

2004:17

**Den svenska konsumentprisindexserien (KPI),
1955–2004**

**En empirisk studie av säsongsmönstret
En tillämpning av TRAMO/SEATS**

I serien Bakgrundsfakta presenteras bakgrundsmaterial till den statistik som avdelningen för ekonomisk statistik vid SCB producerar. Det kan röra sig om produktbeskrivningar, metod-redovisningar samt olika sammanställningar av statistik som kan ge en överblick och underlätta användandet av statistiken.

Utgivna publikationer från 2001 i serien Bakgrundsfakta till Ekonomisk statistik

- 2001:1 Offentlig och privat verksamhet – statistik om anordnare av välfärdstjänster 1995, 1997 och 1999
- 2002:1 Forskar kvinnor mer än män? Resultat från en arbetstidsundersökning riktad till forskande och undervisande personal vid universitet och högskolor år 2000
- 2002:2 Forskning och utveckling (FoU) i företag med färre än 50 anställda år 2000
- 2002:3 Företagsenheten i den ekonomiska statistiken
- 2002:4 Statistik om privatiseringen av välfärdstjänster 1995–2001. En sammanställning från SCB:s statistikällor
- 2003:1 Effekter av minskad detaljeringsgrad i varunomenklaturen i Intrastat – från KN8 till KN6
- 2003:2 Consequences of reduced grade in detail in the nomenclature in Intrastat – from CN8 to CN6
- 2003:3 SAMU. The system for co-ordination of frame populations and samples from the Business Register at Statistics Sweden
- 2003:4 Projekt med anknytning till projektet “Statistik om den nya ekonomin”. En kartläggning av utvecklingsprojekt och uppdrag
- 2003:5 Development of Alternative Methods to Produce Early Estimates of the Swedish Foreign Trade Statistics
- 2003:6 Övergång från SNI 92 till SNI 2002: Underlag för att bedöma effekter av tids-seriebrott
- 2003:7 Sveriges industriproduktionsindex 1913–2002 – Tidsserieanalys
The Swedish Industrial Production Index 1913–2002 – Time Series Analysis
- 2003:8 Cross-country comparison of prices for durable consumer goods: Pilot study – washing machines
- 2003:9 Monthly leading indicators using the leading information in the monthly Business Tendency Survey
- 2003:10 Privat drift av offentligt finansierade välfärdstjänster. En sammanställning av statistik
- 2003:11 Säsongrensning av Nationalräkenskaperna – Översikt
- 2003:12 En tillämpning av TRAMO/SEATS: Den svenska utrikeshandeln 1914–2003
- 2003:13 A note on improving imputations using time series forecasts
- 2003:14 Definitions of goods and services in external trade statistics

Fortsättning på omslagets tredje sida!

Ovannämnda rapporter, liksom övriga SCB-publikationer, kan beställas från:
Statistiska centralbyrån, SCB, Publikationstjänsten, 701 89 ÖREBRO,
telefon 019-17 68 00 eller fax 019-17 64 44.

Du kan också köpa SCB:s publikationer i **Statistikbutiken**:
Karlavägen 100, Stockholm

2004:17

**Den svenska konsumentprisindexserien (KPI),
1955–2004**

**En empirisk studie av säsongsmönstret
En tillämpning av TRAMO/SEATS**

The Swedish Consumerpriceindex (CPI), 1955-2004

An empirical study of seasonality

An application of TRAMO/SEATS

Producent Statistiska centralbyrån
 Avdelningen för ekonomisk statistik

Förfrågningar Lars-Erik Öller, Statistiska centralbyrån och Stockholms universitet,
 Statistiska institutionen,
 e-post: lars-erik.oller@scb.se

Påbyggnadskurs i statistik
Kandidatuppsats, 10 poäng

Författare: Linnea Källqvist
 linnea@kallqvist.com
 Peter Nilsson
 peterboture@yahoo.com

Handledare: Lars-Erik Öller

Innehållsförteckning

1	Inledning	7
1.1	Bakgrund	7
1.2	Syfte	8
1.3	Avgränsning	8
1.4	Disposition	8
2	Metod	9
2.1	TRAMO/SEATS	9
2.2	Säsongrensning	10
2.3	Transformationer	11
2.4	Tillvägagångssätt	11
3	Data	12
4	Resultat	13
4.1	Integrationsordning	13
4.2	Modellbestämning och diagnostik	13
4.2.1	Modell 1	13
4.2.2	Modell 2	14
4.2.3	Modell 3	16
4.2.4	Modell 4	17
4.3	Slutlig dekomponering	19
5	Sammanfattning och avslutande kommentarer	23
5.2	Framtida forskning	24
6	Litteraturförteckning	25
	Appendix A	27
	Appendix B	30
	Appendix C	35
	Appendix D	41
	Appendix E	44

Diagramförteckning

Diagram 3.1	Originalserie av konsumentprisindex, månadsdata, jan 1955 – feb 2004	12
Diagram 4.1	Modell 1, Residualer	14
Diagram 4.2	Modell 2, Total outliereffekt	15
Diagram 4.3	Modell 2, Residualer	16
Diagram 4.4	Modell 3, Residualer	17
Diagram 4.5	Modell 5, Residualer	18
Diagram 4.6	Box Cox-transformerad originalserie	19
Diagram 4.7	Trendcykeln	19
Diagram 4.8	Säsongkomponent	20
Diagram 4.9	Irreguljär komponent	20
Diagram 4.10	Säsongkomponent per månad	21
Diagram C.1	1:a differensen av originalserien	35
Diagram C.2	2:a differensen av originalserien	35
Diagram C.3	1:a differensen och säsongsdifferensen av originalserien	36
Diagram C.4	2:a differensen och säsongsdifferensen av originalserien	36
Diagram C.5	1:a differensen av originalserien, logaritmerad	37

Diagram C.6	2:a differensen av originalserien, logaritmerad	37
Diagram C.7	1:a differensen och säsongsdifferensen, logaritmerad	38
Diagram C.8	2:a differensen och säsongsdifferensen, logaritmerad	38
Diagram C.9	1:a differensen av originalserien, Box Cox-transformerad, $\lambda=0,28$	39
Diagram C.10	2:a differensen av originalserien, Box Cox-transformerad, $\lambda=0,28$	39
Diagram C.11	1:a differensen och säsongsdifferensen, Box Cox-transfor- merad, $\lambda=0,28$	40
Diagram C.12	2: a differensen och säsongsdifferensen, Box Cox-tranfor- merad, $\lambda=0,2$	40
Diagram D.1	Anderson Darlings normalitetstest, Modell 1	41
Diagram D.2	Anderson Darlings normalitetstest, Modell 2	41
Diagram D.3	Anderson Darlings normalitetstest, Modell 3	42
Diagram D.4	Anderson Darlings normalitetstest, Modell 4	42
Diagram D.5	Anderson Darlings normalitetstest, Modell 5	43

Tabellförteckning

Tabell 4.1	Test av originalseriens integrationsordning	13
Tabell 4.2	Diagnostik för modellerna 1 till 5, estimerade i TRAMO	22
Tabell E.1	Modellspecifikationer och diagnostik, originalserie, modell 1 – 12	44
Tabell E.2	Modellspecifikationer och diagnostik, originalserie, modell 13 – 30	45
Tabell E.3	Modellspecifikationer och diagnostik, logaritmerad serie, modell 1 – 18	46
Tabell E.4	Modellspecifikationer och diagnostik, logaritmerad serie, modell 19 – 36	47
Tabell E.5	Modellspecifikationer och diagnostik, logaritmerad serie, modell 37 – 42	48
Tabell E.6	Modellspecifikationer och diagnostik, Box Cox-transf. serie, modell 1 – 12	48
Tabell E.7	Modellspecifikationer och diagnostik, Box Cox-transf. serie, modell 13 – 18	49

Konsumentprisindex (KPI) avser att visa utvecklingen för de genomsnittliga konsumentpriserna med avseende på den privata inhemska konsumtionen. KPI-serien är det mest använda måttet för prisutveckling och används bland annat som inflationsmått. I denna uppsats undersöker vi huruvida det finns något signifikant säsongsmönster i den svenska månadsstatistiken av konsumentprisindex samt av vilken ordning serien är integrerad. Vi tillämpar programmen TRAMO och SEATS för att anpassa en univariat ARIMA-modell till ett dataunderlag, bestående av 590 observationer, under tidsperioden 1955–2004. Vi kommer fram till att KPI-serien låter sig modelleras inom ramen för en ARIMA-modell om serien först transformeras med hjälp av Box-Cox. Modellspecifikationen är dock inte helt adekvat, vilket leder oss till den alternativa slutsatsen att KPI-serien kanske borde modelleras med utgångspunkt i ARCH- och GARCH-familjen. Resultaten från de estimeringar vi utfört indikerar att det existerar ett signifikant säsongsmönster i KPI-serien och att serien karakteriseras av en $I(2)$.

Nyckelord: TRAMO/SEATS, KPI, ARIMA och Box Cox-transformation.

1 Inledning

1.1 Bakgrund

Konsumentprisindex (KPI) avser att mäta den genomsnittliga utvecklingen av priserna för hela den privata inhemska konsumtionen. Det är ett kompensationsindex, som anger hur stor inkomstökning en konsument behöver för att kunna behålla sin konsumtionsstandard när priserna höjts. Ett KPI på exempelvis 110 anger således att inkomsten måste öka med 10 procentenheter för att kompensera för prishöjningen sedan den tidpunkt då KPI låg på 100. Ansvarig för att beräkna och publicera konsumentprisindex är Statistiska centralbyrån (Nationalencyklopedin, 2000).

Man mäter de priser som konsumenterna faktiskt betalar och inkluderar således moms, andra indirekta skatter och subventioner. SCB samlar varje månad in prisnoteringar (i regel den 15:e i varje månad) för ett urval av varor och tjänster. Den relativa betydelsen av olika representantvaror anges genom vägningstal, som visar hur stor andel de olika varorna och tjänsterna representerar i den totala privata konsumtionen. Urvalet av representantvaror och deras vägningstal revideras vid varje årsskifte. Som underlag används uppgifter från nationalräkenskaperna om den privata konsumtionens sammansättning under närmast föregående år. KPI är ett kedjeindex med årslänkar, konstruerat för att tillgodose krav på både kort och lång sikt. Därav följer att SCB publicerar ett korttidsindex för respektive månad samt ett långtidsindex i december (www.scb.se).

KPI-serien är det mest använda måttet för prisutveckling och används bland annat som inflationsmått. Det är därför av central betydelse att öka förståelsen och kunskapen om KPI-seriens struktur och karaktär. Utgångspunkten i föreliggande studie är att undersöka hur den svenska konsumentprisserien har förändrats under tidsperioden 1955–2004 och därmed närmare studera om serien, i likhet med till exempel nationalräkenskaperna och industriproduktionsindex, uppvisar trend, cykler och säsongskomponenter och i så fall hur de ser ut. Vi är särskilt intresserade av att undersöka om det finns någon signifikant säsongsvariation i serien eftersom det, oss veterligen, inte gjorts någon sådan undersökning tidigare. Men även med tanke på att det i dagsläget inte förekommer någon säsongrensning av konsumentpriserna. För att undersöka detta appliceras TRAMO/SEATS på den svenska månadsstatistiken av konsumentprisindex.

Statistikprogramvaran TRAMO/SEATS utvecklades av Agustin Maravall och Victor Gomez vid Spaniens centralbank. TRAMO/SEATS består av de två integrerade metoderna TRAMO och SEATS, som bygger på modellbaserad säsongrensning. De två programmen är officiellt rekommenderade av Eurostat och den Europeiska centralbanken (ECB) (Öhlén, 2003). Metoden för säsongrensning grundar sig på univariat ARIMA-modellering (Box och Jenkins, 1976) och används på SCB för att säsongjustera vissa tidsserier, som exempelvis nationalräkenskaperna och industriproduktionsindex.

Den här uppsatsen undersöker även KPI-seriens integrationsordning. Två enhetsrotstester och ett stationaritetstest kommer därför att utföras. Framförallt är vi intresserade av att testa för om prisserien uppvisar en dubbel enhetsrot, det vill säga karakteriseras av en $I(2)$, för den stokastiska trenden. Om så är fallet skulle det innebära att inflationen karakteriseras av en $I(1)$. Om prisserien skulle uppvisa en dubbel enhetsrot implicerar det att chocker som påverkar inflationen är av permanent karaktär, det vill säga att inflationen uppvisar så kallad persistens¹. Eftersom under-

¹ Persistens i priserna implicerar att priserna är trögrörliga, vilket innebär att anpassningen mot jämvikt går långsamt.

sökningsperioden sträcker sig över femtio år har naturligtvis inflationen varierat kraftigt, men en dubbel enhetsrot i KPI-serien relateras inte till magnituden, utan till långvarigheten hos chocker, vilket gör det teoretiskt möjligt att finna en dubbel enhetsrot i prisserien under en period med låg inflation likväl som i en period med hög inflation. För ett utförligare resonemang om egenskaperna hos en $I(2)$ se Juselius (2004).

