

Lunds universitet
Nationalekonomiska institutionen
Magisteruppsats VT 2015

En modell för prognoser på offentlig konsumtion
– tidsserieanalys och prognosutvärdering

Författare: Anna Fahlén
Handledare: Lina Maria Ellegård

Sammanfattning

Den här uppsatsen utreder möjligheten att använda en modell som skattats med tidsserieanalys i Konjunkturinstitutets (KI) prognoser på offentlig konsumtion. Autoregressiva modeller skattas på realtidsdata från 1994. Modellerna skattas och utvärderas med out-of-sample prognoser. Modellernas prognosförmåga utvärderas mot första publiceringen av statistik för offentlig konsumtion för perioden 2004–2013. Modellprognoserna utvärderas mot KI:s förmåga att göra träffsäkra prognoser. Det utvärderingsmått som används är relativ RMSE. Modellernas och KI:s prognosfel jämförs även med Diebold-Marianos test för lika prognosförmåga. Prognosernas snedvridning utvärderas genom att jämföra modellernas och KI:s bias. Resultaten visar att det är möjligt att använda resultaten från en tidsseriemodell i prognoserna med hjälp av tidsserieanalys.

Nyckelord: Out-of-sample prognoser, offentlig konsumtion, autoregressiva modeller, relativ RMSE, Diebold-Mariano.

Innehållsförteckning

1.	Inledning	4
1.1.	Disposition	4
2.	Offentlig konsumtion	6
2.1.	Offentlig konsumtion i BNP	7
2.2.	Offentlig konsumtion i den Offentliga sektorns finansiella sparande	8
3.	Prognoser på offentlig konsumtion	9
3.1.	Olika metoder i preliminära och definitiva beräkningar	10
3.2.	Brist på månadsstatistik	11
3.3.	Allmän översyn av nationalräkenskaperna revideraras vart femte år	11
3.4.	Offentlig konsumtion redovisas per transaktion	11
3.5.	Konjunkturinstitutets prognoser på offentlig konsumtion	12
4.	Data	14
4.1.	Inga märkbara effekter av finanskrisen för offentlig konsumtion	14
5.	Metod	16
5.1.	Tidsserieanalys	16
5.1.1.	AR(1)-modellen	17
5.1.2.	AR(p)-modellen	17
5.1.3.	Stationäritet	18
5.1.4.	Metod för out-of-sample prognoser	19
5.2.	Metoder för prognosutvärdering	21
5.2.1.	Relativ RMSE	21
5.2.2.	Diebold-Marianos test för lika prognosförmåga	22
5.2.3.	Bias	23
5.2.4.	Rangordning av utvärderingsmått	23
6.	Resultat	25
6.1.	Modellen har bättre prognosförmåga än KI i decemberprognoserna	26
6.2.	Modellerna har större prognosfel än KI i mars-, juni- och augustiprognoserna	28
6.2.1.	Modellernas prognosförmåga i mars	28
6.2.2.	Modellernas prognosförmåga i juni	29
6.2.3.	Modellernas prognosförmåga i augusti	30
7.	Slutsatser	32
7.1.	Små skillnader mellan modellerna	32
7.2.	Modellernas relativa prognosförmåga försämras mellan mars och augusti	33
7.3.	Offentlig konsumtion underskattas både av modellen och av KI	33
7.4.	Modellen kan bidra till bättre prognoser på offentlig konsumtion	34
7.5.	Användning av modellen i KI:s prognoser	34
	Referenser	36
	Appendix	38

1. Inledning

Konjunkturinstitutet (KI) är en statlig myndighet som bland annat gör prognoser på den ekonomiska utvecklingen i Sverige och internationellt. KI:s prognoser används som beslutsunderlag för den ekonomiska politiken i Sverige. Men även av andra statliga myndigheter, företag, organisationer och av arbetsmarknadens parter. KI:s prognoser är så kallade bedömningsprognoser där den senast publicerade statistiken för respektive prognosvariabel analyseras tillsammans med annan relevant statistik. I prognoserna tas även hänsyn till information från omvärldsbevakning som påverkar utvecklingen av en viss variabel. I många prognoser på KI används även modeller som stöd i prognosarbetet. Syftet med att använda modeller är att systematisera information som är av värde i prognosarbetet och inte att en modell ska göra hela prognosarbetet. Prognosmakare talar ofta om att man använder ett *modellstöd* i prognoserna. Modeller kan baseras på månadsdata för att säga något om en variabels kvartalsutveckling eller så kan modellen innehålla information om en eller flera andra variabler som säger något om hur en variabel utvecklas. Om man vet att modellen historiskt sett har gjort bra prognoser på en viss variabel så finns ett bra stöd för prognosmakaren att använda modellens resultat i prognosarbetet. Resultatet från en modell är alltså bara en av många faktorer som vägs samman med annan information för att komma fram till prognosen för en viss variabel.

Den offentliga sektorns konsumtionsutgifter är en av många variabler i KI:s prognoser och det finns i dagsläget inte någon modell som är framtagen för att användas som stöd i den prognosen. *Syftet* med den här uppsatsen är därför att undersöka om det finns något värde i att använda en enkel autoregressiv tidsseriemodell i prognosen på offentlig konsumtion. För att undersöka detta är *frågeställningen* i uppsatsen därför om en tidsseriemodell kan göra bättre prognoser på offentlig konsumtion än KI. Om modellen gör lika bra eller bättre prognoser än KI så tyder det på att en modell har något att tillföra i prognosarbetet. I uppsatsen skattas en tidsseriemodell på data över offentlig konsumtion på kvartal. Modellens kvartalsprognoser summeras till årsprognoser som utvärderas mot den statistik för offentlig konsumtion på år som Statistiska centralbyrån (SCB) har publicerat. För att få reda på om modellens prognoser är bättre än KI:s utvärderas modellprognoserna mot KI-prognoserna. De huvudsakliga resultaten visar att det är möjligt att använda tidsseriemodeller som stöd i prognosen på offentlig konsumtion.

1.1. DISPOSITION

Resterande delar av uppsatsen är strukturerad enligt följande. Kapitel 2 beskriver definitionen av offentlig konsumtion och var i den ekonomiska statistiken den redovisas. Kapi-

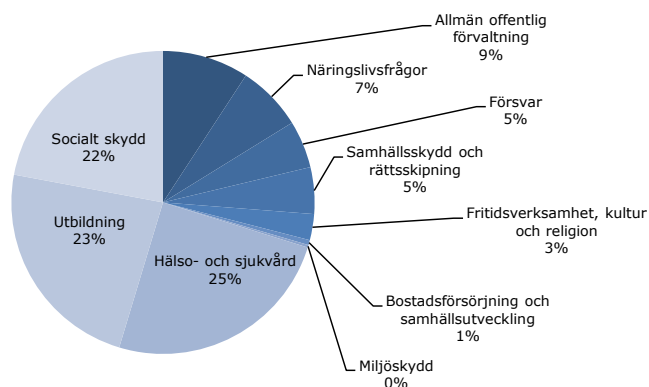
tel 3 belyser svårigheter med att göra prognoser på offentlig konsumtion samt förklarar hur KI gör prognoser. Kapitel 4 beskriver den data som används för att skatta modellerna. I kapitel 5 redogörs för vilken metod som används för att skatta och utvärdera prognosmodellen. Resultaten redovisas och analyseras i kapitel 6. Slutsatserna presenteras i kapitel 7.

2. Offentlig konsumtion

Definitionen av offentlig konsumtion är summan av alla utgifter som offentlig sektor har för att tillhandahålla offentligt finansierade tjänster (Eurostat, 2013). Den offentliga sektorn består av delsektorena stat, primärkommuner, landsting och ålderspensionssystem. Enligt statistik från SCB:s nationalräkenskaper sker två tredjedelar av den offentliga konsumtionen i primärkommuner och landsting och en tredjedel av konsumtionen i staten. Ålderspensionssystemets konsumtion är mycket marginell. Den offentliga konsumtionen består till 70 procent av individuell konsumtion, så som utbildning, hälso- och sjukvård samt äldreomsorg. Resterande del av den offentliga konsumtionen är kollektiv och består bland annat av allmän offentlig förvaltning, försvar, samhällsskydd och rättsskipning, näringslivsfrågor (se diagram 1). Denna indelning avser den offentliga konsumtionens *ändamål*. Ändamålen beskriver vilken typ av verksamhet som den offentliga sektorn har konsumtionsutgifter för.

Diagram 1 Offentlig konsumtion fördelad per ändamål, 2012

Procent av total offentlig konsumtion



Källa: SCB.

I nationalräkenskaperna redovisas även den offentliga konsumtionen i termer av vilka typer av *transaktioner* som konsumtionen avser (se tabell 1). Denna indelning är mer redovisningsorienterad och konsumtionen delas då in i förädlingsvärde, förbrukning, sociala naturaförmåner, försäljning och egenproducerade tillgångar. Förädlingsvärdet motsvarar de utgifter som offentlig sektor har för löner, arbetsgivaravgifter och produktionskostnader. Förbrukning är alla andra utgifter såsom hyror och inköp av material som offentlig sektor har för att kunna tillhandahålla offentligt finansierade tjänster. Sociala naturaförmåner är de utgifter som offentlig sektor har för offentligt finansierade tjänster men som utförs av en privat aktör. Som exempelvis när ett privat företag driver en gymnasieskola men verksamheten finansieras av kommunen. Försäljning är en avdragspost och uppstår

exempelvis om ett privat företag låter kommunen utföra snöröjning på sin parkering och betalar för det. Egenproducerade tillgångar är också en avdragspost och uppkommer exempelvis om en myndighet utvecklar en programvara som säljs vidare till ett privat företag.

Tabell 1 Offentlig konsumtion fördelad per transaktion, 2012

Miljarder kronor

	2012
Förädlingsvärde	668
Förbrukning	344
Sociala naturaförmåner	125
Försäljning (-)	146
Egenproducerade tillgångar (-)	35
Offentlig konsumtion	955

Anm. Tabellen summerar inte pga. avrundning.

Källa: SCB.

