0,0,0)(0,1,1)	3x9	5	Usikkert	0,726	1,073	Usikker

Kilde: Statistisk sentralbyrå.

Tabell A1.2 Ansatte

	ARIMA modell	Ses.filter	Hend filter	t-verdi for LS1k15	Outliere	Identi.ses.?	M7	Q	Kvalitet
A	(1,0,0)(0,1,0)	3x3	5	t=-1,78	SO1k15(t=6,31), SO2k15(t=4,13)	Ja	0,126	0,177	God
B	(0,2,1)(0,0,0)	-	5	t=0,17	LS1k11(t=9,01)	-	-	0,088	God
C	(1,1,0)(0,1,0)	3x3	5	t=1,01	LS1k11(t=-5,46)	Ja	0,292	0,363	God
D+E	(2,1,0)(0,1,0)	3x3	5	t=3,39		Ja	0,322	0,274	God
F	(0,1,0)(0,1,1)	3x3	5	t=-8,45		Ja	0,432	0,419	God
G	(0,0,0)(0,1,0)	3x3	5	t=7,62		Ja	0,456	0,763	God
H	(1,0,0)(0,1,0)	3x5	5	t=1,32		Ja	0,484	0,672	God
I	(0,1,1)(0,1,0)	3x3	5	t=0,61		Ja	0,539	0,498	God
J	(1,1,0)(0,1,1)	3x5	5	t=0,26		Nei	1,060	1,020	God
K	(1,1,0)(0,0,0)	-	7	t=1,17		-	-	0,864	God
L	(0,0,0)(1,1,0)	3x5	5	t=-8,10		Ja	0,636	0,871	God
M	(1,1,0)(0,1,0)	3x5	5	t=-1,00		Nei	1,387	1,020	Usikker
N	(0,1,1)(0,1,0)	3x3	5	t=-6,02		Ja	0,487	0,454	God
O	(1,2,1)(0,1,1)	3x9	5	t=24,73	AO1k15(t=-12,81)	Usikkert	0,571	0,747	God
P	(1,0,1)(0,1,1)	3x5	5	t=32,60	SO3k15(t=8,92)	Ja	0,101	0,336	God
Q86	(0,1,0)(0,1,1)	3x5	5	t=10,51		Ja	0,219	0,180	God
Q87	(0,0,0)(0,1,0)	3x5	5	t=21,31		Ja	0,320	0,434	God
Q88	(2,1,0)(0,1,0)	3x3	7	t=27,80		Nei	1,168	1,142	Usikker
R	(0,1,0)(1,0,1)	3x5	5	t=11,22		Ja	0,407	0,457	God
S	(1,1,0)(0,1,1)	3x5	5	t=8,08		Ja	0,550	0,692	God

Kilde: Statistisk sentralbyrå.

Statistisk sentralbyrå

Postadresse:
Postboks 8131 Dep
NO-0033 Oslo

Besøksadresse:
Akersveien 26, Oslo
Oterveien 23, Kongsvinger

E-post: ssb@ssb.no
Internett: www.ssb.no
Telefon: 62 88 50 00

ISBN 978-82-537-9623-9 (elektronisk)



Statistisk sentralbyrå
Statistics Norway