1.2 Syfte

Uppsatsens syfte är framförallt att försöka besvara följande frågor:

- Av vilken ordning är KPI-serien integrerad?
- Finns det ett signifikant säsongsmönster i KPI-serien?
- Är det rimligt att under den givna tidsperioden approximera KPI-serien med en ARIMA-modell?

1.3 Avgränsning

Den här uppsatsen använder sig av statistikprogrammen TRAMO och SEATS för att estimerar månadsstatistiken av KPI. Vi kommer däremot inte att undersöka om vi får andra resultat med alternativa programvaror. Vi är medvetna om att det ideala förfaringssättet är att estimerar modeller utifrån olika programvaror för att sedan jämföra resultaten.

Valet av tidsperiod kan även det ses som en avgränsning, eftersom observationer av konsumentprisindex finns från 1914. Ett alternativ hade därför varit att utgå från kvartalsdata, vilket finns att tillgå från 1917. Eftersom vi i den här uppsatsen undersöker huruvida det finns någon säsongsvariation i KPI-serien är valet av tidsperioden 1955–2004 att föredra, då månadsstatistik finns för hela den perioden. Detta kan därmed ge oss en precisare beräkning av säsongsmönstret.

Den här uppsatsen utgår även från antagandet att det saknas kalendereffekt i KPI-serien. Vi är medvetna om att detta kan ifrågasättas eftersom årets månader är av olika längd, vilket till exempel innebär att helger faller ut oftare under vissa månader än andra. Vi har dock inte hittat något stöd för att kalendereffekter har en effekt på konsumentpriserna, varken i den litteratur vi gått igenom eller hos SCB.

1.4 Disposition

Avsnitt 1 inleder. Avsnitt 2 tar bland annat upp metodiken bakom de båda programmen TRAMO och SEATS samt det tillvägagångssätt vi använder oss av för att finna den lämpligaste ARIMA-modellen. Avsnitt 3 redovisar vårt datamaterial och i avsnitt 4 presenteras resultaten. Avsnitt 5 sammanfattar våra resultat samt ger avslutande kommentarer.

2 Metod

I detta kapitel presenteras de metoder som används. Det första och andra avsnittet behandlar statistikprogrammet TRAMO/SEATS respektive säsongsrensning. Sedan beskrivs den Box-Cox-transformation, som, bland annat, har nyttjats för att stabilisera KPI-serien i studien. Avslutningsvis redogörs för vårt tillvägagångssätt.

2.1 TRAMO/SEATS

TRAMO, "Time series Regression with ARIMA noise, Missing observations and Outliers", är ett Fortran-program, som vid SCB används på PC under MS-DOS och Windows. Det är ett program för skattning och prognostisering av regressionsmodeller, vars slumpterm specificeras med en ARIMA-modell. Programmet identifierar och korregerar för de olika typerna av extremvärden och kan även skatta speciella effekter såsom kalendereffekter (Öhlén, 2003).

SEATS, "Signal Extraction in ARIMA Time Series", är ett program för estimering av icke observerbara komponenter i tidsserier. Programmet använder en ARIMA-baserad metod. De olika komponenterna skattas och prognostiseras med signalextraheringsteknik applicerad på ARIMA-modeller (Öhlén, 2003).

En utförlig beskrivning av metodologin bakom de båda programmen går att läsa i Gomez och Maravall (1994, 1996, 1998), Maravall (1995) och Maravall och Sánchez (2000). Förfaringssättet kan, kortfattat, summeras enligt följande:

Givet tidsserien $y_t = (y_1, \dots, y_t)$ av observerade värden skattar TRAMO regressionsmodellen:

$$y_t = x_t' \beta + v_t, \quad (2.1.1)$$

där x_t' är en matris av regressionsvariabler: $x_t' = (x_{1t}, \dots, x_{mt})$ och β representerar en vektor av regressionskoefficienter: $\beta' = (\beta_1, \dots, \beta_m)$ och v_t följer den stokastiska, generella ARIMA-processen:

$$\varphi(B) \Phi(B^s) \Delta^d \Delta_s^D y_t = \theta(B) \Theta(B^s) e_t.^2 \quad (2.1.2)$$

SEATS dekomponerar sedan v_t , enligt:

$$v_t = TC_t + S_t + I_t + T_t, \quad (2.1.3)$$

där TC står för trendcykeln, S för säsongen, I för den irreguljära komponenten och T för kalendereffekter samt outliereffekter. Dessa komponenter följer i sin varsin ARIMA-modell, vilket gör det möjligt att få enskilda prognoser för de olika komponenterna i SEATS.

Säsongsjustering är det speciella fallet då:

$$v_t = N_t + S_t, \quad (2.1.4)$$

där N_t representerar den säsongsjusterade serien det vill säga:

$$N_t = TC_t + I_t + T_t. \quad (2.1.5)$$

² Se appendix A för en teoretisk genomgång av ARIMA-modellering.

2.2 Säsongrensning

Sedan 1999 använder SCB TRAMO/SEATS vid säsongrensning av den officiella statistiken. Dessförinnan användes programmet X-11-ARIMA. Bytet av säsongrensningsprogram ägde rum efter ett utvecklingsarbete, som genomfördes på SCB under slutet av 1990-talet, delvis i samarbete med Eurostat. Det visade sig att metoden för säsongrensning med programmen TRAMO och SEATS gav högst kvalitet (Öhlén, 2003).

De två programmen är strukturerade så att de kan användas tillsammans. När de används för att säsongjustera en tidsserie så förjusteras serien i TRAMO för att sedan dekomponeras i SEATS. En säsongrensad serie kan sedan användas till att göra jämförelser mellan alla tidpunkter i en serie. Det finns dock alltid kvar ett visst brus, vars betydelse varierar mellan olika serier, men även för olika metoder av säsongrensning. Detta innebär att tidsmässiga förändringar, beräknade på den säsongrensade serien, ofta kan vara svåra att tolka. I dessa fall kan det vara lämpligt att även redovisa den långsiktiga utvecklingen med beräkningen av trendcykeln. Trendcykeln representerar den långsiktiga utvecklingen bättre på grund av att den eliminerar bruset i högre grad jämfört med den säsongrensade serien. SCB:s officiella statistik, som redovisar månadsförändringar/kvartalsförändringar uppräknat till årstakt, är därför oftast beräknade på basis av trendcykeln (www.scb.se).

Den säsongjusterade serien innehåller, förutom trendcykeln och bruset, även eventuella extremvärden (outliers). Extremvärden är observationer, som avviker från de övriga värdena i serien även om vi tar hänsyn till säsong- och kalendereffekter (för att kunna definiera en avvikelse måste naturligtvis en specifik gräns fastställas). Extremvärden definieras utifrån tre skilda typer: (Planas, 1997)

- Additivt extremvärde – karakteriseras av att den påverkar serien endast vid en tidpunkt, varvid serien "hoppar till" men genast återgår till ungefär samma nivå som tidigare.
- Nivåskifte – karakteriseras av att serien, från en viss nivå hoppar till en annan nivå och sedan stannar på den nya nivån.
- Temporär förändring – karakteriseras av att serien hoppar till en ny nivå, men återgår till den ursprungliga nivån efter några perioder.

Orsakerna till att extremvärden existerar i serien kan variera. I vissa fall kan förändringar i den ekonomiska politiken ge upphov till extremvärden. I andra fall kan externa händelser påverka serien, till exempel att världsmarknadspriset på olja av någon anledning drastiskt förändras. Även förändringar av representantvaror i konsumentpriskorgen kan ge upphov till nivåskiften och i vissa fall kan statistiska fel orsaka extremvärden i serien (www.scb.se).

För att kunna beräkna de olika komponenterna i tidsserien på ett adekvat sätt, måste flera grundläggande statistiska krav tas i beaktande. Först är det viktigt att komponenterna är oberoende av varandra. Säsongrensningen försämras till exempel avsevärt om säsongsfaktorn samvarierar med den cykliska komponenten, vilket kan vara svårt att avgöra. Vidare är det centralt att den irreguljära faktorn, bruset, antas följa en oberoende normalfördelning (Kaiser och Maravall, 2001).

I praktiken är det självfallet så att statistiska tidsseriemodeller endast är en förenkling av verkligheten. Detta innebär att förutsättningen för att den valda modellen är en acceptabel approximation med nödvändighet måste prövas med lämpliga statistiska tester. I den här uppsatsen kommer en rad olika statistiska tester att genomföras, vilka syftar till att undersöka en given modells lämplighet. Dessa tester redovisas i appendix B.

2.3 Transformationer

Vid estimering av univariata ARIMA-modeller, men även vid estimering av många andra modeller, utgår man från att tidsseriens varians är konstant över tiden, det vill säga homoskedastisk. När det gäller ekonomiska tidsserier är de dock ofta växande över tiden, vilket även gäller för prisserier i allmänhet och konsumentpris-indexserien i synnerhet. Generellt sett ökar variansen med nivån. Det absolut vanligaste förfaringssättet för att stabilisera serien är då att logaritmera tidsserien, vilket kan ha flera praktiska fördelar. Dels får man ett direkt estimat på elasticiteten i regressionslikvationer med ekonomiska variabler och dels gäller, för små förändringar, att differensen av en logaritmisk serie är approximativt en procentuell förändring. Den viktigaste egenskapen är dock att vi kan få tidsserien homoskedastisk (Gujarati, 1999).

En mer generell transformation som kan användas för att stabilisera en tidsserie med varierande varians är den så kallade Box Cox-transformationen (Box och Cox, 1964). För tidsserien y_t definieras Box Cox-transformationen $T(y_t)$ som:

$$T(y_t) = y_t^{(\lambda)} = \frac{y_t^\lambda - 1}{\lambda}, \quad (2.3.1)$$

där λ representerar transformationsparametern.

2.4 Tillvägagångssätt

Det första som kommer att undersökas i denna uppsats är KPI-seriens integrationsordning. För att undersöka detta använder vi oss av två olika *enhetsrotstester*, dels det utökade Dickey Fuller (ADF) och dels Phillips Perron. Vi kommer också att använda oss av ett *stationaritetstest*, KPSS-testet. Dessa finns beskrivna i appendix B. Testen utförs i statistikprogrammet STATA.

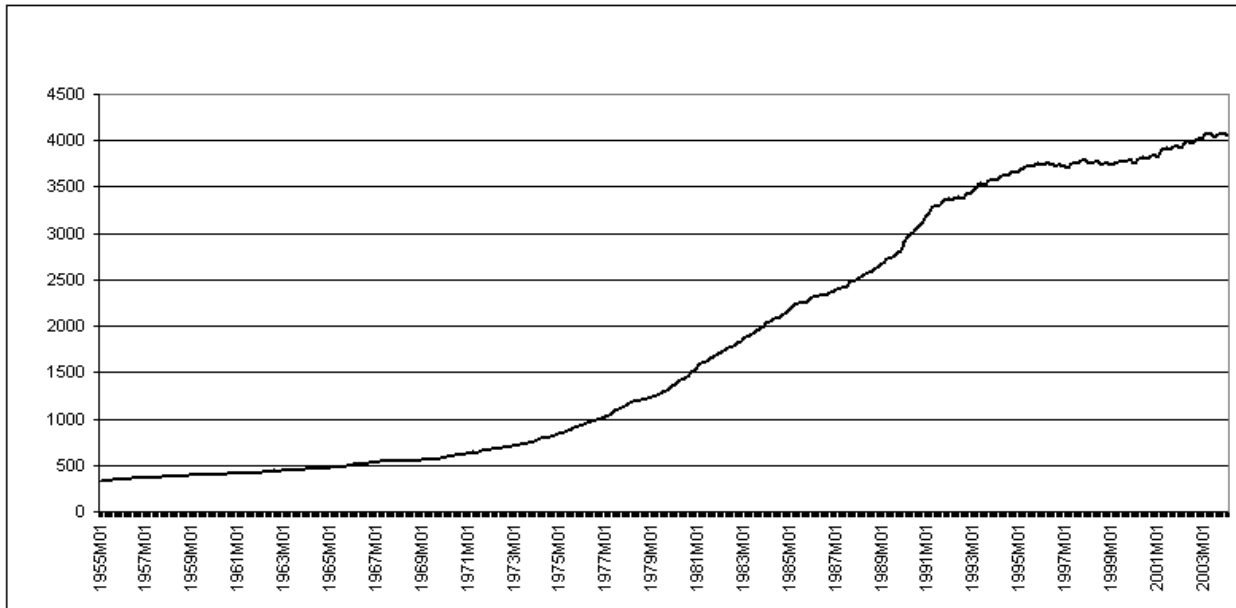
Efter detta går vi vidare med att tillämpa programmen TRAMO och SEATS på KPI-serien. Vi kommer att utgå från originalserien och den fullt automatiska inställningen i TRAMO. Den första modellen analyseras sedan utifrån den valda diagnosen. Förfaringssättet kommer sedan att utgå från en iterativ metod, det vill säga att vi genom en metod för upprepning testar oss fram. Inom ramen för ARIMA-modellering innebär detta att vi kommer att testa olika kombinationer av ingående parametrar i modellen. Om ingen lösning verkar möjlig utifrån att endast ändra de ingångna parametrarna, går vi vidare med att testa alternativa tillvägagångssätt manuellt utifrån den eventuella feldiagnosen. För att testa huruvida fördelningen för de beräknade residualerna är symmetrisk utförs två olika normalitetstest, dels det i TRAMO/SEATS inbyggda Bowman Shenton-testet och dels Anderson Darling-testet.