2.1. OFFENTLIG KONSUMTION I BNP

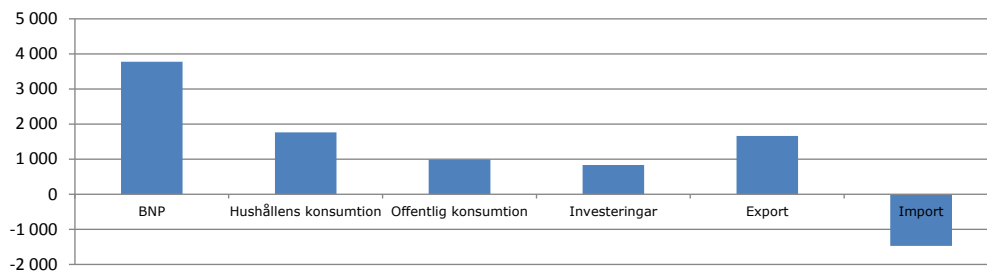
BNP är ett mått på ekonomisk aktivitet i ett land och definieras för en tidsperiod, vanligtvis ett år eller ett kvartal (Lequiller och Blades, 2014). Uppgifter om BNP används dels för att mäta hur den ekonomiska aktiviteten utvecklas i ett land dels för att jämföra ekonomisk aktivitet mellan olika länder. Det finns flera sätt att definiera BNP. De två vanligaste sätten är att mäta summan av alla utgifter alternativt summan av all produktion i ett land. BNP kan även definieras som summan av alla inkomster som genereras i ett land. Den offentliga konsumtionen är en delkomponent i BNP från utgiftssidan som ofta benämns försörjningsbalansen och definieras som:

$$Y = C + G + I + X - M \quad (1)$$

Där Y står för BNP, C för hushållens konsumtion, G för offentlig konsumtion, I för investeringar, X för export och M för import. Offentlig konsumtion uppgick 2013 till 26 procent av Sveriges BNP (se diagram 2). SCB sammanställer och redovisar nationalräkenskaperna, där BNP ingår i produkträkenskaperna.

Diagram 2 BNP och försörjningsbalansens delar, 2012

Miljarder kronor



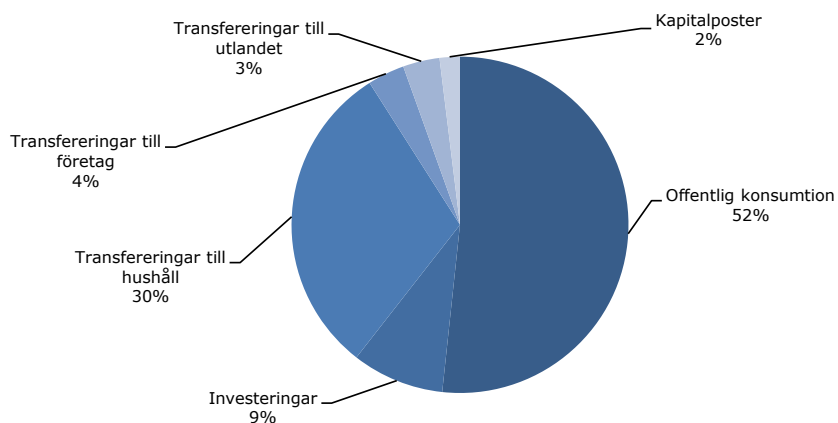
Källa: SCB.

2.2. OFFENTLIG KONSUMTION I DEN OFFENTLIGA SEKTORNS FINANSIELLA SPARANDE

Utöver produkträkenskaperna sammanställer nationalräkenskaperna även sektorräkenskaperna. Där redovisas alla transaktioner som sker mellan sektorerna i ekonomin. Ekonomin delas in i offentlig sektor, hushålls, företags- och utlandssektorn. Sektorernas inkomster och utgifter redovisas i syfte att bland annat få fram det finansiella sparandet i respektive sektor. Offentlig konsumtion är en del av den offentliga sektorns utgifter och är därför en komponent i den offentliga sektorns finansiella sparande. Definitionen av finansiellt sparande är offentliga inkomster minus offentliga utgifter. De offentliga utgifterna består även av investeringar, transfereringar till hushåll, företag och utlandet samt kapitalutgifter. Offentlig konsumtion uppgick 2013 till 52 procent av de offentliga utgifterna (se diagram 3). Den offentliga sektorns finansiella sparande är ett intressant mått eftersom ett av de budgetpolitiska målen, överskottsmålet, är att det finansiella sparandet i genomsnitt ska vara 1 procent av BNP över en konjunkturcykel.

Diagram 3 Offentliga sektorns utgifter, 2013

Procent av offentliga sektorns totala utgifter



Källa: SCB.

3. Prognoser på offentlig konsumtion

I det här kapitlet beskrivs vilka svårigheter som finns i prognoserna på offentlig konsumtion samt hur KI gör dessa prognoser. Syftet med kapitlet är att ge en bild av vilka utmaningar man ställs inför när man gör prognoser på variabeln offentlig konsumtion och vad som utmärker detta område. Uppsatsen syftar till att undersöka om en tidsseriemodell kan bidra till bättre prognoser på offentlig konsumtion. En prognosmodell löser emellertid inte de problem som beskrivs här men kan om modellen har en god prognosförmåga hjälpa till att öka träffsäkerheten i prognoserna.

Ett av de mest centrala problemen med att göra prognoser på offentlig konsumtion är att statistiken revideras relativt mycket. Med revidering menas hur mycket en viss variabel ändras mellan att första och senaste utfall publiceras. Med första utfall avses den statistik som publicerades första gången statistiken sammanställdes. Med senaste utfall avses den statistik som har publicerats senast, detta utfall kan innehålla mer information än det första utfallet då SCB kan ha fått in mer information som är av värde för beräkningarna.

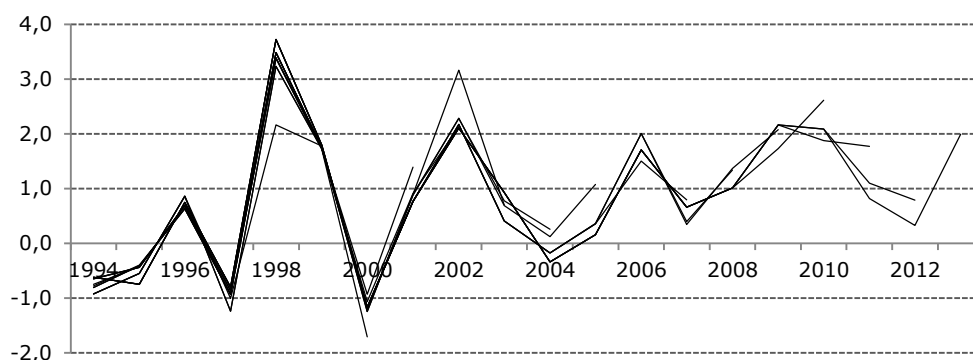
diagram 4 visar hur mycket tidsserien för offentlig konsumtion har reviderats för perioden 1994–2013. Varje linje i diagrammet motsvarar hur tidsserien såg ut vid det tillfälle då statistiken publicerades. Om statistiken för offentlig konsumtion har reviderats ett visst år ligger linjen på flera olika nivåer i diagrammet. Ett exempel är 2012 som det finns två linjer för, vilket innebär att det året har offentlig konsumtion reviderats två gånger hittills. Första gången statistiken publicerades ökade offentlig konsumtion med 0,3 procent jämfört med 2011. När de reviderade uppgifterna om 2012 publicerades ökade offentlig konsumtion med 1,1 procent jämfört med 2011.

I utredningen om översyn av den ekonomiska statistiken (SOU2002:118, 2002) framhölls att studier av revideringar i nationalräkenskaperna visar att de första preliminära utfallen för offentlig konsumtion är skakiga och ibland kan följas av stora revideringar. Problemen spände då över den offentliga konsumtionen som helhet, både vad gäller statlig och kommunal konsumtion. Utredningen pekade på att offentlig konsumtion tillhörde ett av flera förbättringsområden inom den ekonomiska statistiken. Sedan utredningen presenterade sitt slutbetänkande har beräkningarna av den offentliga konsumtionen genomgått flera metodförändringar i syfte att förbättra statistiken och minska revideringarna. Statistiken revideras emellertid fortfarande relativt mycket och orsaken till dessa revideringar

beror på flera olika saker. En orsak är att SCB:s beräkningsmetoder skiljer sig åt mellan preliminära och definitiva beräkningar. En annan orsak är att beräkningarna ses över och förbättras vart femte år. Som en konsekvens av att samhället och ekonomin förändras gör även Eurostat ändringar i det internationella regelverk där alla definitioner anges.

Diagram 4 Revideringar av offentlig konsumtion, 2004–2013

Procentuell förändring



Anm. Diagrammet visar tidsserier för offentlig konsumtion som publicerats i samband med att SCB redovisat nationalräkenskaper för fjärde kvartalet. Första tidsserien innehåller 1994–1999 och senaste tidsserien innehåller 1994–2013.

Källa: SCB.

3.1. OLIKA METODER I PRELIMINÄRA OCH DEFINITIVA BERÄKNINGAR

En förklaring till revideringarna är att det tar lång tid innan SCB får tillgång till ett fullständigt underlag för att kunna sammanställa den offentliga konsumtionen. Den första preliminära versionen av statistiken redovisas 60 dagar efter det kvartal som avses. I samband med varje kvartalsberäkning uppdateras tidigare publicerade uppgifter för innevarande år. Vid publicering av det första kvartalet varje år revideras även föregående års kvartalsutfall. De definitiva årsberäkningarna redovisas i september året efter respektive år, dvs. knappt två år efter det år som avses. När de definitiva beräkningarna redovisas brukar statistiken revideras vilket beror på att beräkningsmetoderna skiljer sig åt. Definitiva beräkningar av offentlig konsumtion baseras på detaljerad information från stat, primärkommuner och landsting som gör det möjligt att beräkna konsumtion i både fasta och löpande priser. I preliminära beräkningar finns en mycket begränsad mängd information tillgänglig. Konsumtionen i löpande priser baseras på ett underlag som täcker staten och landstingen samt de större primärkommunerna. För att beräkna konsumtionen i fasta priser finns det, på ett fåtal områden, uppgifter om volymutveckling men i huvudsak utgör utfallen i praktiken prognoser. Exempelvis görs prognoser på antalet elever i skolan det aktuella året. För att få fram fasta priser sker en kvartalsfördelning av konsumtionen med uppgifter om antalet arbetade timmar inom skolan. Följderna av detta är att utfallet

för den offentliga konsumtionen vanligtvis revideras flera gånger mellan att det första preliminära utfallet redovisas och det definitiva utfallet redovisas. Det kan vara bra att komma ihåg att senaste utfall inte är samma som det definitiva utfallet för alla år i en tidsserie. Åren före 2012 har genomgått en definitiv årsberäkning medan åren därefter bara är summeringar av årens preliminära kvartalsberäkningar.