3 Data

Konsumentprisindex började registreras redan 1914. De tre första åren, 1914–1916, finns beräkningar för indexet årsvis och mellan 1917–1954 finns kvartalsstatistik. Tidsserien som ligger till grund för analysen i den här uppsatsen utgörs av 590 observationer och har som basår 1914. Månadsstatistik för KPI började föras under sommaren 1954. Vi använder oss av observationer från januari 1955 fram till och med februari 2004.

Diagram 3.1

Originalserie av konsumentprisindex, månadsdata, jan 1955 – feb 2004



För att serien ska kunna modelleras med ARIMA krävs att serien är stationär, vilket ofta kan åstadkommas genom en lämplig differentiering och/eller transformering (se appendix A). För att få en överskådlig bild av seriens utseende efter genomförda differentieringar och transformeringar visas diagram över detta samlade i appendix C.

4 Resultat

Kapitlet redovisar studiens resultat med avseende på KPI-seriens integrationsordning, modellbestämning och undersökning av säsongsmönster. Kapitlet innehåller även en dekomponering av vår slutgiltiga modell.

4.1 Integrationsordning

Olika specifikationer av univariata ARIMA-modeller kommer att appliceras på KPI-serien i detta avsnitt av uppsatsen. För att få en uppfattning om hur dessa bör vara uppbyggda, inleder vi analysen med att testa av vilken ordning serien är integrerad. För att undvika störningen av den möjliga säsongsvariationen utgår vi här från årsdata, vilket självklart eliminerar fluktuationer inom året.

Tabell 4.1, nedan, redogör för resultaten av de olika testerna angående seriens integrationsordning. Båda enhetsrottesten indikerar att originalserien är av integrationsordning 2. Vi får även samma resultat med stationaritetestet, KPSS. Vi kan därmed konstatera att KPI-serien, intressant nog, verkar uppvisa en dubbel enhetsrot för den långsiktiga utvecklingen, det vill säga för den stokastiska trendmodellen. Rimligtvis borde detta resultat medföra att det samma ska gälla även för de modeller vi ska estimerar i TRAMO.

Tabell 4.1
Test av originalseriens integrationsordning

Test	Nivå		Första differensen		Andra differensen		Tredje differensen	
	konstant, trend	konstant, utan trend	konstant, trend	konstant, utan trend	konstant, trend	konstant, utan trend	konstant, trend	konstant, utan trend
ADF	—	—	-2.28 [0.45] (1)	0.15 [0.97] (1)	-2.05 [0.58] (1)	-2.04 [0.27] (1)	-5.39 [0.00] (1)	-5.42 [0.00] (1)
PP	—	—	-2.10 [0.55]	1.14 [0.99]	-2.52 [0.32]	-2.38 [0.15]	-8.53 [0.00]	-8.58 [0.00]
KPSS	0.52	2.43	0.33	0.84	0.04	0.06		
Kritiskt 1%	0.12	0.74	0.12	0.74	0.12	0.74		
värde 5%	0.15	0.46	0.15	0.46	0.15	0.46	—	—
10%	0.22	0.35	0.22	0.35	0.22	0.35		

ADF t-test, PP t-test $H_0: \psi = \theta - 1 = 0$ MacKinnons (1991) approximerade p-värde inom hakparantes.
ADF-laggar i parantes, valda utefter BIC.

4.2 Modellbestämning och diagnostik

Detta avsnitt presenterar våra estimeringar i TRAMO och SEATS. Framställningen utgår från fem modeller, vilka är av särskilt intresse. Övriga modeller finns redovisade i appendix E.

4.2.1 Modell 1

Vi går sedan vidare med sökandet efter en adekvat ARIMA-modell, genom att applicera programmet TRAMO på KPI-serien. Den första modellen som estimeras är med automatisk inställning i TRAMO (undantaget att TRAMO ställs in på outlier-korrektion).

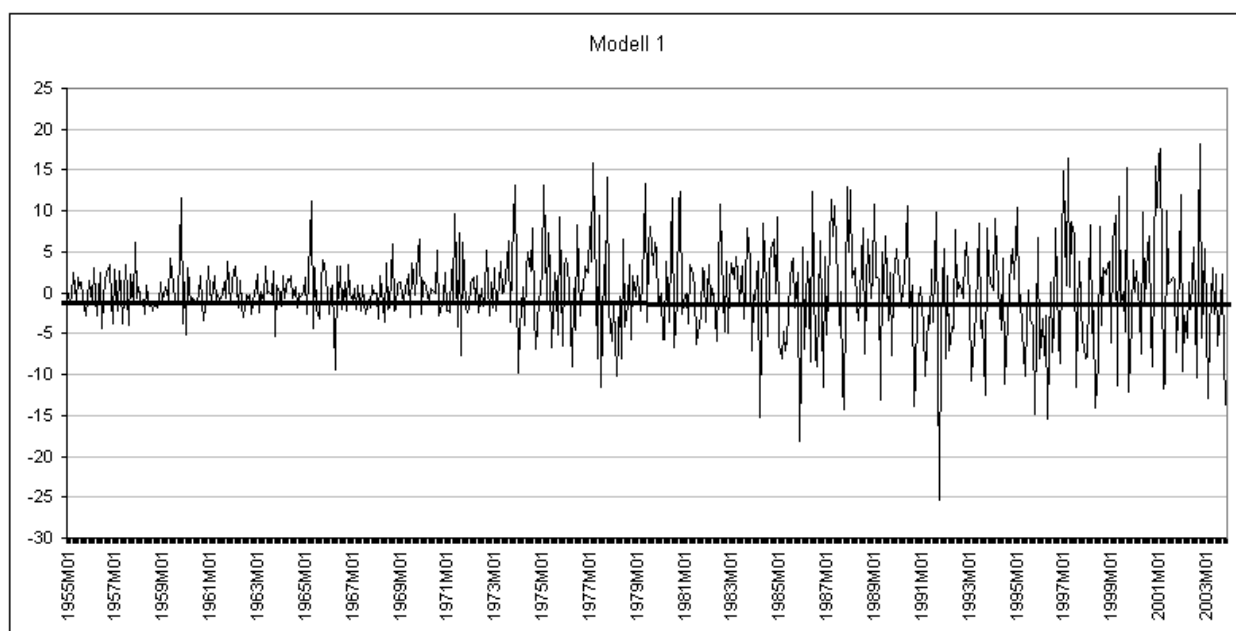
Med denna procedur erhålls ARIMA-modellen (3,1,1)(0,1,1), vilket skrivs

$$(1 - 0,98B + 0,20B^2 - 0,13B^3) \Delta \Delta_{12} x_t = (1 - 0,76B)(1 - 0,64B^{12}) e_t, \quad (4.2.1)$$

$\begin{matrix} (-10,09) & (3,16) & (-2,60) & (-8,51) & (-17,74) & t\text{-värden} \end{matrix}$

för seriens nivåer. All diagnostik för modellerna finns samlad i tabell 4.2 på sid 17. Modellen innehåller 26 outliers, vilket nog får betraktas som ett mycket högt antal. Ser vi till diagnostiken för modellens residualer visar normalitetstestet, N , att residualerna inte är normalfördelade eftersom $N = 30,65$ är större än det kritiska värdet 6 (se även Anderson Darlings-normalitetstest i appendix B). Värdena för skevhet och kurtosis visar tydligt att icke-normaliteten kan associeras med den senare. Fördelningen karakteriseras därmed av leptokurtosis, se appendix B, vilket visserligen inte får så stor effekt på punktprognoser (Maravall och Sánchez, 2000). Autokorrelationstestet, Ljung Box-Q, är däremot bra och likaså värdet på Box Pierce-Q. Ett diagram över residualerna, diagram 4.1 nedan, visar dock tydligt att modellen inte är lämplig för att beskriva KPI-serien, eftersom den lider av heteroskedasticitet. Detta bekräftas även av värdet för kvadrerade Ljung Box-Q. Därmed är serien inte stationär. Run-testet visar att residualerna inte heller är okorrelerade, vilket, även det, kan tyda på ickestationaritet.

Diagram 4.1
Modell 1, Residualer



Efter att ha testat ett trettioital modeller i TRAMO (se appendix E för modellspecifikationer och diagnostik) kunde vi konstatera att problemet med heteroskedasticitet troligen inte går att lösa med en modell som specificeras i nivåer. Detta leder oss till att låta TRAMO testa för en logaritmerad specifikation istället (diagram av olika differentieringar av den logaritmerade serien visas i appendix C).

4.2.2 Modell 2

Efter att TRAMO logaritmerat serien anpassas nu även en annan modell. Istället erhålls ARIMA-modellen $(0,2,1)(0,1,1)$, vilket kan skrivas som:

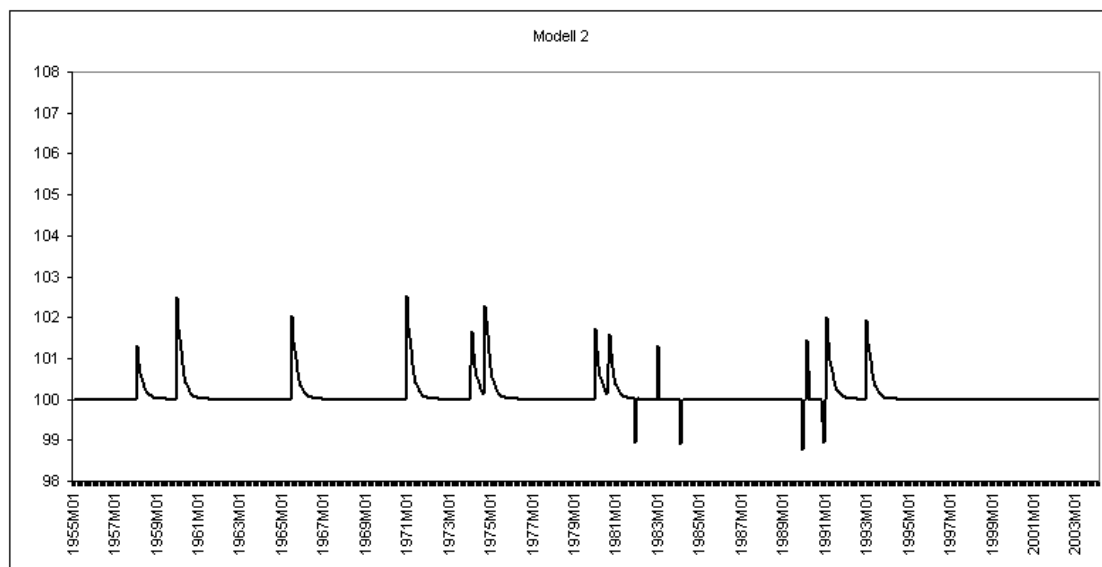
$$\Delta^2 \Delta_{12} x_t = (1 - 0,92B)_{(-56,27)} (1 - 0,84B^{12})_{(-36,71)} e_t, \quad t\text{-värden} \quad (4.2.2)$$

med de estimerade parametrarna inkluderade. Modell 2 implicerar att trendmodellen karakteriseras av en $I(3)$ natur. Men om MA-polynomen faktoriseras ges att en av rötterna är lika med $(1 - 0,99B)$ och därmed utbytbar mot Δ . Därav följer att modellen, i praktiken, förenklas till en $I(2)$, vilket även bekräftas av SEATS. Samti-

diget är den andra roten, $(1 - 0,92B)$, inte heller långt ifrån 1, vilket skulle kunna innebära att modellen är nära en $I(1)$. Standaravvikelsen på parametern från SEATS visar dock att så inte är fallet, vilket innebär att modellen stämmer väl med vad vi testat för under avsnitt 4.1; att trendmodellen karakteriseras av en $I(2)$.

Antalet outliers minskar nu till 16 stycken och den totala outliereffekten på serien åskådliggörs i diagram 4.2, nedan.

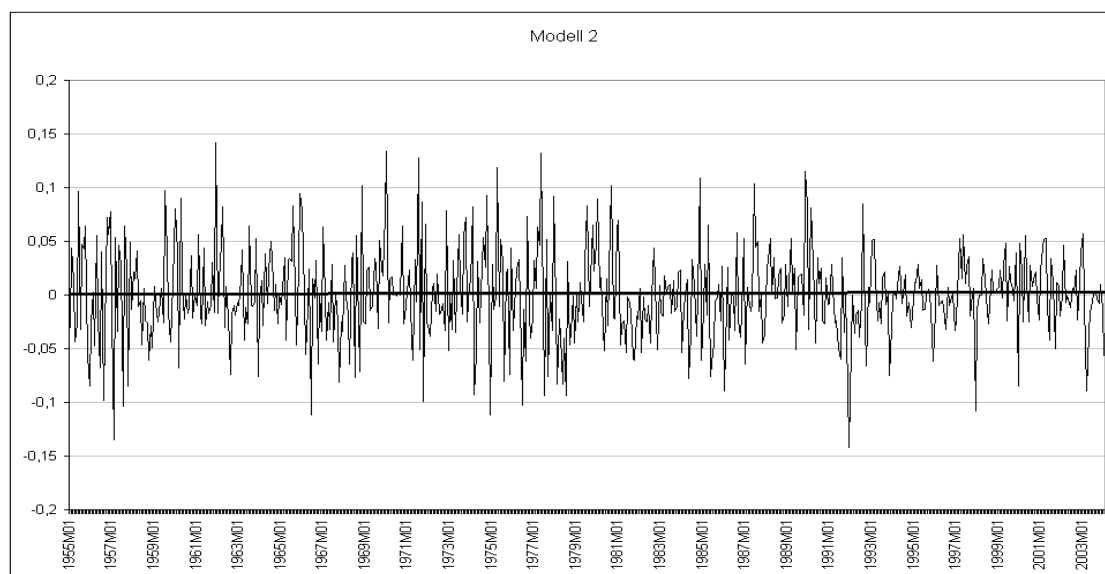
Diagram 4.2
Modell 2, Total outliereffekt



Diagrammet ovan visar att outliereffekten, framförallt, karakteriseras av en rad temporära förändringar, vilket tyder på att prisserien efter en viss tid återgår till sin ursprungliga nivå då någon extrem händelse initialt har påverkat priserna.

Diagnostiken för modell 2 är bra och parametrarna är signifikanta. Värdet på Ljung Box Q^2 indikerar att heteroskedasticiteten också har försvunnit. Likaså är autokorrelationstestet (Ljung Box-Q) fortfarande godkänt och Run-testet är inte längre signifikant. Grafiskt ser residualerna nu bättre ut jämfört med modell 1, se diagram 4.3. Men det finns en viss tendens till att residualerna är mindre i den andra halvan av serien jämfört med den första halvan.

Diagram 4.3
Modell 2, Residualer



Modellen har fortfarande problem med icke-normalitet i residualerna, även fast värdet på normalitetstestet har sjunkit något. Skevhets- och kurtosisvärdena indikerar att detta, fortfarande framförallt beror på en viss leptokurtosis. SEATS dekomponerar modellen och en summering av diagnostiken finns samlad i tabell 4.2 på sid 17. De olika komponenternas slumpmässiga varians tyder på en relativt stokastisk trendcykel, en stabil säsongsfaktor och en liten irreguljär komponent.³ Den historiska estimeringen av säsongsvariationen visar att fem månader är signifikanta. Ser vi till estimeringen för de två senaste åren och för ettårsprognosen är fyra månader signifikanta.

Efter att ha testat drygt fyrtio modeller i TRAMO (se appendix E för modellspecifikationer och diagnostik) kunde vi konstatera att problemet med ickenormaliteten i residualerna troligen inte går att lösa genom att endast specificera andra parametrar. Detta leder oss till att testa en annan modifikation.

4.2.3 Modell 3

I diagram 4.3 kunde vi ana oss till att den logaritmerade transformationen av KPI-serien möjligen var för stor. Residualerna har, från att ha uppvisat ett ökande mönster i diagram 4.1, nu snarare ett svagt avtagande utseende i diagram 4.3. Ett antal olika Box Cox-transformationer testades därför i syfte att finna en mer lämplig transformation av data. Det visade sig att Box Cox- transformationen med transformationsparametern, $\lambda = 0.28$, resulterade i den bästa modellen då TRAMO, liksom tidigare, körs med automatisk inställning (diagram av olika differentieringar av den Box Cox-transformerade serien visas i appendix C).

TRAMO ändrar nu modell till en ARIMA(1,2,1)(0,1,1), vilken med de estimerade parametervärdena kan skrivas som:

$$(1 + 0,07B)\Delta_{12}^2 x_t = (1 - 0,93B)(1 - 0,78B^{12})e_t, \quad (4.2.3)$$

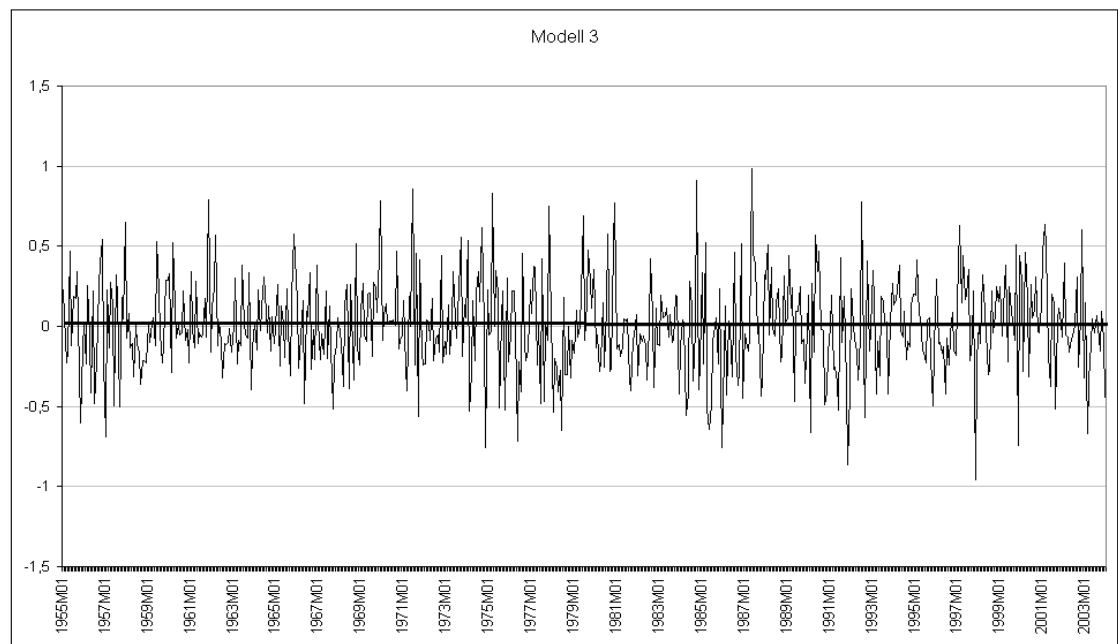
(1,70) (-58,69) (-27,90) t-värden

³ Önskvärt är att trendcykeln och säsongsmönstret uppvisar så liten felvariens som möjligt, eftersom en stor felvariens innebär en osäker komponent. Den irreguljära faktorn däremot vill vi ska absorbera så mycket brus som möjligt, det vill säga ha ett högt värde.

AR-parametern som inkluderas är, som synes, inte signifikant på 95-procentig signifikansnivå (t -värde=1,70) och samtidigt är det en nackdel med fler parametrar enligt kravet på parsimonitet.

Diagnostiken för modell 3 är dock bra, till och med aningen bättre än för modell 2. Både de observerade värdena för Ljung Box-Q och Ljung Box-Q² indikerar att residualerna varken uppvisar autokorrelation eller heteroskedasticitet. Run-testet är godkänt, vilket indikerar att residualerna också är oberoende, se diagram 4.4 för residualdiagram. Framför allt sjunker nu värdet på normalitetstestet och dessutom avsevärt då $N = 8,19$, se även appendix B för Anderson Darling-testet.

Diagram 4.4
Modell 3, Residualer



Även modell 3 karakteriseras, enligt SEATS, av en $I(2)$ för trendmodellen. SEATS dekomponerar modellen och den estimerade säsongen utgörs av 8 signifikanta månader per år, dels för den historiska estimeringen och dels för de två senaste åren samt för prognosen. Samtidigt är osäkerheten nu aningen större, ty säsongskomponentens varians ökar något jämfört med modell 2. Trendcykeln uppvisar dock ett stabilare mönster. Den estimerade felvariansen ökar marginellt både för trendcykeln och den säsongjusterade serien.

4.2.4 Modell 4

Eftersom residualerna i modell 3 fortfarande visade på viss icke-normalitet testas också en modifikation, som innebär att det kritiska värdet för outlierdetektion minskas något. Detta leder till att TRAMO estimerar en annan modell. Modell 4 är en $ARIMA(3,1,1)(0,1,1)$ med en AR-parameter, som inte är signifikant. Antalet outliers som definieras blir nu hela 30 stycken, vilket nog måste anses som orimligt många.

Modifikationen faller ut som väntat med avseende på normalitetstestet. Värdet på kurtosis sjunker och är inte längre signifikant, vilket medför att normaliteten nu också är på en godkänd nivå (se även appendix D för Anderson Darlington-testet). Den övriga diagnostiken i TRAMO är också bra. Modellen innehåller dock betydligt fler parametrar, vilket inte är önskvärt enligt parsimonitetskravet, det vill säga att modellen bör ha så få parametrar som möjligt. Värdena på BIC och AIC ökar också.

Diagnostiken från SEATS visar även att modellen estimerar en mer osäker säsongs-komponent och att variansen hos den irreguljära komponenten minskar. Samtidigt ökar den estimerade felvariansen i både trendcykeln och den säsongsjusterade serien.

4.2.5 Modell 5

Snarare än att minska på gränsen för outlierdetektion, borde kanske en modell specificeras utan att definiera några extremvärden. Orsaken till detta är att extremvärden kan antas ingå i prisbildningsmekanismen och därför borde inkorporeras i den stokastiska processen. TRAMO specificerades därför med ett kritiskt värde för outlierdetektion så högt att inga outliers kunde identifieras.

I TRAMO estimeras nu ARIMA-modellen (0,2,1)(0,1,1), vilken med parametervärdena kan skrivas som:

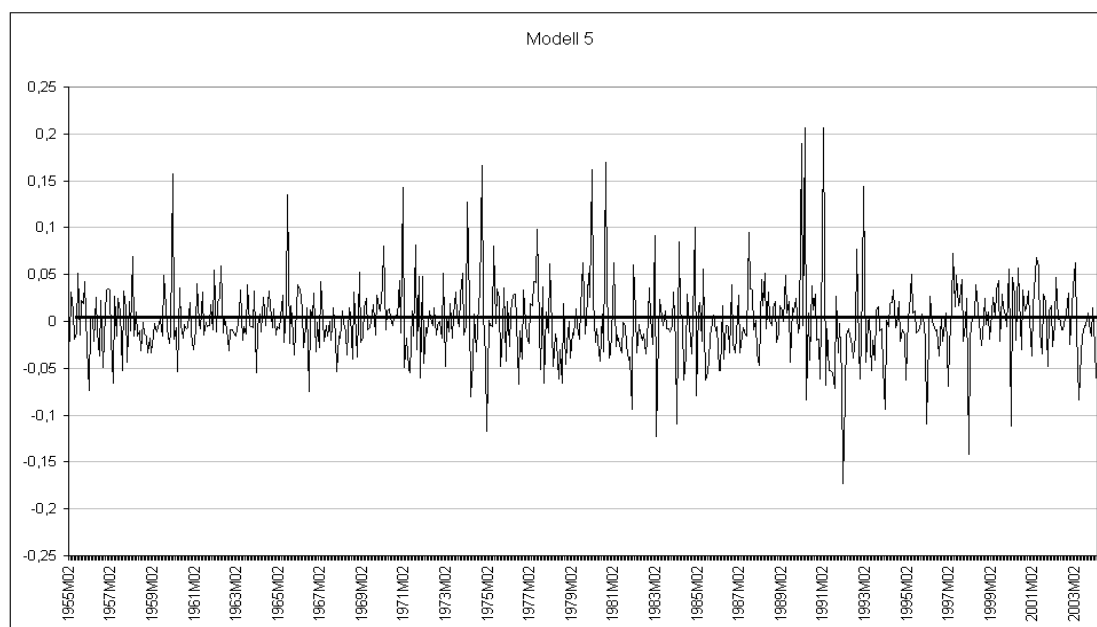
$$\Delta^2 \Delta_{12} x_t = (1 - 0,94B)_{(-66,42)} (1 - 0,84B^{12})_{(-37,06)} e_t \quad (4.2.4)$$

t -värden

Värdet för normalitetstestet skjuter nu som väntat i höjden, eftersom alla extremvärden nu inkorporeras i modellen. Detta beror framförallt, liksom tidigare, på att kurtosisvärdet är högt.