3.2. BRIST PÅ MÅNADSSTATISTIK

På många prognosområden såsom prognoser på hushållens konsumtion och utrikeshandeln med varor och tjänster finns statistik med månadsfrekvens som har stort informativt värde i prognosarbetet. Eftersom all offentlig statistik redovisas med så stor fördröjning så finns det ingen sådan månadsstatistik som kan användas i prognosen på offentlig konsumtion.

3.3. ALLMÄN ÖVERSYN AV NATIONALRÄKENSKAPERNA REVIDERARAS VART FEMTE ÅR

Utöver de definitiva årsberäkningarna gör SCB även en allmän översyn av nationalräkenskaperna vart femte år, så kallad generalrevidering. I en allmän översyn tas hänsyn till ny information som tillkommit och eventuella felaktigheter som upptäckts sedan den senaste definitiva årsberäkningen gjordes. Nyligen gjordes även en stor förändring av flera definitioner i nationalräkenskaperna till följd av ändrat regelverk för redovisningen av nationalräkenskaperna. Regelverket fastställs av Eurostat och följs av alla EU-länder.

3.4. OFFENTLIG KONSUMTION REDOVISAS PER TRANSAKTION

I de preliminära beräkningarna redovisas offentlig konsumtion endast per transaktion. De definitiva årsberäkningarna innehåller däremot såväl transaktions- som ändamålsfördelad offentlig konsumtion men redovisas med nästan två års eftersläpning. KI gör därför prognoser på transaktionsfördelad offentlig konsumtion eftersom det är det som prognoserna kan utvärderas mot på kort sikt. Det finns både praktiska och pedagogiska problem med att göra prognoser per transaktion istället för på ändamål. Det innebär att det både är svårare att göra prognoser eftersom transaktionsindelningen säger väldigt lite om den faktiska verksamheten i offentlig sektor. Statistiken beskriver då exempelvis hur stor den offentliga sektorns konsumtion är i termer av förädlingsvärde, förbrukning och sociala natura förmåner. Men exempelvis ökningen av den offentliga sektorns konsumtion av sjukvård och utbildning framgår inte. En sådan uppdelning av den offentliga konsumtionen kommer först efter två år. Om statistiken skulle varit uppdelad på ändamål i stället för på transaktioner skulle det varit enklare att förklara och förstå vad utvecklingen av offentlig konsumtion beror på, både för SCB och för användarna av statistiken.

3.5. KONJUNKTURINSTITUTETS PROGNOSE PÅ OFFENTLIG KONSUMTION

Trots att den definitiva beräkningen av offentlig konsumtion bygger på långt mycket mer detaljerad information gör KI prognos på den första och preliminära statistiken som publiceras för en viss variabel. Det beror på att när de definitiva beräkningarna redovisas så har det året passerat för längesedan och eventuella beslut om finans- och penningpolitik har redan fattats och genomförts.

KI gör prognoser både på kort och medellång sikt. Med kort sikt avses innevarande och nästkommande år. Medellång sikt innebär att prognoser görs för ytterligare tre år framåt i tiden. Det huvudsakliga syftet är att göra årsprognoser men för att nå fram till dessa görs i de allra flesta fall även kvartalsprognoser.

För offentlig konsumtion görs en prognos på helåret för delsektorena stat, primärkommuner och landsting. Helårsprognosen fördelas på de kvarvarande kvartalen innevarande år och kvartalen nästkommande år. Prognosen fördelas även på transaktioner för att ta fram en prognos på förädlingsvärde, förbrukning, sociala naturaförmåner, försäljning och egenproducerade tillgångar. Prognosen på den offentliga sektorns transaktioner används delvis i andra prognoser på KI.

I de kortsiktiga prognoserna på offentlig konsumtion görs bedömningsprognoser där flera olika underlag vägs samman. Kvartalsstatistik analyseras, regeringens beslut om finanspolitiska åtgärder vägs in och en bedömning av det demografiska behovet av offentligt finansierade tjänster görs. Hänsyn tas även till andra förändringar som har inverkan på den offentliga konsumtionens utveckling. Den modell som kan användas som stöd i de kortsiktiga prognoserna är den demografiska modell, DEMOG, som egentligen är avsedd att användas för prognoser på medellång sikt. DEMOG är en modell som väger samman kostnaderna per individ för hälso- och sjukvård, utbildning och socialt skydd med den demografiska utvecklingen enligt SCB:s befolkningsprognos. Något förenklat antar modellen att konsumtionen av välfärdstjänster per individ är konstant över tid men att befolkningens sammansättning förändras. Resultatet av detta används för att skriva fram individuell konsumtion i prognoserna på medellång sikt. Den kollektiva delen av offentlig konsumtion skrivs fram med utvecklingen för hela befolkningen.

Som nämndes tidigare finns det inom många prognosområden tillgång till månadsstatistik som innehåller information som säger något om hur en variabel kommer att utvecklas under exempelvis ett kvartal. Ingen sådan månadsstatistik finns tillgänglig som kan användas för analyser av offentlig konsumtion. Ett annat sätt att ta in information i prognos-

sen är att se vad olika modeller ger för resultat. Inom flera prognosområden på KI används tidsseriemodeller som stöd i prognosarbetet. Dessa modeller kan antingen vara baserade på månadsindikatorer eller laggade variabler men även kombinationer förekommer. Modellerna används antingen i prognoser på ett till flera kvartal framåt i tiden eller för prognoser som avser ett helt år. Prognoserna på offentlig konsumtion saknar en sådan modell vilket till viss del förklaras av att relativt liten vikt läggs vid enskilda kvartalsutfall i prognosen på offentlig konsumtion eftersom revideringarna ofta är stora. En annan anledning är att användandet av tidsserieanalys historiskt sett varit sparsam på området för offentlig konsumtion. För offentlig konsumtion är alltså möjligheten att använda en tidsseriemodell i prognosen än så länge outforskad. Därför undersöker den här uppsatsen om det skulle vara möjligt att använda enkla tidsseriemodeller i prognoserna på offentlig konsumtion. Syftet är alltså inte att ersätta befintlig prognosmetod med en modell utan snarare att se om en tidsseriemodell kan tillföra något i prognosarbetet.

En sådan modell löser emellertid inte de problem som just har beskrivits men innebär att den som gör prognosen har mer information att förlita sig på i prognosarbetet. I nästa kapitel redogörs för vilken data som ligger till grund för skattning och utvärdering av modellerna.

4. Data

I uppsatsen används SCB:s statistik över offentlig konsumtion enligt nationalräkenskaperna för att skatta modellen. Den data som används är kvartalsvisa uppgifter om offentlig konsumtion i fasta priser. För att efterlikna de förutsättningar som råder då KI gör sina prognoser används realtidsdata för att skatta modellen. Definitionen av realtidsdata är den data som finns tillgänglig vid ett givet tillfälle. SCB reviderar tidsserierna i nationalräkenskaperna allteftersom nya uppgifter kommer in men också till följd av att definitioner ändras. Detta gör exempelvis att tidsserien för perioden 1993–2000 ser annorlunda ut 2014 än vad den gjorde 2001. Vid skattning av modeller med realtidsdata fångas bara den information som fanns tillgänglig vid ett givet tillfälle det vill säga ingen information som tillkommit efteråt kommer med. Modellen skattas på kvartalsvis realtidsdata från SCB från 1994. Modellens kvartalsprognoser summeras till årsprognoser som utvärderas för perioden 2004–2013.

Eftersom KI gör prognoser på första utfall beräknas modellens prognosfel som skillnaden mellan prognos och det första utfallet som SCB publicerat för variabeln offentlig konsumtion, i utvärderingen. Första utfall avser den statistik som SCB redovisade den första gången uppgifterna för en viss period publicerades. Även KI:s prognosfel beräknas som skillnaden mellan prognos och första utfall för offentlig konsumtion. KI gör prognoser fyra gånger per år, i mars, juni, augusti respektive december. Data över KI:s prognoser på den årliga utvecklingen av offentlig konsumtion i fasta priser används för att utvärdera modellens prognosförmåga mot KI:s prognosförmåga. Eftersom inte bara realtidsdata, utan också realtidsprognoser, arkiveras systematiskt på KI kan modellprognosen jämföras med motsvarande bedömningsprognoser som KI gjort vid respektive tillfälle.

4.1. INGA MÄRKBARA EFFEKTER AV FINANSKRISEN FÖR OFFENTLIG KONSUMTION

Modellerna i den här uppsatsen skattas på data från 1994 och utvärderas för perioden 2004–2013. Vid prognosutvärdering och modellutveckling är det vanligt att man delar upp analysen i perioden före att ekonomin utsatts för en chock som varit omöjlig att förutse och en period efter. Hösten 2008 inleddes finanskrisen till följd av en sådan oförutsedd händelse och Sveriges BNP utvecklades därför mycket svagt 2008 och 2009, detta berodde i huvudsak på att exporten och investeringarna minskade dessa år men också på att de svenska hushållens konsumtion inte ökade så mycket som tidigare år. Konsekvenserna för utvecklingen av offentlig konsumtion till följd av den finanskrisen var emellertid obefintliga. Den offentliga konsumtionen ökade med 1,3 respektive 2,3 procent dessa år

vilket är mer än det historiska genomsnittet sedan 1993 som är 0,8 procent. Eftersom effekterna på den offentliga konsumtionens utveckling av finanskrisen var obefintliga har ingen korrigering av utvärderingsperioden gjorts till följd av detta i uppsatsen.

5. Metod

I det här kapitlet beskrivs vilka metoder som används för att skatta och utvärdera modellerna. I den första delen beskrivs AR-modellerna, krav på stationaritet och vilken metod för out-of-sampleprognoser som används. I den andra delen beskrivs vilka metoder som används i prognosutvärderingen det vill säga hur prognosfelen mäts, vilka tester som används och hur jag analyserar prognosernas snedvridning.

5.1. TIDSSERIEANALYS

I den traditionella ekonometrin betonades till en början vikten av att använda ekonomisk teori i kombination med studier av ekonomiska samband för att förklara förhållanden mellan beroende och förklarande variabler. Historiskt sett infördes användandet av laggade variabler successivt men till en början gjordes det inte på ett systematiskt sätt. Syftet var då inte heller att analysera strukturen och dynamiken i datan. I tidsserieekonometrin är däremot utgångspunkten att utnyttja all befintlig information som kan fås av en variabls tidsserie. Ett av de största användningsområdena för tidsserieanalys är att göra prognoser. När tidsserieanalys används för att skatta prognosmodeller finns varken intresse för att skapa strukturella modeller, förstå ekonomiska samband eller att testa hypoteser. Det enda syftet är att ta fram en modell som gör träffsäkra prognoser (Asteriou och Hall, 2011).