Både värdena för BIC och AIC minskar dock jämfört med modell 3 och 4 och modellen uppvisar heller ingen autokorrelation, varken på icke-säsong eller på säsong. Inte heller run-testet ger något utslag, vilket innebär att residualerna också är okorrelerade. Däremot visar Ljung Box- Q^2 på att residualerna nu lider av heteroskedasticitet, vilket också kan anas i diagram 4.5, nedan.

Diagram 4.5
Modell 5, Residualer



Vi genomförde därför även Engles ARCH-test med två laggar. Testet indikerar att modell 5 lider av ARCH-störningar då teststatistikan, NR^2 , ger ett värde på 6.81. Det kritiska värdet enligt χ^2 -fördelningen är 5.99 vid en 95% signifikansnivå. Detta innebär att nollhypotesen om ingen ARCH-störning kan förkastas.

4.3 Slutlig dekomponering

I valet mellan de estimerade modellerna står det klart att modell 1 är direkt olämplig för att dekomponeras. Jämförs modell 2 och 3 kunde vi se att diagnostiken för båda modellerna var bra med undantaget för att modell 2 uppvisade betydligt sämre värde på normalitetstestet. Detta ger att Box Cox-tranformation troligen är mer lämplig än logaritmering. Visserligen är säsongskomponenten mer stabil i modell 2, men modell 3 estimerar å andra sidan en stabilare trendcykel. Värden för variansen hos den irreguljära komponenten är densamma för de två modellerna. Modell 3 borde därmed vara att föredra, eftersom både modell 4 och 5 uppvisade klara brister blir det nu denna modell som slutligen dekomponeras.

Från diagnostiken i SEATS ges även att modellen har två rötter på trendcykeln, vilket överensstämmer med våra tidigare utförda tester för integrationsordning. Den estimerade modellen är därmed integrerad av andra ordningen, det vill säga en $I(2)$.

Nedan visas diagram på de olika komponenterna:

Diagram 4.6
Box Cox-transformerad originalserie

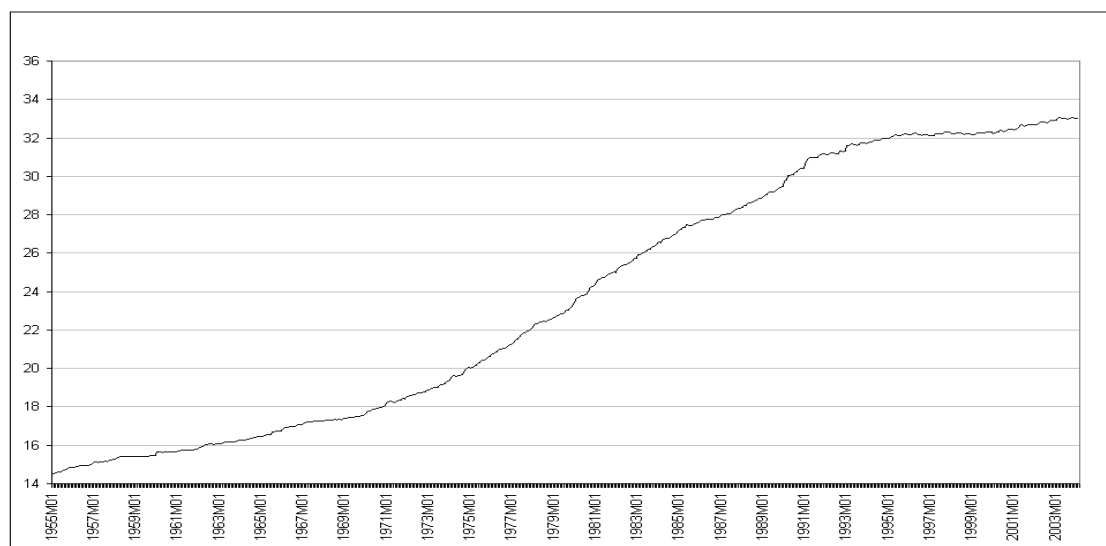
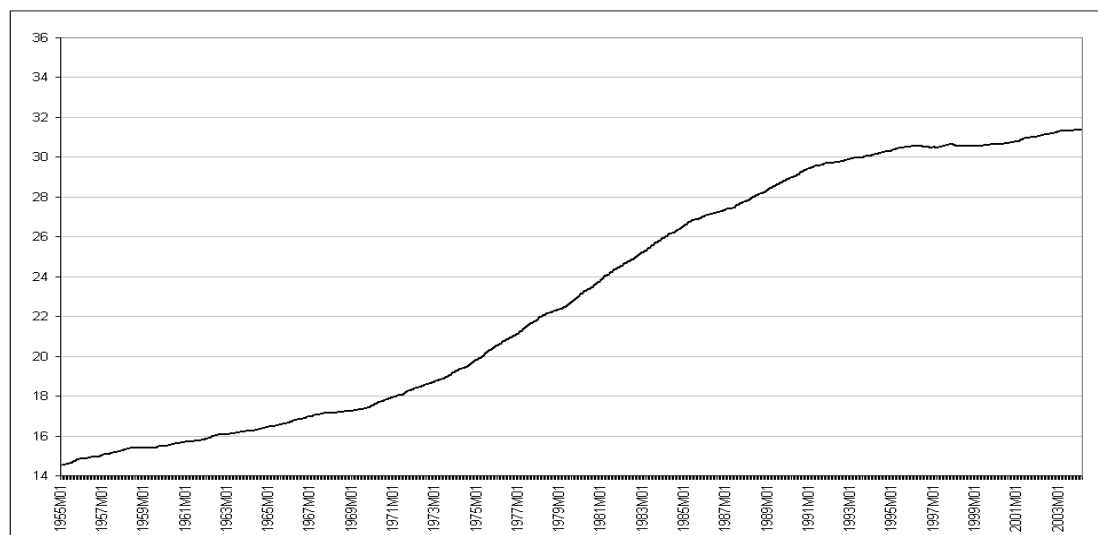
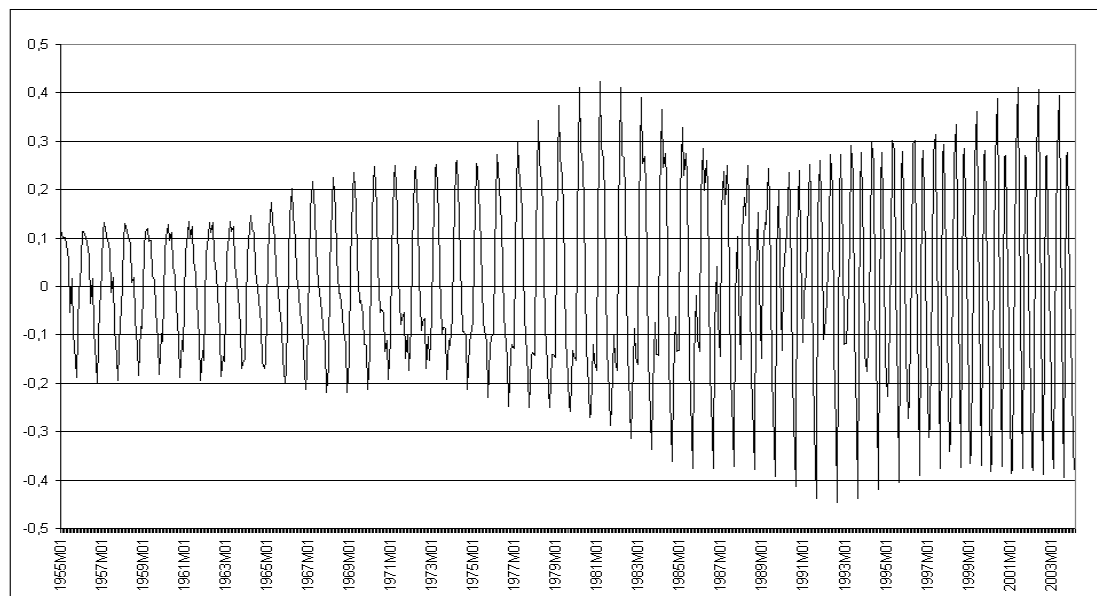


Diagram 4.7
Trendcykeln



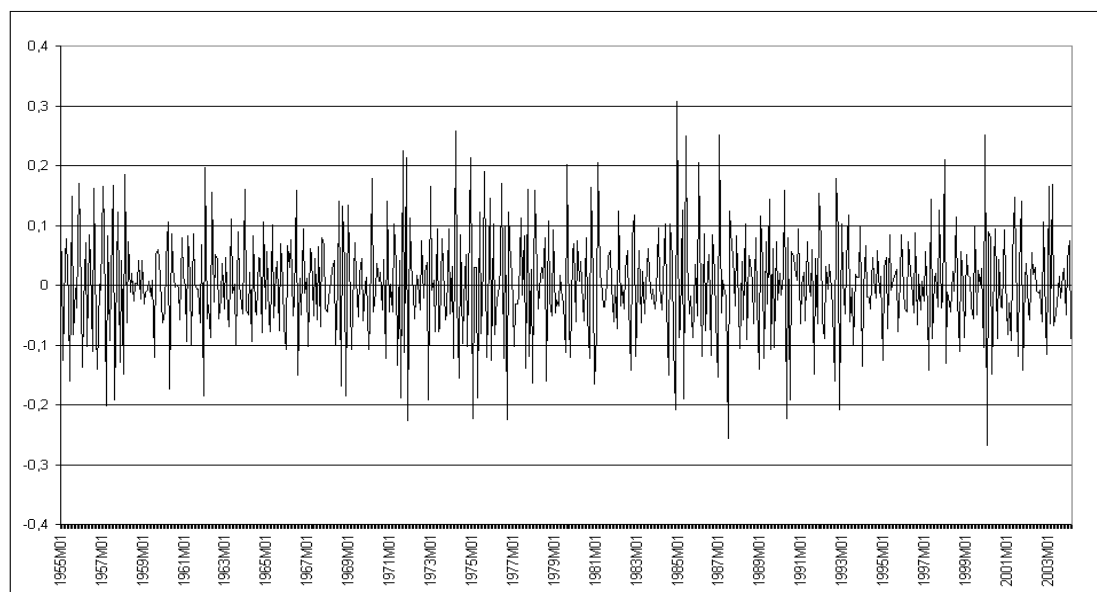
Vi ser tydligt att trendcykeln är den mest dominerande komponenten eftersom den följer originalserien väldigt väl.

Diagram 4.8
Säsongskomponent



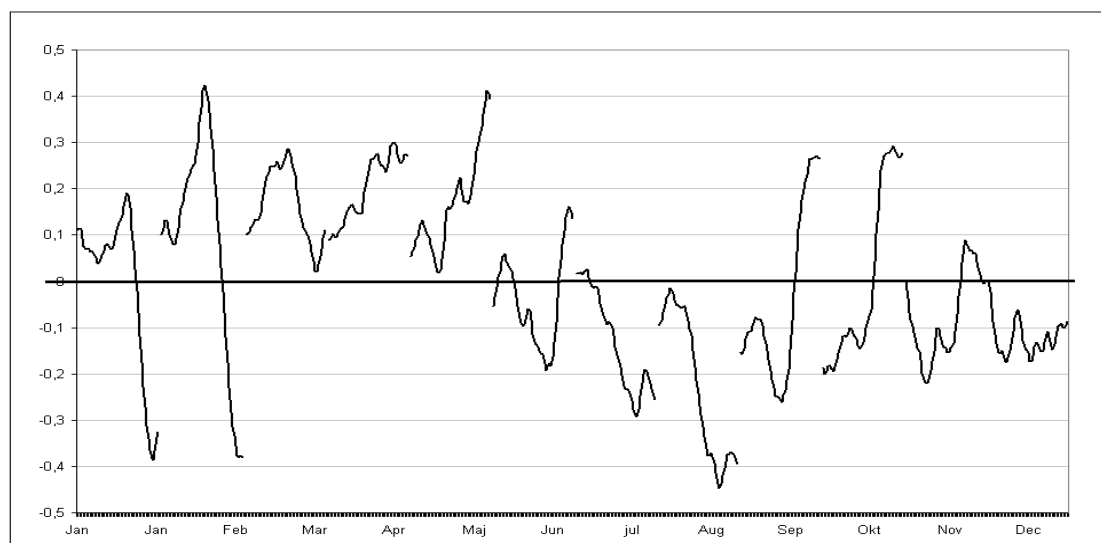
Diagrammet visar hur säsongsmönstret förändras under tidsperioden. Det är tydligt att säsongskomponenten ökar i magnitud under tidsperioden. Samtidigt verkar också komponenten bli alltmer komplex under den andra halvan av tidsperioden. Inte bara förändras mönstret, utan flera månader inkluderas också. Ett annat sätt att illustrera hur säsongsmönstret för de olika månaderna förändras under tidsperioden är att visa ett diagram bestående av de olika månadernas förändringar var för sig, se diagram 7.10.