Det finns flera metoder inom tidsserieekonometrin som kan användas för att skatta en enkel prognosmodell med laggade variabler. En vanlig benchmark är AR(1)-modellen på grund av dess enkelhet och för att den ofta ger relativt bra resultat (se Mitchel, 2009 och Pesaran et al., 2009). Många tidsserier uppvisar autokorrelation, det vill säga linjära samband mellan tidigare observationer. Detta tyder på att tidigare observationer kan förutse kommande observationer. En autoregressiv process modellerar medelvärden för y_t som en funktion av tidigare observationer, $y_{t-1}, y_{t-2}, \dots, y_{t-p}$. Den statistiskt mest enkla och rena tidsseriemodellen är den autoregressiva modellen av ordningen 1, AR(1), som innehåller en lagg. Antagandet bakom en AR(1)-modell är att tidsseriens beteende för Y_t till största del bestäms av värdet på Y i föregående period, $t - 1$. Vad som händer i period t beror på vad som hände i $t - 1$. Alternativt, vad som händer i $t + 1$ beror på vad som hände i t .

En AR-modell kan ha en eller flera laggar. För att veta hur många laggar som är optimalt kan man skatta modeller med olika lagglängd och undersöka vilken modell som ger lägst prognosfel vid utvärdering. Det finns även olika tester som visar vilken lagglängd som är optimal i en AR-modell. Exempelvis kan man använda Akaikes informationskriterie

(AIC). När man gör ett sådant test på den senaste tidsserien för offentlig konsumtion visar det att en AR-modell som består av tre laggade variabler är mest optimal¹. Här skattas därför modeller med 1–4 laggar.

I resterande del av det här kapitlet redogörs för vilka ekvationer som ligger till grund för AR-modeller med olika lagglängd, det vill säga AR(1)-modellen som innehåller en lagg och AR(p)-modellen som innehåller två eller fler laggar.

5.1.1. AR(1)-modellen

Ekvationen för en AR(1)-modell ges av:

$$Y_t = \phi Y_{t-1} + u_t \quad (2)$$

För enkelhetens skull inkluderas ingen konstant, $|\phi| < 1$ och u_t är en Gaussisk felterm, även kallat vitt brus. Feltermen är den enklaste formen av stokastiska processer. Feltermen har inte någon trend eller säsong. Den består av en följd stokastiska variabler u_t som alla sinsemellan är oberoende och därmed saknar korrelation. De är lika fördelade med ett väntevärde som är noll och har en konstant varians σ^2 . Om feltermen dessutom är normalfördelad kallas det Gaussiskt vitt brus.

5.1.2. AR(p)-modellen

En generalisering av AR(1)-modellen är AR(p)-modellen. Siffran inom parentes, p , anger den autoregressiva processens ordning och antalet laggade beroende variabler som modellen har. Exempelvis är en AR(2)-modell en autoregressiv modell av ordningen 2 och har formeln:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t \quad (3)$$

På liknande sätt beror en AR(p)-modell på p tidigare observationer och kallas för en AR-modell av graden p . Ekvationen ges av:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t \quad (4)$$

¹ Testresultatet redovisas i Appendix 1.

Eller om ekvationen skrivs med summeringssymbolen:

$$Y_t = \sum_{i=1}^p \phi_i Y_{t-i} + u_t \quad (5)$$

Och slutligen, om ekvationen skrivs med laggoperatoren:

$$Y_t(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) = u_t \quad (6)$$

$$\phi(L)Y_t = u_t \quad (7)$$

Där $\phi(L)Y_t$ är en polynom funktion av Y_t .

5.1.3. Stationäritet

Ett begrepp som ligger till grund för all tidsserieanalys är stationäritet. En tidsserie måste vara stationär för att det ska vara möjligt att skatta en AR-modell. I en stationär tidsserie är eventuella chocker temporära och på sikt återgår tidsserien till sitt långsiktiga medelvärde. En icke-stationär tidsserie har därför inget långsiktigt medelvärde som serien återgår till, variansen är tidsberoende och oändlig. Stationära tidsserier utsätts däremot ibland för chocker men effekterna av chocken avtar och tidsserien återgår då till sitt långsiktiga medelvärde. En tidsserie är därför stationär när medelvärdet, variansen och kovariansen är konstanta över tid. Tidsserien för Y_t är därför stationär om:

- (a) $E(Y_t)$ är konstant för alla t ,
- (b) $Var(Y_t)$ är konstant för alla t och
- (c) $Cov(Y_t, Y_{t+k})$ är konstant för alla t och alla $k \neq 0$.

Värdena på dessa mått är därför desamma oavsett om observationerna i en tidsserie är från 1993–2003 eller från 2004–2013. Om en tidsserie inte är stationär skulle exempelvis resultaten från en klassisk regression vara ogiltig. Granger och Newbold (1974) visar att regressioner som innehåller icke-stationära tidsserier blir missvisande även om analysen uppvisar signifikanta relationer mellan icke-stationära tidsserier, detta kallas *spurious regressions*.

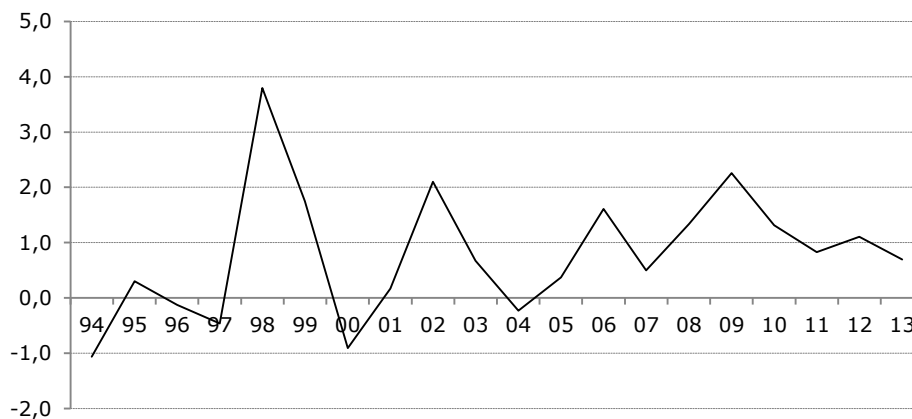
En tidsserie i en AR(1)-modell är stationär om $-1 < \phi < 1$. Om $\phi = 1$ eller -1 råder ett instabilt läge och tidsserien kan förändras men behöver inte göra det. Om $\phi = 1$ och $\alpha = 0$ är tidsserien icke-stationär. Ekvation 2 introducerar restriktionen $|\phi| < 1$ i syfte att garantera stationäritet. Om $|\phi| > 1$ skulle Y_t tendera att bli större i varje period vilket skulle ge en explosiv tidsserie.

Det är enkelt att avgöra om en tidsserie är stationär, icke-stationär eller explosiv genom att studera dess graf. Stationäritet kan också avgöras genom ett enkelt test som utformades av Dickey och Fuller (1979, 1981), testet kallas för Augmented Dickey-Fullertest, ADF-test. Asteriou och Hall (2011) visar ytterligare ett sätt att se om en tidsserie är stationär. Om summan av koefficienterna i en $AR(p)$ -modell är mindre än ett är tidsserien stationär. Eftersom antalet modeller som skattas i den här uppsatsen är relativt många är det enklare att studera tidsseriens graf och att göra ett ADF-test än att summera koefficienterna i alla modeller för att avgöra om vår tidsserie är stationär eller inte.

I diagram 4 i kapitel 3 visades en graf över de revideringar som har gjorts mellan 2004 och 2013, där ser vi att revideringarna är stora men att alla tidsserier ändå visar samma mönster. För enkelhetens skull har jag därför valt att endast studera grafen för det senaste publicerade utfallet för offentlig konsumtion (se diagram 5). Det framgår att tidsserien är stationär eftersom seriens medelvärde tycks vara konstant över tid. Det så kallade ADF-testet visar också att tidsserien är stationär. Sannolikheten att serien skulle vara icke-stationär är mindre än 0,01 procent. Testet visar att t -värdet är $-6,2$ vilket är mindre än det kritiska värdet på 1-procentsnivån som är $-3,8$ (Gujarati, 1995).

Diagram 5 Offentlig konsumtion, 1994–2012

Procentuell förändring



Källa: SCB.

5.1.4. Metod för out-of-sample prognoser

En metod som ofta används för makroekonomiska prognoser är så kallade out-of-sample prognoser. Metoden går ut på att modellen skattas på en uppsättning observationer och utvärderas mot en annan uppsättning observationer som inte ingår i urvalet som skattningen baseras på. I den här typen av prognoser används typiskt sett en av följande två metoder (Zivot, 2013 och Franses, 1998).

1. *Rekursiva prognoser med expanderande informationsmängd.* Metoden innebär att man har en ursprunglig tidsserie från $t = 1, \dots, T$ som den första modellskattningen baseras på. Prognoser görs för h steg framåt i tiden, som perioder saknas för, med start i period T . Tidsserien utökas med en period i taget och modellen skattas om varje gång en ny period läggs till i tidsserien.

För den här uppsatsen innebär metoden att den första prognosen som modellen gör, vilket är 2004, baseras på en skattning på data för perioden 1994q1–2003q4. När nästa prognos ska göras, fortfarande för 2004 men nu med utfall för ytterligare ett kvartal, skattas modellen på data för perioden 1994q1–2004q1. Den sista prognosen som modellen gör avser 2013 och baseras på en skattning med data för perioden 1993q–2013q3.

2. *Prognoser med rullande informationsmängd.* I den här metoden utgår man från en ursprunglig tidsserie från $t = 1, \dots, T$ som den första modellskattningen baseras på. Prognoser görs fortfarande för h steg framåt i tiden, som perioder saknas för, med start i period T . Men antalet perioder som ska ingå i alla kommande skattningar avgörs av hur många perioder som ingår i den första skattningen. När en period tillkommer i tidsserien skattas modellen på perioderna $t = 2, \dots, T + 1$. Prognoser görs för h steg framåt i tiden, som perioder saknas för, med start i period $T + 1$.