Diagram 4.9
Irreguljär komponent



Den irreguljära komponenten karakteriseras av ett konstant medelvärde och ser även ut att vara homoskedastisk.

Diagram 4.10
Säsongskomponent per månad



Diagrammet ovan visar att säsongsmönstret för de olika månaderna följer tydliga mönster och inte beter sig slumpmässigt. För månaderna januari och februari ses detta mycket tydligt. Från att ha uppvisat en uppåtgående trend mellan 1950 och 1980-talet sker sedan ett trendbrott och de båda månaderna övergår till negativa värden i början på 1990-talet. Samtidigt ser vi att månaderna mars, april och maj uppvisar positiva värden under hela tidsperioden, vilket implicerar att prisserien har varit som högst, i genomsnitt, under denna period. Ser vi till nuvarande säsongsmönster i prisserien är KPI som högst under april och maj samt på hösten, under september och oktober. Under sommarmånaderna juli och augusti tillsammans med januari och februari är de månader då priserna i konsumentpriskorgen verkar vara som lägst.

Tabell 4.2
Diagnostik för modellerna 1 till 5, estimerade i TRAMO

TRAMO					
1955-2004	Ej transf.	Log	Box-Cox	Box-Cox	Box-Cox
ARIMA-modell:	(3,1,1)(0,1,1)	(0,2,1)(0,1,1)	(1,2,1)(0,1,1)	(3,1,1)(0,1,1)	(0,2,1)(0,1,1)
Modell:	1	2	3	4	5
1. AIC	3705	-4594	-2372	-2473	-2012
2. BIC	3,81	-10,71	-6,83	-6,88	-6,35
3. Standard avvikelse (medelvärde)	0,25	0,000	0,00	0,00	0,00
4. Standard avvikelse (residualerna)	5,81	0,00	0,03	0,03	0,04
5. Normalitet < 6, (95%, 2 d.f)	30,65	25,60	8,19	4,62	682,00
6. Skevhet	-0,01	0,18	0,10	0,21	0,91
7. Kurtosis	4,16	3,98	3,56	3,18	8,01
8. Ljung-Box Q (24 autok) <34.1 (95%, 20 d.f)	27,03	19,80	22,94	19,63	19,54
9. Ljung-Box Q ² <36,4 (95%, 24 d.f)	171,98	25,12	21,50	20,36	97,02
10. Pierce Q (2 autok)<9 (95%, d.f)	0,00	1,64	1,15	0,00	0,00
11. Run-test, t-värde	2,90	1,44	0,42	1,3695	-0,17
12. Parametervärden (t-värden)	-0,98 (-10,09)	-0,92 (-56,27)	0,07 (1,70)	-0,86 (-11,98)	-0,94 (-66,42)
	0,20 (3,16)	-0,84 (-36,71)	-0,93 (-58,69)	0,06 (1,14)	-0,84 (-37,06)
	-0,13 (-2,60)		-0,78 (-27,90)	-0,13 (-2,83)	
	-0,76 (-8,51)			-0,80 (-12,97)	
	-0,64 (-17,74)			-0,74 (-22,09)	
13. Antal outliers	26	16	17	30	0
SEATS					
1955-2004	Nivåer	Log	Box Cox	Box Cox	Box Cox
ARIMA-modell:	(3,1,1)(0,1,1)	(0,2,1)(0,1,1)	(1,2,1)(0,1,1)	(3,1,1)(0,1,1)	(0,2,1)(0,1,1)
Modell:	1	2	3	4	5
1. Varians i komponenterna (i enheter of Va)					
Trend-cykel	0,080	0,261	0,178	0,063	0,258
Säsong	0,167	0,004	0,011	0,018	0,004
Irreguljär	0,030	0,090	0,090	0,050	0,091
2. Estimerad felvarians (i enheter av Va)					
Trend-cykel	0,453	0,272	0,296	0,348	0,273
Säsongsjusterad	0,396	0,139	0,170	0,192	0,138
3. Signifikant säsong-månader per år (95%)					
Hela serien	10	5	8	10	6
Senaste två åren	9	4	8	8	3
Prognos för nästa år	9	4	8	8	3

5 Sammanfattning och avslutande kommentarer

Vi har i föreliggande uppsats visat ett tydligt säsongsmönster i konsumentpriserna. Vi har även visat att den svenska konsumentpriserien, mellan 1955 och 2004, är integrerad av andra ordningen, det vill säga att trendmodellen uppvisar en dubbel enhetsrot. Detta konstaterade vi efter att två olika enhetsrotstest samt ett test för stationaritetstest genomförts. Dessa resultat överensstämde sedan med de modeller vi estimerade genom att tillämpa programmet TRAMO. Innebörden av att KPI-serien uppvisar en dubbel enhetsrot är att inflationen under tidsperioden inte är stationär, utan integrerad av första ordningen. Detta implicerar att en chock på prisserien påverkar inflationen på ett permanent sätt Juselius (2004).

Ett flertal olika ARIMA-modeller har estimerats i denna uppsats. Den första modellen erhöles genom att använda den automatiska inställningen i TRAMO och var en $ARIMA(3,1,1)(0,1,1)$ för seriens nivåer. Modell 1 uppvisade dock betydliga problem. Framförallt karakteriserades modellens residualer av heteroskedasticitet och leptokurtosis. Vi försökte lösa dessa problem genom att specificera om modellen med andra parametrar. Problemen försvann dock inte genom denna metod. Istället gick vi över till en logaritmerad specifikation av KPI-serien och använde återigen den automatiska inställningen i TRAMO.

Resultatet blev modell 2, en $ARIMA(0,2,1)(0,1,1)$ och problemet med heteroskedasticitet var löst. Residualerna uppvisade dock fortfarande ickenormalitet, trots en viss förbättring. Vi försökte därför, på samma sätt som tidigare, att lösa problemet genom att specificera om modellen med andra parametrar. Återigen misslyckades vi med denna metod. Genom att studera residualdiagrammet från modell 2 kunde vi dock ana att den logaritmerade transformationen varit alltför kraftfull. Vi undersökte därför om en annan transformation kunde passa bättre.

Vi testade ett flertal olika Box Cox-transformationer och den som gav bäst resultat, utifrån diagnostiken i TRAMO, resulterade i modell 3, en $ARIMA(1,2,1)(0,1,1)$. Problemet med ickenormaliteten hos residualerna visade sig nu vara nästintill löst och den övriga diagnostiken var också bra. En AR-parameter, som inte var signifikant på 95-procentnivån ingick dock i modellen. Osäkerheten i den estimerade säsongskomponenten ökade marginellt. Å andra sidan minskade osäkerheten i trendcykelkomponenten. Det visade sig att försöken med att specificera om modellens parametrar inte gav något resultat denna gång heller, vilket medförde att vi undersökte en alternativ modifikation.

För att försöka komma till rätta med ickenormaliteten minskade vi därför det kritiska värdet för outlierdetektion, vilket resulterade i att residualerna, slutligen, kunde antas vara normalfördelade. Modell 4, en $ARIMA(3,1,1)(0,1,1)$, var dock betydligt mer komplex än modell 3 och kan därför kritiseras utifrån parsimonitetskravet. Modellen innefattade även 30 outliers, vilket vi ansåg vara orimligt.

Istället testades en sista modell, som utgick från antagandet att extremvärden ingår i prisbildningsmekanismen och därmed inte borde korrigeras. I modell 5 ändrades därför inställningen i TRAMO så att inga extremvärden definierades. Som väntat gav detta oss, återigen, problem med normaliteten (leptokurtosis), men även med heteroskedasticitet. I övrigt var diagnostiken bra. Samtidigt kan vi naturligtvis inte utesluta att prisserier ska karakteriseras av kurtosis och möjligen även heteroskedasticitet.

En tänkbar lösning på problemen i modell 5 vore dock att utveckla modellen till att inkludera en betingad variansstruktur, det vill säga att modellera KPI-serien inom ramen för ARCH- och GARCH-familjen. Denna tänkbara lösning ryms dock inte

inom ramen för den här uppsatsen utan får betraktas som ett projekt för en vidare undersökning.

Gemensamt för de estimerade modellerna i den föreliggande uppsatsen var att de alla uppvisade ett signifikant säsongsmönster. Intressant nog innehöll alla våra estimerade modeller en signifikant MA-parameter på säsongsnivå. Hur denna säsongskomponent ska beskrivas är däremot inte givet, ty det beror på vilken modell som appliceras på serien. Vi valde att dekomponera modell 3, vilket medförde att ett säsongsmönster med åtta signifikanta månader estimerades i SEATS. Detta estimat för säsongskomponenten gällde likväl för den historiska estimeringen som för estimeringen av de senaste två åren och ettårsprognosen. Vi har också visat hur säsongsmönstret förändrats under tidsperioden genom ett diagram på de enskilda månadernas säsongsmönster över tid, se diagram 4.10, sidan 16.

För att avslutningsvis återknyta till frågorna i vår syftesformulering:

- Vi har, genom att applicera tre olika hypotestester, visat att KPI-serien är integrerad av andra ordningen. Inför skrivandet av den här uppsatsen inspirerades vi av Professor Katarina Juselius, som tidigare i år skrivit om $I(2)$ -processer. Vi fann sedan, intressant nog, att den svenska KPI-serien karakteriserades av en $I(2)$. Innebörden, av detta är att inflationen, i sin tur, inte är stationär utan karakteriseras av en $I(1)$, vilket betyder att en chock på inflationen karakteriseras av persistens.
- Vi har även, genom att tillämpa TRAMO, anpassat ett antal ARIMA-modeller till serien. Ingen av dessa kunde dessvärre gälla som en fullt adekvat modell till serien.
- De estimerade modellerna i den här uppsatsen innehöll dock alla en signifikant säsongskomponent. De fem modeller, vilka vi redovisat för under resultatavsnittet uppvisade alla en signifikant MA-parameter på säsongsfrekvensen, vilket vi finner högst anmärkningsvärt. Detta indikerar att det existerar ett signifikant säsongsmönster i konsumentprisserien under tidsperioden 1955-2004. Säsongsmönstret har självklart förändrats under tidsperioden, vilket vi illustrerade i ett diagram för de enskilda månaderna, men verkar också ha ökat och blivit mer komplext (se diagram 7.8 och 7.10). Slutsatsen blir därmed att en säsongsjusterings av konsumentprisindexserien borde beaktas.

5.1 Framtida forskning

Vi ser framemot att ta del av en studie som undersöker KPI-serien på en disaggregerad nivå. Med andra ord en studie som undersöker KPI-serien olika komponenter: "Kläder och Skor", "Livsmedel", "Bostad, Värme och Hushållsel". Eftersom vi i föreliggande uppsats har estimerat ett signifikant säsongsmönster i konsumentpriserna vore det intressant att vidare undersöka hur säsongsmönstret ser ut utifrån indexets olika varugrupper.

En annan intressant studie vore att modellera KPI-serien inom ramen för ARCH- och GARCH-familjen. Detta då Engles ARCH-test indikerade att modell 5, det vill säga den modell där extremvärden antogs vara inkorporerade i KPI-processen, led av ARCH-störningar.

6 Litteraturförteckning

Black, F. (1976), "Studies in Stock Price Volatility Changes", Proceeding of the 1976 meeting of the Business and Economic Statistics Society, *American Statistical Association*, 177-181.

Bowerman, B. L. och O'Connell, R. T. (1993), *Forecasting and Time Series - an Applied Approach*, Third Edition, Duxbury Press, Belmont.

Box, G.E., Jenkins, G.M. och Reinsel, G.C. (1976), *Time Series Analysis, Forecasting and Control*, Second Edition, Holden-Day, San Francisco.

Box, G.E, Jenkins, G.M. och Reinsel, G.C. (1994), *Time Series Analysis, Forecasting and Control*. Third Edition. Holden-Day: San Francisco.

Box, G.E och Cox, D.R. (1964), "An Analysis of Transformation", *Journal of Royal Statistical Society*, B 26, s 211-243.

Gomez, V. och Maravall, A. (1994), " Estimation, Prediction and Interpolation for Nonstationary Series with the Kalman Filter", *Journal of the American Statistical Association*, 89, 611-624.

Gómez, F. och Maravall, A. (1996), "Programs TRAMO and SEATS. Instructions for the User (Beta version: September 1996)", *Working Paper No. 9628*, Bank of Spain.

Gómez, V. och Maravall, A. (1998). "Guide for Using the Programs TRAMO and SEATS". *Working Paper 9805*, Research Department, Bank of Spain.

Gujarati, D. (1999), *Essentials of Econometrics*, 2nd edition, Irwin/McGraw-Hill, Singapore.

Harvey, A.C. (1981), *Time Series Models*, Phillip Alan, Oxford.

Juselius, K. (2004), "Inflation, Money Growth and I(2) Analysis", Institute of Economics, University of Copenhagen.

Kaiser, R. och Maravall, A. (2001), *"Measuring Business Cycles in Economic Time Series"*, Springer-Verlag, New York.

Kleinbaum, G. D., Kupper, L. L. och Muller, E. K. (1988). *Applied Regression Analysis and other Multivariable methods*, Second Edition, Duxbury Press, Belmont.

MacKinnon, J. G., (1991), "Critical Values of Cointegration Tests", in Engle, R. E. och Granger C. W. J., eds., *Long-Run Economic Relationships: Readings in Cointegration*, Chapter 13, Oxford University Press, New York.

Makridakis, S., Wheelwright, S. och Hyndman, R. J. (1998), *Forecasting Methods and Applications*, Wiley, New York.

Maravall, A. (1995). "Unobserved Components in Economic Time Series", in Pesaran, H. & Wickens, M. (eds), *The Handbook of Applied Econometrics*, vol. 1, Basil Blackwell, Oxford.

Maravall, A, och Sánchez, F. (2000). "An Application of TRAMO-SEATS: Model Selection and Out-of-Sample Performance: the Swiss CPI series", *Working Paper 9914*, Research Department, Bank of Spain. To appear in *Seasonal Adjustment Procedures. Experiences and Perspectives*, Rome.

Mcleod, A.I och Li W.K (1983), "Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelation", *Journal of Time Series Analysis*, 4, 269-273.

NE Nationalencyklopedin AB (2000), Multimedia 2000 plus, DVD skiva.

Planas, C. (1997), *Applied Time Series Analysis, Modelling, Forecasting, Unobserved Components Analysis and the Wiener-Kolmogorov Filter*, Office for official publications of the European Communities, Luxembourg.

Pierce, D.A. (1978). "Seasonal Adjustment when Both Deterministic and Stochastic Seasonality are Present, *Seasonal Analysis of Economic Time Series*", ed. A.Zellner, Washington, D.C. U.S. Dept. of Commerce, Bureau of the census, 242-269.

Priestley, M. P. (1981), *Spectral Analysis and Time Series*, Academic Press Ltd, London.

Öhlén, S. (2003), "Säsongrensning av nationalräkenskaperna, en översikt", Statistiska centralbyrån.

Appendix A

I programmen TRAMO och SEATS används ARIMA-modeller som en integrerad del både i beräkningen av extremvärden och vid dekomponeringen av serien i en säsongsfaktor och i andra faktorer. För att läsaren lätt ska kunna följa med i användandet av ARIMA-modeller i uppsatsen presenteras här en sammanfattande genomgång av den teoretiska uppbyggnaden för ARIMA-modeller.

Wolds dekomponering

Wolds dekomponeringsteorem fastslår att en andra ordningens stationär process, y_t , kan uttryckas som summan av en deterministisk funktion tillsammans med en oändlig linjär kombination av okorrelerade slumpvariabler enligt:

$$y_t = c(t) + e_t + \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \dots \quad (\text{A.1})$$

$$y_t = c(t) + \sum_{i=0}^{\infty} \varphi_i e_{t-i} \quad (\text{A.2})$$

$$y_t = c(t) + \varphi(B)e_t, \quad (\text{A.3})$$

där $c(t)$ representerar den deterministiska komponenten, $\varphi_0 = 1$, B är en bakåtskiftande operator och $\sum_{j=0}^{\infty} \varphi_j^2 < \infty$. Processen kan därmed uttryckas som summan av en deterministisk komponent och en linjär stokastisk process. Den deterministiska delen uttrycker vanligtvis seriens medelvärde och den stokastiska delen representerar en oändlig glidande medelvärdesprocess.

Arima-modeller

Enligt Wolds teorem kan sökandet efter en modell till serien y_t betraktas som sökandet efter en lämplig representation till den stokastiska processen $\varphi(B)e_t$. Gruppen av ARIMA-modeller erbjuder ett kraftfullt verktyg för att approximera denna del i Wolds dekomponering, vilket dessutom kan göras med relativt få parametrar. För en utförligare beskrivning av Wolds teorem och ARIMA-modeller se exempelvis Planas (1997).

Autoregressiv Modell (AR)

Om vi antar att φ -koefficienten är sådan att $\varphi_i = \varphi^i$ och $|\varphi| < 1$ kan y_t skrivas som:

$$y_t = (1 + \varphi B + \varphi^2 B^2 + \dots + \varphi^n B^n) e_t = \quad (\text{A.4})$$

$$= \frac{1 - \varphi^{n+1} B^{n+1}}{1 - \varphi B} e_t = \quad (\text{A.5})$$

$$= \frac{1}{1 - \varphi B} e_t \quad \text{då} \quad \lim_{n \rightarrow \infty} \varphi^n = 0. \quad (\text{A.6})$$

Arrangeras ovanstående ekvation om så erhålls en AR(1)-process, vilken kan uttryckas enligt:

$$y_t = \varphi y_{t-1} + e_t. \quad (\text{A.7})$$

AR(1)-modellen kan sedan utvecklas för att även beskriva mer generella mönster. En autoregressiv process av ordningen p kan skrivas enligt:

$$y_t + \varphi y_{t-1} + \dots + \varphi_p y_{t-p} = e_t, \quad (\text{A.8})$$

vilket benämns som en AR(p)-process.

Glidande Medelvärdesmodell (MA)

Om vi sedan antar att φ -koefficienten i ovanstående generella AR-process (ekvation A.8) är sådan att $\varphi_i = \theta^i$ med $|\theta| < 1$ så kan följande ekvation skrivas:

$$e_t = (1 + \theta B + \theta^2 B^2 + \dots) y_t = \quad (\text{A.9})$$

$$= \frac{1}{1 - \theta B} y_t \quad \text{då } \lim_{n \rightarrow \infty} \theta^n = 0. \quad (\text{A.10})$$

Genom att arrangera om ovanstående ekvation erhålls en MA(1)-process, vilken kan uttryckas enligt:

$$y_t = (1 - \theta B) e_t. \quad (\text{A.11})$$

MA(1)-processen kan därmed uttrycka en kort representation av en oändlig men konvergent AR-process. Detta kräver dock med nödvändighet att $|\theta| < 1$, vilket kallas för inverterbarhetskriteriet.

En generell MA-processen av ordningen q skrivs:

$$y_t = e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} \quad (\text{A.12})$$

och benämns som en MA(q)-process.

Autoregressiv – Glidande Medelvärdesmodell (ARMA)

Vissa processer uttrycks dock bäst av modeller, som kombinerar en AR(p)-process med en MA(q)-process. Dessa modeller benämns som ARMA-modeller av ordningen p och q och skrivs som:

$$y_t + \phi y_{t-1} + \dots + \phi_p y_{t-p} = e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q}. \quad (\text{A.13})$$

Genom att använda sig av den bakåtskiftande operatören kan ekvationen dock uttryckas i betydligt kompaktare form enligt:

$$\phi(B) y_t = \theta(B) e_t. \quad (\text{A.14})$$

Integrerad Autoregressiv – Glidande Medelvärdesmodell (ARIMA)

Ekonomiska tidsserier uppfyller ofta inte villkoren för stationaritet, vilket vanligtvis beror på att det förekommer långsiktiga trender i dessa serier⁴. Många ickestationära tidsserier kan däremot omvandlas till att bli stationära genom en lämplig transfor-

⁴ Ekonomiska tidsserier liknar ofta slumpvandringar, vilka inte är stationära.

mation. Processer som gjorts stationära genom första ordningens differens, det vill säga:

$$z_t = y_t - y_{t-1} = \Delta y_t, \quad (\text{A.15})$$

benämns som integrerade av första ordningen, vilket skrivs $I(1)$. En process som gjorts stationär genom andra ordningens differens benämns således som en $I(2)$ och så vidare.

En generell stokastisk ARIMA-process, det vill säga en autoregressiv integrerad glidande medeltalsmodell, utgörs av en blandad ARMA-process för en tidsserie som gjorts stationär genom differentiering. För tidsserier utan säsongsvariation kan en sådan modell skrivas som:

$$\varphi(B)\Delta^d y_t = \theta(B) e_t, \quad (\text{A.16})$$

där $\varphi(B)$ och $\theta(B)$ är polynom, som uppfyller stationaritets- och inverterbarhets-kriterierna⁵. $\Delta = 1 - B$ är en differentieringsoperator, d står för antal differentieringar som använts för att göra serien stationär och e_t är vitt brus.

Om polynomen $\varphi(B)$ och $\theta(B)$ är av ordningen p respektive q , så sägs y_t följa en $ARIMA(p,d,q)$ -modell. Box och Jenkins (1976) utvecklade sedan ARIMA-modelleringen för att även kunna beskriva tidsserier med säsongsvariation, vilket kan skrivas som:

$$\varphi(B)\Phi(B^s)\Delta^d\Delta_s^D y_t = \theta(B)\Theta(B^s) e_t, \quad (\text{A.17})$$

där parametern D representerar det minsta antal differentieringar som behövs för att göra serien stationär och $\Delta_s = 1 - B^s$ står för differentieringsoperatoren avseende säsongen. Ovanstående ARIMA-modell sägs vara av ordning $(p,d,q)(P,D,Q)$.

⁵ För en definition av begreppet stationaritet och inverterbarhetskriteriet se exempelvis Bowerman och O'Connell (1993).

Appendix B

Appendixet redogör, kortfattat, för de statistiska tester som används i studien, vilket bland annat inkluderar test för integrationsordning, normalitet, autokorrelation och heteroskedasticitet.

Integrationsordning

I uppsatsen har vi använt sig av två olika test för att undersöka KPI-seriens integrationsordning. Dessa har varit det utvidgade Dickey Fuller-testet och Phillips Perron-testet.

Utvidgat Dickey Fuller

Dickey Fuller-testet (Dickey och Fuller 1979, Fuller 1976) undersöker antalet enhetsrötter i en tidsserie. Vid en AR(1)-process utgår härledningen från följande ekvation:

$$y_t = \phi y_{t-1} + v_t. \quad (\text{B.1})$$

Därefter subtraheras y_{t-1} från bägge membra:

$$y_t - y_{t-1} = (\phi - 1)y_{t-1} + v_t, \quad (\text{B.2})$$

där $\phi - 1 = \delta$, vilket innebär att formel (B.2) kan omformuleras till:

$$\Delta y_t = \delta y_{t-1} + v_t. \quad (\text{B.4})$$

Nollhypotes är att serien har en enhetsrot, det vill säga att $\phi - 1 = 0$. Alternativhypotesen är att serien är stationär.

Testet gäller dock endast om residualerna (v_t) i genererar vitt brus. Om det finns autokorrelation i termen v_t utökas modellen med AR-termer tills det att termen v_t är blekt. Det utökade Dickey Fuller-testet (augmented DF, ADF) testet kan skrivas enligt:

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-1} + v_t. \quad (\text{B.5})$$

Phillips Perron

Phillips Perron-testet (Phillips och Perron, 1988) är ett annat test som kan användas för att undersöka antalet enhetsrötter. Testet bygger på DF-testet, men tillämpar en något annorlunda strategi jämfört med ADF-testet. Problemet med seriekorrelation i feltermen hanteras inte med laggade differentierade variabler, utan istället med parametriska metoder. Vanligtvis ger de båda testerna samma resultat.

Stationaritet

En alternativ metod för att undersöka integrationsordningen hos serien är att applicera ett så kallat stationaritetstest.

Kpss

KPSS-testet (Kwiatkowski, Phillips, Schmidt och Shin, 1992) utgår istället från nollhypotesen att serien är nivåstationär, alternativt trendstationär. Alternativhypotesen är att serien innehar en enhetsrot (till skillnad från Phillips Perron-testet och ADF-testet, där nollhypotesen är ickestationär och alternativhypotesen är stationär).

Teststatistikan i KPSS-testet definieras enligt:

$$KPSS = \sum_{t=1}^T \frac{R_t^2}{\hat{\sigma}^2}, \quad (\text{B.6})$$

$$\text{där } R_t = \sum_{r=1}^t e_r. \quad (\text{B.7})$$

T representerar antalet observationer, $\hat{\sigma}^2$ är feltermens varians och e_r är residualerna från en OLS-regression, där Y_t estimerats med en konstant och eventuellt en tidstrend.