För den här uppsatsen innebär metoden att den första prognosen som modellen gör, vilket är 2004, baseras på en skattning på data för perioden 1994q1–2003q4. När nästa prognos ska göras, fortfarande för 2004 men nu med utfall för ytterligare ett kvartal, så skattas modellen på data för perioden 1994q2–2004q1. Den sista prognosen som modellen gör avser 2013 och baseras på en skattning med data för perioden 2003q4–2013q3.

För att efterlikna de förutsättningar som råder när KI:s prognoser görs används metoden med expanderande informationsmängd. Modellerna skattas på kvartalsdata för den årliga procentuella förändringen av offentlig konsumtion i fasta priser. Modellens kvartalsprognoser för innevarande år räknas sedan om till offentlig konsumtionen i nivå uttryckt i antal miljoner kronor. Därefter summeras eventuellt kvartalsutfall i nivå med modellprognosen i nivå för att få fram en helårsprognos i antal miljoner kronor. Helårsprognosen för innevarande år räknas sedan om till en prognos som uttrycks som den offentliga konsumtionens procentuella förändring innevarande år jämfört med föregående år. Helårs-

prognosen på den offentliga konsumtionens procentuella förändring utvärderas sedan mot första helårsutfall och modellens prognosfel utvärderas mot KI:s prognosfel.

5.2. METODER FÖR PROGNOOUTVÄRDERING

KI gör kortsiktiga prognoser för innevarande och nästkommande år. I den här uppsatsen utvärderas modellernas prognosförmåga för innevarande år eftersom sannolikheten att en tidsseriemodell skulle kunna göra träffsäkra prognoser på längre sikt än fyra kvartal framåt i tiden, bedöms som låg.

Modellernas och KI:s prognoser utvärderas mot det första utfallet för offentlig konsumtion som har publicerats. Det är även intressant att undersöka hur prognoserna förhåller sig gentemot senaste publicerade utfall, men en sådan analys är inte möjlig eftersom en förändring av definitionerna i nationalräkenskaperna infördes så sent som hösten 2014. De nya definitionerna påverkade särskilt tidsserien för offentlig konsumtion, i huvudsak beror det på att försvarets vapensystem numera definieras som offentliga investeringar istället för offentlig konsumtion.

KI gör prognoser fyra gånger per år, i mars, juni, augusti och december. I mars finns inget utfall för innevarande år, i juni finns utfall för första kvartalet, i augusti finns utfall för andra kvartalet och i december finns utfall för tredje kvartalet. Prognosförmågan utvärderas för respektive prognosomgång för att ta reda på om modellens prognosprecision skiljer sig mellan olika prognosomgångar. När prognosförmågan utvärderas för olika prognosomgångar och man exempelvis vill veta hur bra modellen står sig mot KI:s prognoser i en marsprognos så väljs endast modellens och KI:s prognosfel i mars ut. Nedan presenteras vilka metoder som används för att utvärdera modellens prognosförmåga.

5.2.1. Relativ RMSE

Vid prognosutvärdering är man intresserad av att mäta storleken på prognosfelet. Prognosfelet definieras här som skillnaden mellan prognos och utfall för en viss variabel. Prognosfelet kan vara positiva om prognosmakaren har gjort en överskattning och negativa om prognosmakaren har gjort en underskattning. Här är vi i första hand intresserade av att utvärdera träffsäkerheten i modellens prognos. Om modellen överskattar eller underskattar offentlig konsumtion återkommer vi till. En vanlig metod för att utvärdera träffsäkerheten hos en prognosmodell är att beräkna rotmedelkvadratfelet som ofta förkortas RMSE (se Yu och Zivot 2011). RMSE står för Root Mean Squared Error och är roten ur genomsnittet av de kvadrerade prognosfelet. Måttet straffar stora fel genom att alla prognosfel kvadreras. Ekvationen för RMSE ges av:

$$RMSE = \sqrt{\frac{\sum_{i=t}^n (y_i^{modell} - y_i^{utfall})^2}{n}} \quad (8)$$

I den här uppsatsen vill vi ju veta om den skattade modellens prognoser är bättre än KI:s prognoser. För att bedöma prognosförmågan hos modellen måste därför modellens prognosprecision ställas i relation till KI:s prognoser. För den här typen av prognosutvärdering används ofta det som kallas för relativ RMSE (se Hjelm och Jönsson, 2010). Ekvationen för relativ RMSE uttrycks:

$$Relativ\ RMSE = \frac{\sqrt{\frac{\sum_{i=t}^n (y_i^{modell} - y_i^{utfall})^2}{n}}}{\sqrt{\frac{\sum_{i=t}^n (y_i^{KI} - y_i^{utfall})^2}{n}}} \quad (9)$$

Om relativ RMSE < 1 är modellens prognoser bättre än att vid varje prognostillfälle använda KI:s prognoser.

5.2.2. Diebold-Marianos test för lika prognosförmåga

Analysen av RMSE säger om en metod ger lägre prognosfel än en annan metod. I det så kallade Diebold-Marianotestet ställs två prognosmetoder mot varandra i syfte att ta reda på om det är möjligt att med statistisk metod avgöra om den ena prognosmetoden är bättre än den andra (Diebold och Mariano, 1995). Resultatet av testet säger om skillnaden i prognosfel mellan två prognosmetoder är statistiskt säkerställd, det vill säga om skillnaden i prognosfel är tillräckligt stor för att man ska kunna säga att den ena metoden faktiskt är bättre än den andra. I testet görs parvisa jämförelser där man utgår från de två prognoserna, $\hat{y}_{modell,t}$ och $\hat{y}_{KI,t}$ för tidsserien y_t med prognosfelen $u_{modell,t}$ och $u_{KI,t}$. Jämförelsen består i att skillnaden av de kvadrerade prognosfelen, $d_{t+1} = \hat{u}_{modell,t+1}^2 - \hat{u}_{KI,t+1}^2$ beräknas för att sedan räkna ut skillnaden mellan medelkvadratfelet, MSE, för de båda prognoserna. Funktionen för testvärdet följer av:

$$t_{DM} = \frac{MSE_{modell} - MSE_{KI}}{\sqrt{Var(MSE_{modell} - MSE_{KI})}} \quad (10)$$

Där nollhypotesen $H_0: MSE_{modell} = MSE_{KI}$ är att metoderna är lika bra och mothypotesen $H_1: MSE_{modell} < MSE_{KI}$ är att modellen gör bättre prognoser än KI.

Teststatistikan, t_{DM} , följer en t-fördelning med beslutsregel $P(t_{DM} > \alpha)$ där $\alpha = t_{0,05}(H - 1)$ där H är antalet prognoser. Vid ett 95-procentigt konfidensintervall är pro-

gnoserna lika bra om $-1,96 < t_{DM} < 1,96$. Om $t_{DM} < -1,96$ gör modellen bättre prognoser än KI och om $t_{DM} > 1,96$ gör prognosen sämre prognoser än KI.

5.2.3. Bias

För att utvärdera modellens och KI:s prognosförmåga jämförs prognosernas genomsnittliga fel som ofta kallas *bias* (engelska för snedvridning). Prognosernas bias definieras som prognos minus utfall. En positiv bias innebär att prognoserna i genomsnitt överskattar variabelns utvecklingstakt och en negativ bias innebär att prognoserna i genomsnitt underskattar variabelns utvecklingstakt. Om en variabels bias är 0,1 så överskattas variabeln systematiskt med i genomsnitt 0,1 procentenhet. Syftet med att analysera prognosernas bias är att se om det finns någon systematisk snedvridning. Vi får då veta om det sker en systematisk över- eller underskattning i prognoserna samt hur stor snedvridningen är. Biasen kan ge en missvisande bild då stora negativa fel ett år kan kompenseras av stora positiva fel ett annat år. Uppgifter om prognosers bias bör därför tolkas med viss försiktighet och gärna tillsammans med att man tittar på hur prognosfelen faktiskt ser ut.

5.2.4. Rangordning av utvärderingsmått

De utvärderingsmått som just presenterats visar på olika sätt hur bra en prognosmetod är i förhållande till en annan. En intressant fråga är emellertid vad som händer om de olika utvärderingsmåten pekar åt olika håll och vilket mått som avgör om modellerna har något att tillföra i prognosen på offentlig konsumtion

En modell skulle exempelvis kunna ha ett högt relativt RMSE som betyder att modellen har större prognosfel och sämre prognosförmåga än KI och samtidigt ha en lägre bias än KI. Detta förklaras av att biasen endast är ett genomsnitt av samtliga prognosfel och om modellens prognosfel har olika tecken kan biasen bli låg som en konsekvens av att stora negativa prognosfel tas ut av stora positiva prognosfel.

I utvärderingen har jag valt att i första hand bedöma modellernas prognosförmåga utifrån deras relativa RMSE. Frågeställningen i uppsatsen är om en modell kan göra lika bra eller bättre prognoser än KI och eftersom måttet relativt RMSE uttrycker hur stort prognosfel modellen har i förhållande till KI är just det resultatet det som är av störst vikt i analysen. I andra hand tar jag hänsyn till resultaten av Diebold-Marianotestet det vill säga om det finns någon statistiskt säkerställd skillnad i metodernas prognosförmåga. Biasen är framförallt intressant om det går att se om måttet speglar en systematiskt över eller underskattning i prognoserna på en viss variabel. Därför är det vid all biasanalys viktigt att studera prognosfelen för de enskilda åren för att ta reda på vilken typ av prognosfel som

har gjorts. Att beräkna prognosernas bias är av stor betydelse när modellerna ska användas i prognosarbetet.

6. Resultat

I detta kapitel redovisas och analyseras resultaten av skattningarna och prognosutvärderingen. KI gör fyra prognoser per år, i mars, juni, augusti respektive december och för varje prognosomgång har fyra olika modeller skattats med 1–4 laggar. Modellen har skattats om varje gång ett nytt kvartalsutfall tillkommer vilket innebär att modellen skattats om fyra gånger per år under utvärderingsperioden, som består av tio år. Sammantaget har därmed 160 modeller skattats (4 AR-modeller, 4 kvartal och 10 utvärderingsår). Modellens och KI:s prognoser utvärderas för respektive prognosomgång för att se om prognosförmågan skiljer sig mellan olika prognostillfällen. Prognosförmågan utvärderas mot första utfall vilket är det första helårsutfallet för offentlig konsumtion som redovisas 60 dagar efter respektive år. Perioden som utvärderas är 2004–2013. En detaljerad redovisning av resultaten från prognosutvärderingen för respektive prognosomgång följer i avsnitt 6.1–6.2.

Syftet med uppsatsen är att undersöka värdet av att använda en modell som stöd i prognosen på offentlig konsumtion. Frågeställningen är därför om modellen kan göra lika bra eller bättre prognoser än KI. Resultaten av *relativa RMSE* är av den anledningen avgörande för modellernas prognosförmåga i förhållande till KI. Ett relativt RMSE som är mindre än ett innebär att modellens prognosfel är mindre än KI:s. Utvärderingen av relativa RMSE visar att det med enkel tidsserieanalys går att skatta en modell som gör bättre årsprognoser på offentlig konsumtion än KI för innevarande år. Utvärderingen visar emellertid att modellen bara är bättre än KI i decemberprognosen när utfall finns tillgängligt för tre av årets fyra kvartal. Det är alltså bara i decemberprognoserna som modellerna har ett lägre prognosfel än KI och utvärderingen visar på ett relativt RMSE som är mindre än ett.

Som komplement till utvärderingen av relativa RMSE introducerades i avsnitt 5.2.2 det så kallade *Diebold-Marianotestet* som är en metod för att testa om det finns någon skillnad i prognosfelen (MSE) mellan modellen och KI. Resultaten av testet visar att det inte går att statistiskt säkerställa att modellens prognosfel är mindre än KI:s i de fall modellens prognoser är bättre än KI:s. Testet visar inte heller på att KI:s prognoser skulle vara bättre än modellens i de fall då relativ RMSE är större än ett. Det är alltså inte statistiskt säkerställt att den ena prognosen är bättre än den andra i någon prognosomgång eftersom skillnaderna i prognosfel är för små. Det kan därmed konstateras att prognosfelen är mycket små och att för långtgående slutsatser av utvärderingen av relativa RMSE inte bör dras.

För att kunna tolka och använda modellernas prognoser är det intressant att veta hur modellernas och KI:s *bias* förhåller sig till varandra. Det är därför viktigt att veta vilka enskilda prognosfel som ligger bakom en bias, detta analyseras i avsnitt 6.1–6.2. I korthet visar modellens och KI:s bias att bådars prognoser underskattar den offentliga konsumtionen. Modellens bias är som minst i decemberprognosen då prognoserna underskattar den offentliga konsumtionen med i genomsnitt drygt 0,3 procentenheter. KI:s bias är som minst i marsprognosen då prognoserna underskattar den offentliga konsumtionen med i genomsnitt 0,3 procentenheter. KI:s prognoser har därmed en något mindre snedvridning än modellen när den jämförs med den modell som har lägst bias. Det finns inget mönster i förändringen av modellens och KI:s bias som man på förhand skulle kunnat tro. Snedvridningen minskar alltså inte i takt med att ytterligare ett kvartalsutfall finns tillgängligt.

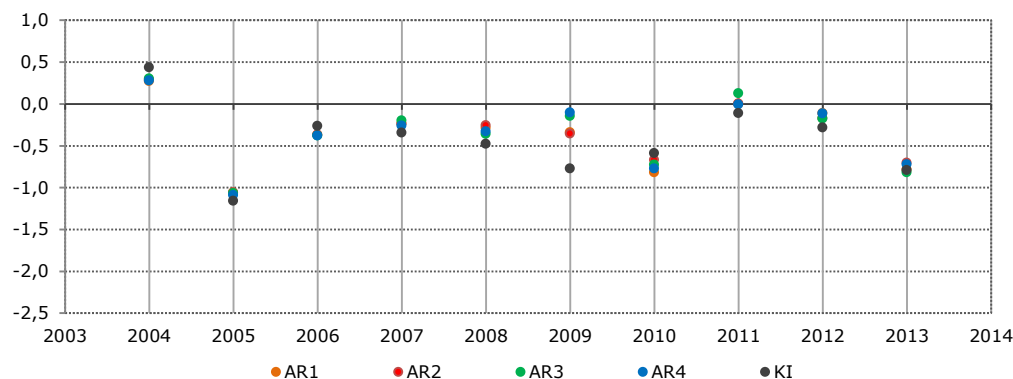
6.1. MODELLEN HAR BÄTTRE PROGNOSEFÖRMÅGA ÄN KI I DECEMBERPROGNOSENA

För att avgöra om modellerna kan göra bättre prognoser än KI jämförs modellernas och KI:s prognosfel genom att det relativa RMSE analyseras. De *relativa RMSE* för decemberprognosen visar att modellen kan göra bättre årsprognoser på offentlig konsumtion än KI för innevarande år när utfall finns för tre av fyra kvartal. Detta gäller oavsett hur många laggar som används i AR-modellen. Modellen som ger lägst relativt RMSE är en AR(2)-modell. Skillnaderna mellan de olika modellerna är mycket små. Det relativa RMSE är mindre än 0,9, vilket innebär att modellerna har drygt 10 procents lägre prognosfel än KI. Den bästa modellen, AR(2)-modellen, har ett relativt RMSE som är 0,84 och den sämsta modellen, AR(1)-modellen, har ett relativt RMSE som är 0,89. Detta betyder att den bästa modellens prognosfel är 16 procent lägre än KI:s och att den sämsta modellens prognosfel är 11 procent lägre än KI:s.

diagram 6 visar modellernas och KI:s prognosfel i samtliga decemberprognoser under perioden 2004–2013. Skillnaden i prognosfel mellan modellerna och KI är liten alla år utom 2009 då samtliga modeller har ett mindre prognosfel än KI. Det är just skillnaden i prognosfel 2009 som bidrar till att det relativa RMSE är så mycket mindre än ett. Men modellerna gör ändå bättre eller lika bra prognoser som KI i åtta av tio decemberprognoser. Då modellen gör sämre prognoser än KI, 2006 och 2010, skiljer det bara 0,1 procentenhet i prognosfel.

Diagram 6 Modellernas och KI:s prognosfel i decemberprognosen, 2004–2013

Procentenheter



Källor: SCB och Konjunkturinstitutet.

Om hänsyn endast tas till modellernas relativa RMSE kan det ändå konstateras att modellerna gör bättre prognoser än KI i decemberprognoserna. Resultaten av *Diebold-Marianotestet* visar däremot att det inte går att statistiskt säkerställa att modellens prognosfel är mindre än KI:s i de fall modellens prognoser är bättre. Att testet inte visar på någon statistisk signifikans gör inte resultatet från relativa RMSE ointressanta, det säger egentligen bara att det inte går att statistiskt säkerställa att den ena prognosen är bättre än den andra eftersom skillnaderna i prognosfelen är för små. Ett relativt RMSE under ett indikerar ändå att modellerna har ett lägre prognosfel än KI och att det finns ett värde i att använda resultaten från modellerna i KI:s prognosarbete. Men man måste komma ihåg att skillnaderna mellan metoderna är mycket små.

Tabell 2 Utvärdering av modellernas och KI:s prognosförmåga i decemberprognosen

Relativ RMSE, Diebold-Marianos test och bias

Första utfall	AR(1)	AR(2)	AR(3)	AR(4)
Relativ RMSE	0,89	0,84	0,88	0,87
Diebold-Mariano	-0,35	-0,65	-0,39	-0,41
Modellens bias	-0,37	-0,35	-0,34	-0,35
KI:s bias	-0,43	-0,43	-0,43	-0,43

Källor: SCB och Konjunkturinstitutet.

Om modellerna ska användas för prognoser på offentlig konsumtion är det viktigast att veta vilken modell som har lägst relativt RMSE. Det är också bra att känna till vilken *bias* både modellerna och KI har för att kunna värdera modellernas resultat på bästa sätt. I Diagram 6 kunde vi se att modellerna underskattat offentlig konsumtion alla år utom 2004 och 2011. Både modellernas och KI:s bias visar också att den offentliga konsumtionen underskattas systematiskt (se tabell 2). Modellernas bias i decemberprognosen innebär att modellen underskattar den offentliga konsumtionen med i genomsnitt drygt 0,3

procentenheter. KI:s bias är något större med sina -0,43. Det ska emellertid poängteras att skillnaderna mellan modellernas och KI:s bias är väldigt små. Dessa skillnader kan helt och hållet förklaras av det prognosfel som KI gjorde 2009. Om 2009 utesluts är biasen -0,35, det vill säga i princip identisk med modellernas bias.

6.2. MODELLERNA HAR STÖRRE PROGNOSEFEL ÄN KI I MARS-, JUNI- OCH AUGUSTIPROGNOSENA

Vi har nu konstaterat att modellernas prognosförmåga är bättre än KI:s i decemberprognoserna men sämre än KI i övriga prognosomgångar. I det här avsnittet studeras prognosfelen i mars-, juni- och augustiprognoserna. Sammanfattningsvis är resultaten från utvärderingen av modellerna väldigt lika mellan de tre prognosomgångarna.

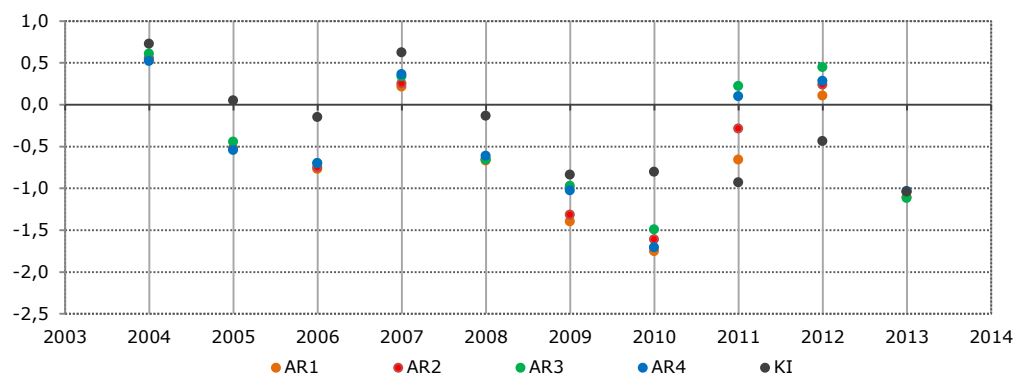
6.2.1. Modellernas prognosförmåga i mars

I marsprognosen finns väldigt lite information tillgänglig och inget utfall för offentlig konsumtion har publicerats för innevarande år. Det huvudsakliga resultatet från utvärderingen visar att modellerna har sämre prognosförmåga än KI i mars. I tabell 3 redovisas modellernas relativa RMSE som är det utvärderingsmått som avgör om modellerna är det utvärderingsmått som avgör modellernas prognosförmåga i förhållande till KI:s. Alla modeller har ett relativt RMSE som är större än ett och gör därmed sämre prognoser än KI. Den bästa modellen har ett relativt RMSE som är 1,19 vilket betyder att prognosfelen är 19 procent större än KI:s. Diebold-Marianotestet är ett komplement i utvärderingen som visar om det är statistiskt säkerställt om modellerna har en bättre prognosförmåga än KI eller inte. Resultaten från testet för marsprognoserna visar som förväntat inte på någon statistiskt signifikant skillnad i prognosfel mellan KI och modellerna.

När vi studerar prognosfelen i diagram 7 framgår det emellertid att modellernas prognoser inte är så dåliga för alla prognosomgångar. Modellerna är faktiskt bättre än KI för fyra år och ungefär lika bra som KI ett år. Modellerna har varit lika bra eller bättre än KI i de senaste tre årens marsprognoser. Modellerna gör både överskattningar och underskattningar av offentlig konsumtion men till följd av stora prognosfel 2009 och 2010 är biasen negativ. Modellerna underskattar offentlig konsumtion systematiskt med 0,4–0,6 procentenheter.

Diagram 7 Prognosfel i marsprognoser, 2004–2013

Procentenheter



Källor: SCB och Konjunkturinstitutet,

Tabell 3 Utvärdering av modellernas och KI:s prognosförmåga i marsprognosen

Relativ RMSE, Diebold-Marianos test och bias

Första utfall	AR(1)	AR(2)	AR(3)	AR(4)
Relativ RMSE	1,37	1,27	1,19	1,23
Diebold-Mariano	0,45	0,35	0,30	0,27
Modellens bias	-0,60	-0,52	-0,38	-0,44
KI:s bias	-0,30	-0,30	-0,30	-0,30

Källor: SCB och Konjunkturinstitutet.

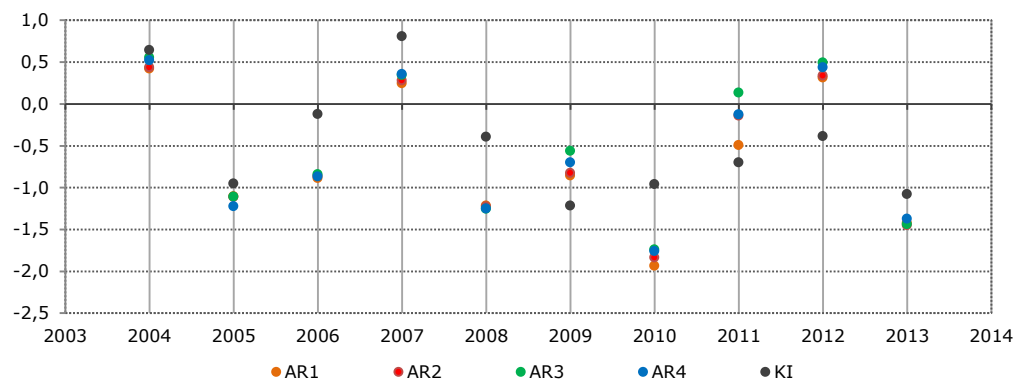
6.2.2. Modellernas prognosförmåga i juni

I juniprognosen har SCB publicerat utfall för första kvartalet, vilket borde tillföra information som är av värde för både modellerna och KI:s prognos. Utvärderingen av relativa RMSE visar däremot att modellernas prognosförmåga fortfarande är sämre än KI:s i juniprognosen. En jämförelse av juniprognosen med marsprognosen visar även att modellernas lägsta relativa RMSE är något högre i juniprognosen än i marsprognosen (se tabell 3). Det kan troligtvis förklaras av att KI då har tillgång till annan information som kan påverka offentlig konsumtion innevarande år som en modell inte kan känna till. Exempel på sådan information är finanspolitiska förslag som läggs fram i vårpropositionen och som påverkar den offentliga konsumtionen det aktuella året.

Trots att modellernas lägsta RMSE är högre i juniprognosen än i marsprognosen så gör modellen bättre prognoser än KI för fem av de tio utvärderingsåren och nästan lika bra prognoser som KI för två år. De senaste fem åren har modellerna gjort bättre prognoser än KI för tre år. Liksom i december- och marsprognoserna underskattar modellerna den offentliga konsumtionen även i juniprognosen och biasen är då också större än i december och mars.

Diagram 8 Prognosfel i juniprognoser, 2004–2013

Procentenheter



Källor: SCB och Konjunkturinstitutet.

Tabell 4 Utvärdering av modellernas och KI:s prognosförmåga i juniprognosen

Relativ RMSE, Diebold-Marianos test och bias

Första utfall	AR(1)	AR(2)	AR(3)	AR(4)
Relativ RMSE	1,29	1,26	1,23	1,25
Diebold-Mariano	0,39	0,37	0,33	0,37
Modellens bias	-0,69	-0,64	-0,54	-0,60
KI:s bias	-0,44	-0,44	-0,44	-0,44

Källor: SCB och Konjunkturinstitutet.

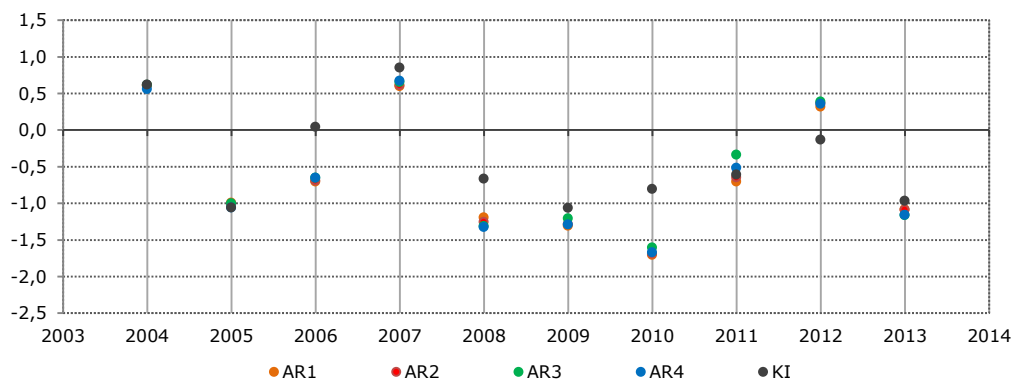
6.2.3. Modellernas prognosförmåga i augusti

Utvärderingen visar att modellernas prognosförmåga försämras ytterligare i augusti jämfört med i juni då modellernas relativa RMSE är högre i samtliga fall (se tabell 5). Resultaten från Diebold-Marianotestet visar också att modellernas prognosförmåga försämras mellan juni och augusti. En förklaring till detta kan vara att det utfall, som SCB publicerar för andra kvartalet i slutet av juli eller början av augusti och som utgör underlag för augustiprognosen, är en så kallad snabbberäkning av nationalräkenskaperna. Snabbberäkningen baseras på mindre omfattande underlag än de ordinarie beräkningarna vilket gör att det utfallet kan komma att revideras mycket när statistiken för offentlig konsumtion för hela året publiceras

Av de tio utvärderingsåren gör modellerna ändå bättre prognoser än KI för fyra år och nästa lika bra prognoser för två år (se diagram 9). Modellerna gör emellertid mycket sämre prognoser än KI för åren 2006, 2008 och 2010. Modellerna underskattar den offentliga konsumtionen i sju fall av tio och biasen är 0,6 procentenheter vilket är ungefär detsamma som i juniprognosen.

Diagram 9 Prognosfel i augustiprognoserna, 2004–2013

Procentenheter



Källor: SCB och Konjunkturinstitutet.

Tabell 5 Utvärdering av modellernas och KI:s prognosförmåga i augustiprognoserna

Relativ RMSE, Diebold-Marianos test och bias

Första utfall	AR(1)	AR(2)	AR(3)	AR(4)
Relativ RMSE	1,32	1,32	1,29	1,33
Diebold-Mariano	0,57	0,57	0,54	0,60
Modellens bias	-0,62	-0,61	-0,56	-0,61
KI:s bias	-0,38	-0,38	-0,38	-0,38

Källor: SCB och Konjunkturinstitutet.

7. Slutsatser

I den här uppsatsen utreds om det är möjligt att använda resultaten från en tidsseriemodell i KI:s prognoser på offentlig konsumtion. Frågan är om det går att utnyttja kvartalsdata för att skatta en enkel modell som kan användas för att göra årsprognoser på offentlig konsumtion. En enkel AR-modell med fyra olika lagglängder har skattats för att undersöka detta. För att besvara frågan om en sådan modell kan tillföra något i prognoserna på offentlig konsumtion har frågeställningen varit om modellen kan göra bättre prognoser än KI. Därför har modellens prognosfel utvärderats och jämförts med KI:s. Metoden för utvärderingen har i första hand varit att analysera modellernas relativa RMSE, dvs. modellernas RMSE i förhållande till KI:s RMSE. Där ett relativt RMSE som är mindre än ett indikerar att modellen har en bättre prognosförmåga än KI. Nästa steg i utvärderingen har varit att genomföra Diebold-Marianos test för lika prognosförmåga. Testet visar om modellens prognosfel är lägre än KI:s och om skillnaden är statistiskt signifikant. Det tredje steget i utvärderingen har varit att studera modellens och KI:s bias för att se om prognoserna har någon snedvridning, hur stor den i så fall är och vad den beror på.

Resultaten av utvärderingen visar att en tidsseriemodell kan göra bättre prognoser än KI men att det bara är i decemberprognosen. Då är modellernas relativa RMSE mindre än ett. I övriga prognosomgångar är modellernas relativa RMSE större än ett. Vilket innebär att modellerna har sämre prognosförmåga än KI då.

Diebold-Marianoestet visar emellertid att det inte är statistiskt säkerställt att modellens prognoser är bättre än KI:s eftersom skillnaden i prognosfelen är för små i decemberprognoserna. I prognosomgångarna i mars, juni och augusti är modellernas prognoser sämre än KI:s prognoser enligt relativa RMSE. När de enskilda prognosfelen vid respektive tidpunkt analyserats är modellerna däremot bättre än KI i ett flertal fall. Modellernas RMSE är däremot högre vilket förklaras av att prognosfelen för vissa år är stora och inte på grund av att alla prognosfel är större än KI:s.

Analysen av modellernas och KI:s bias visar att både modellerna och KI systematiskt underskattar offentlig konsumtion. Modellerna har lägst bias i decemberprognosen och har då också en mindre bias än KI. I övriga prognosomgångar har KI en mindre bias än modellerna och KI:s bias är förvånande nog som minst i marsprognoserna.

7.1. SMÅ SKILLNADER MELLAN MODELLERNA

Resultaten från Diebold-Marianotestet visar att skillnaderna mellan modellernas och KI:s prognosfel är mycket små. Det är även skillnaderna mellan de olika modellerna. Det gäller

både mellan modeller med olika lagglängd och mellan modeller som avser olika prognosomgångar. Att skillnaden är så liten mellan olika modeller kan förklaras av att variansen i den offentliga konsumtionens utveckling är relativt liten. Eftersom det i huvudsak är befolkningens utveckling som styr behovet av välfärdstjänster som vård skola och omsorg och därför påverkar den offentliga konsumtionen så blir svängningarna i offentlig konsumtion ganska små. Det är ovanligt med efterfrågestörningar som påverkar offentlig konsumtion på samma sätt som utvecklingen av andra delar i BNP. Den typ av tillfälliga förändringar som påverkar offentlig konsumtion och som kan ske från ett år till ett annat är exempelvis politiska beslut som innebär att mer eller mindre pengar tillförs offentlig sektor i syfte att öka eller minska mängden utförda välfärdstjänster. Andra saker som kan påverka den offentliga konsumtionen är när antalet personer som söker asyl ökar mycket eller när arbetslösheten är hög och antalet personer i så kallade arbetsmarknadspolitiska program ökar mycket.

7.2. MODELLERNAS RELATIVA PROGNOSEFÖRMÅGA FÖRSÄMRAS MELLAN MARS OCH AUGUSTI

I decemberprognosen är modellernas relativa RMSE som lägst men därefter stiger de allteftersom i mars-, juni- och augustiprognoserna. Resultaten innebär att modellernas prognosfel är större än KI:s i mars och att skillnaden i prognosfel ökar sedan ytterligare i juni och augusti. Modellernas försämrade prognosförmåga i förhållande till KI från mars till juni kan till viss del förklaras av att regeringen lägger fram vårpropositionen i april. I KI:s bedömningsprognos tas eventuella förslag in som påverkar den offentliga konsumtionen innevarande år. En orsak till att modellernas relativa prognosförmåga försämras ytterligare i augustiprognosen kan vara det utfall för andra kvartalet som SCB publicerar i slutet av juli. Utfallet baseras på en så kallad snabberäkning som innehåller en mindre mängd information än en normal beräkning. Snabberäkningen görs för att finansdepartementet ska ha underlag i god tid till budgetpropositionen. Utfallet i snabberäkningen revideras ofta mer än ett ordinarie utfall och kan förklara varför modellernas prognoser är som sämst i augusti i relation till KI.

7.3. OFFENTLIG KONSUMTION UNDERSKATTAS BÅDE AV MODELLEN OCH AV KI

Båda metoderna underskattar utfallet av den offentliga konsumtionen och har en negativ bias i samtliga prognosomgångar. Modellernas underskattning med 0,3–0,7 procentenheter är större än KI:s som underskattar med 0,3–0,5 procentenheter. Orsaken till att KI underskattar offentlig konsumtion i sina prognoser har inte utretts men en möjlig förklaring skulle kunna vara att man är försiktig när prognosen görs eftersom en stark utveckling av offentlig konsumtion både har en stor effekt på BNP-tillväxten och på den offent-

liga sektorns utgifter och därigenom det finansiella sparandet i offentlig sektor. Eftersom prognosen på offentlig konsumtion anses vara relativt svår är det därför vanligt att man på grund av försiktighet underskattar utvecklingen. En potentiell förklaring till varför modellerna underskattar utfallet kan vara att SCB konsekvent reviderar upp tidigare kvartal innevarande år när det fjärde kvartalet publiceras. En sådan analys av utfallet har emellertid inte genomförts.

7.4. MODELLEN KAN BIDRA TILL BÄTTRE PROGNOSE PÅ OFFENTLIG KONSUMTION

En enkel AR-modell kan alltså göra bättre helårsprognoser på variabeln offentlig konsumtion än KI, men bara i december då statistik finns tillgänglig för tre av fyra kvartal. Eftersom resultaten visar på mycket små skillnader i prognosfel mellan modellerna och KI så ska inte för långtgående slutsatser om modellernas prognosförmåga dras. Möjligheten att använda resultat från modellen i mars-, juni- och augustiprognoserna ska av samma anledning inte heller uteslutas. En modell behöver inte nödvändigtvis ha ett lägre prognosfel än en bedömningsprognos för att den ska kunna användas som stöd i prognosarbetet. Om modellen är mer träffsäker och har ett lägre prognosfel än bedömningsprognosen så tyder det snarare på att det finns utrymme att förbättra bedömningsprognosen genom att exempelvis försöka utnyttja information i den här typen av modeller. I takt med att metoderna för bedömningprognoser förbättras och förfinas blir det också svårare för en enkel modell att ha mindre prognosfel. Det är det alltså rimligt att modeller med högre prognosfel än bedömningsprognosen användas som stöd i prognosarbetet.

7.5. ANVÄNDNING AV MODELLEN I KI:S PROGNOSE

Den typ av modeller som skattats och utvärderats i uppsatsen skulle kunna användas som stöd i KI:s prognoser på offentlig konsumtion. Det är emellertid viktigt att den som gör prognosen är medveten om vad som fångas i modellen och vet hur resultaten tolkas på bästa sätt. Den här typen av modeller säger något om hur helårsutfallet för offentlig konsumtion brukar bli när det finns utfall för inget, ett, två eller tre kvartal. Om modellen används i prognosarbetet måste prognosmakaren väga samman modellresultatet med all annan information som normalt används och som är av värde för prognosen. Det är också viktigt att modellen skattas om varje gång ett nytt kvartalsutfall publiceras för att modellen ska baseras på all tillgänglig information. Den som gör prognosen bör också veta hur stora prognosfel modellen brukar göra och hur stor modellens bias är för att kunna använda resultaten på bästa sätt. Det är också intressant för den som gör prognosen att veta hur modellernas prognoser står sig på senare år. Det vill säga att en modell

som används nu har ett lågt RMSE på grund av små prognosfel på senare år och inte små prognosfel för tio år sedan.

Referenser

- Asteriou, D and S G Hall, (2011), *Applied econometrics* (Second edition), New York: Palgrave Macmillan, s. 265–286, 334–355.
- Granger, C W J och P Newbold (1974), Spurious regressions in econometrics, *Journal of econometrics*, vol 2, s. 111–120.
- D’Agostino, A, L. Gambetti och D Giannone (2010), *Macroeconomic forecasting and structural change*, European central bank, Working Paper Nr 1167, s.7–10.
- Dickey D A och W A Fuller (1979), Distribution of the estimators for autoregressive time series with a unit root, *Journal of the american statistical association*, 74, s. 427–431.
- Dickey D A och W A Fuller (1981), Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica* 49, s 1057–1072.
- Diebold, F. X. and R. S. Mariano (1995), Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 2–9.
- Franses, P.H. (1998), *Time series models for business and economic forecasting* Cambridge: Cambridge University press, s. 60–66.
- Gujarati, N. D (1995) *Basic econometrics* (Third edition), New York: McGraw Hill Book Co. S. 809.
- Hjelm, G och Jönsson K (2010), In search of a method for measuring the output gap of the Swedish economy, economic, econometric and practical considerations, Working Paper Nr 115, Konjunkturinstitutet, s. 45.
- Lequiller, F., D. Blades (2014), *Understanding national accounts* (Second edition), Paris: Economica.
- Mitchell, J (2009), Where are we now? The UK recession and nowcasting GDP growth using statistical models, *National Institute Economic Review*, 209, s. 60–69.
- Pesaran, M. H., Schuermann, T. and Smith, L. V. (2009), Forecasting economic and financial variables with global VARs, *International Journal of Forecasting*, 25, s.642–75.

Stock, J.H. och M. W. Watson (1998), A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series, National Bureau of Economic Analysis, Working paper 6607, s. 3–9.

Yu, W S och E Zivot (2011) Forecasting the term structures of treasury and corporate yields using dynamic Nelson-Siegel models, *International journal of forecasting*, 27, s. 579–591.

Österholm P (2014), Survey data and short-term forecasts of Swedish GDP growth, *Applied Economics Letters*, 21:2, s.135–139.

SOU 2002:118, (2002), Utveckling och förbättring av den ekonomiska statistiken, Utredningen om översyn av den ekonomiska statistiken, s. 162–176.

SOU 2002:118 (2002) Beräkningsrutiner för nationalräkenskaperna, Bilaga 3 till slutbetänkande från utredningen om översyn av den ekonomiska statistiken, s. 25–68.

Zivot, E. (2013), Econ 582, Forecast evaluation, University of Washington, föreläsningbilder, 17 april 2013.

Eurostat (2013), European system of accounts – ESA 2010, Publications office of the European Union, s. 69–73.

Appendix

Tabell 6 Enhetsrottest för offentlig konsumtion

1993–2003, senaste utfall

Null Hypothesis: @PC(NKOOF) has a unit root

Exogenous: Constant

Lag Length: 2 (Automatic - based on SIC, maxlag=9)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-70.10290	0.0001
Test critical values:		
1% level	-3.588509	
5% level	-2.929734	
10% level	-2.603064	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(@PC(NKOOF))

Method: Least Squares

Date: 22/01/15 Time: 15:28

Sample: 1993Q1 2003Q4

Included observations: 44

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@PC(NKOOF(-1))	-3.945757	0.056285	-70.10290	0.0000
D(@PC(NKOOF(-1)))	1.955549	0.041541	47.07556	0.0000
D(@PC(NKOOF(-2)))	0.966731	0.021445	45.07894	0.0000
C	5.108844	0.285892	17.86981	0.0000
R-squared	0.995788	Mean dependent var		-0.164120
Adjusted R-squared	0.995472	S.D. dependent var		27.24774
S.E. of regression	1.833429	Akaike info criterion		4.136761
Sum squared resid	134.4585	Schwarz criterion		4.298960
Log likelihood	-87.00874	Hannan-Quinn criter.		4.196912
F-statistic	3152.441	Durbin-Watson stat		1.618834
Prob(F-statistic)	0.000000			

Källor: SCB och egna beräkningar.