Bic

Vid identifieringen av ARIMA-modeller i programmet TRAMO läggs ofta speciell vikt vid måttet *BIC* (Bayes Information Criterion) (Box med flera, 1994). Den modell som minimerar absolutvärdet på *BIC* är enligt kriteriet också den bästa modellen. För att *BIC*-värdet ska anses gälla måste dock förutsättningarna för inferens vara uppfyllda. För en utförlig diskussion om modellval hänvisas till Schwarz (1978) och Shibata (1976). *BIC*-värdet används följaktligen till att jämföra de olika modellerna som anpassats till en given tidsserie. Värdet bygger på likelihoodfunktionen, L , och definieras enligt:

$$BIC = -2 \log(L) + K \log(T), \quad (\text{B.8})$$

där K representerar antalet parametrar i modellen och T antalet observationer, som används vid skattningen av modellens parametrar. Genom att öka komplexiteten hos en tidsseriemodell, det vill säga att inkludera flera parametrar, kan värdet på L ökas, vilket dock kan leda till en så kallad överparametrisering av modellen. *BIC*-måttet kan därmed ses som en justering av likelihoodfunktionen på så sätt att ett straff utdelas för antalet parametrar, K , som modellen innehåller.

Aic

Ett mått som liknar *BIC* är *AIC* (Akaike's Information Criterion), som definieras enligt:

$$AIC = -2 \log(L) + 2K. \quad (\text{B.9})$$

En skillnad mellan *AIC* och *BIC* är att *AIC* har en tendens att godkänna överparametriserade modeller (Harvey, 1981).

Skevhhet och Kurtosis

Skevhhet kan översättas med ordet asymmetri och innebär att den ena svansen i den klockformade normalfördelningskurvan är mer utdragen än den andra. För normalfördelningen ska värdet på skevheten vara 0. Värdet definieras enligt:

$$\frac{E((x - \mu)^3)}{(Var(x))^{3/2}} \quad (B.10)$$

Kurtosis är ett mått på i vilken utsträckning värdena befinner sig nära medelvärdet av fördelningen eller i svansarna. Man skiljer på flera olika typer av kurtosis, vilka benämns som:

- Mesokurtosis är benämningen på fördelningar med kurtosisvärde likt normalfördelningen.
- Platykurtosis innebär att fördelningen har ett kurtosisvärde, som är mindre än normalfördelningens värde. Fördelningen har ett relativt tjockt mittparti på båda sidorna av medelvärdet och en låg toppighet
- Leptokurtosis är en beteckning för ett kurtosisvärde, som är större än normalfördelningens värde. Fördelningen har mer sannolikhetsmassa i svansarna, så kallade "tjocka" svansar, och är toppigare än normalfördelningen.

För normalfördelningen ska värdet på kurtosis vara 3 och beräknas enligt följande:

$$\frac{E((x - \mu)^4)}{(Var(x))^2} \quad (B.11)$$

Normalitet

Ett mer formellt sätt att undersöka om fördelningen för de skattade residualerna är symmetriskt fördelade är att genomföra ett normalitetstest. De så kallade skevhets- och kurtosistesten beräknas utifrån det tredje respektive fjärde momentet hos den antagna stokastiskt normalfördelade variabeln i Wold-dekomponeringen. Skevhets- och kurtosistesten definieras enligt:

$$S = \sqrt{T} \frac{m_3}{\hat{\sigma}^3} \sim N(0,6) \quad (B.12)$$

$$K = \sqrt{T} \frac{m_4 - 3}{\hat{\sigma}^4} \sim N(0,24) \quad (B.13)$$

Testerna blir signifikanta då alltför många residualer har stora värden (i absoluta tal) och kan därför användas för att testa förekomsten av extremvärden hos residualerna.

Bowman Shenton

Statistikorna S och K kan sedan kombineras till ett test för normalitet, vilket benämns Bowman Shenton-testet. Testet är asymptotiskt fördelat som en χ^2 -fördelning (Planas, 1997) och teststatistikan, N , beräknas utifrån t-värdena för de estimateerade statistikorna S och K enligt:

$$N = S^2 + K^2 \quad (B.14)$$

Anderson Darling

Anderson Darlings normalitetstest utgår från teststatistikan A^2 . Den underliggande nollhypotesen är att variabeln är normalfördelad och statistikan definieras enligt:

$$A^2 = -n - S \quad (\text{B.15})$$

där

$$S = \sum_{i=1}^n \frac{(2i-1)}{n} [\ln F(z_i) + \ln(1 - F(Z_{n+1-i}))]$$

F antas vara normalt fördelat, n är urvalsstorleken och z_i är det i :te sorterade, standardiserade, beräknade värdet.

Autokorrelation

Ett formellt test för autokorrelation utgår från att den skattade autokorrelationen är normalfördelad med variansen $1/T$. Testet innebär att följande teststatistiska beräknas:

$$Q_m = T \sum_{i=1}^m r^2(\hat{e}) \quad (\text{B.15})$$

Statistikan utgår från signifikansen hos de m första autokorrelationerna. Valet av m är inte givet, men vid månadsdata är det vanligt att betrakta autokorrelationen två år tillbaka, vilket ger $m=24$. Eftersom $r(\hat{e})$ är normalt och oberoende fördelat så har Q_m -statistikan en χ_m^2 -fördelning för vitt brus (Planas, 1997).

Ljung box-Q

Om residualerna kommer från en anpassad ARIMA-modell måste dock två korrigeringar göras, vilka föreslogs av Ljung och Box (1978). De visade att, för små skattningar, uttrycket $(T-i)/T(T+2)$ bättre kan approximera variansen hos $r(\hat{e})$ än $1/T$, vilket ger följande form på Ljung Box-Q:

$$Q_m = T(T+2) \sum_{i=1}^m r^2(\hat{e}) / (T-i), \quad (\text{B.16})$$

där $T=n-d$ och n står för antalet observationer. De visade också att en bättre approximation av fördelningen ges av en χ^2 -fördelning med $(m-p-q)$ frihetsgrader (Planas, 1997).

Box Pierce-Q

Ljung Box-teststatistikan kan sedan modifieras så att den även innefattar specifika säsongslaggar. Pierce (1978) visade, till exempel för månadsdata, att signifikansen för autokorrelation hos residualerna vid 12 respektive 24 månaders lag kan testas genom följande statistika:

$$Q_s = T(T+2) [r^2(12)/(T-12) + r^2(24)/(T-24)], \quad (\text{B.17})$$

vilken kan approximeras med en χ^2 -fördelning. Statistikan testar alltså för frånvaron av säsongsmässig autokorrelation (Pierce, 1978).

Run-test

Run-testet är ett ickeparametriskt test som testar slumpmässigheten med avseende på residualernas tecken⁶. Nollhypotesen är att residualerna är oberoende och testet utgår från det förväntade antalet runs, sekvenser av positiva eller negativa residualer, som erhålls för en slumpmässig serie av N observationer.

Heteroskedasticitet

Om ingen signifikant autokorrelation kan upptäckas blir slutsatsen att det saknas en linjär struktur hos modellens residualer. Men ett beroende kan ändå existera på grund av en underliggande icke linjär struktur. McLeod och Li (1983) visade att Ljung Box-teststatistika $Q_m(a_t)$ har samma fördelning som statistikan $Q_m(a_t^2)$, vilket kan användas som en indikation på icke linjäritet det vill säga heteroskedasticitet, se exempelvis Mills (1990). Detta i sin tur kan vara en indikation på att den specificerade modellen lider av ARCH-effekter. På liknande sätt kan även Box Pierce-Q testa för heteroskedasticitet.

Engles Arch-test

Engles ARCH-test används för att undersöka om den estimerade modellen har betingad tidsvarierande volatilitet. Testet baseras på en OLS-regression. Regressionens residualer (ε_t) prediktieras först för att sedan kvadreras. Därefter estimeras variabeln med de kvadrerade residualerna mot tidsfördröjda kvadrerade residualer och en konstant enligt följande:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_p \hat{\varepsilon}_{t-p}^2 \quad (\text{B.18})$$

Teststatistikan, NR^2 , består av regressionens förklaringsgrad, R^2 , multiplicerad med antalet observationer, N . Statistikan är asymptotiskt χ^2 -fördelad med p frihetsgrader under nollhypotesen: ingen ARCH-effekt (Engle, 1982). Testet gäller även för GARCH-modeller enligt Bera och Higgins (1993).

⁶ I ett ickeparametriskt test gör man inga eller få antaganden om den fördelning från vilken observationerna kommer.

Appendix C

I Appendix C redovisas diagram som visar första och andra differensen samt första och andra differensen och säsongdifferensen för originalserien, den logaritmerade serien samt den Box-Cox transformerade serien med transformationsparametern satt till 0,28.

Diagram C.1

1:a differensen av originalserien

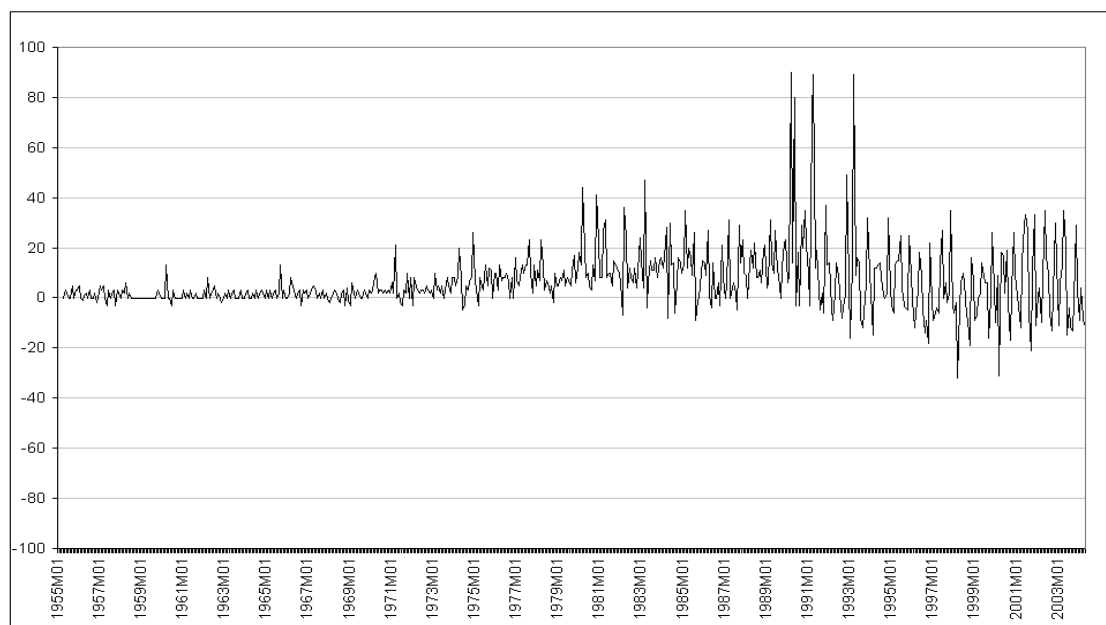


Diagram C.2

2:a differensen av originalserien

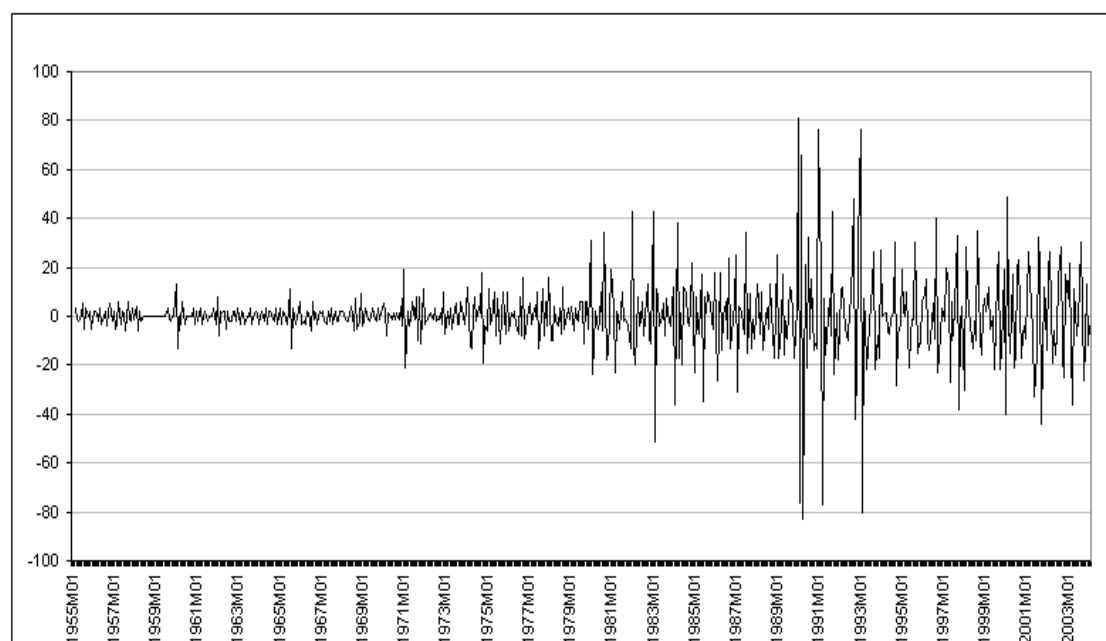


Diagram C.3

1:a differensen och säsongsdifferensen av originalserien

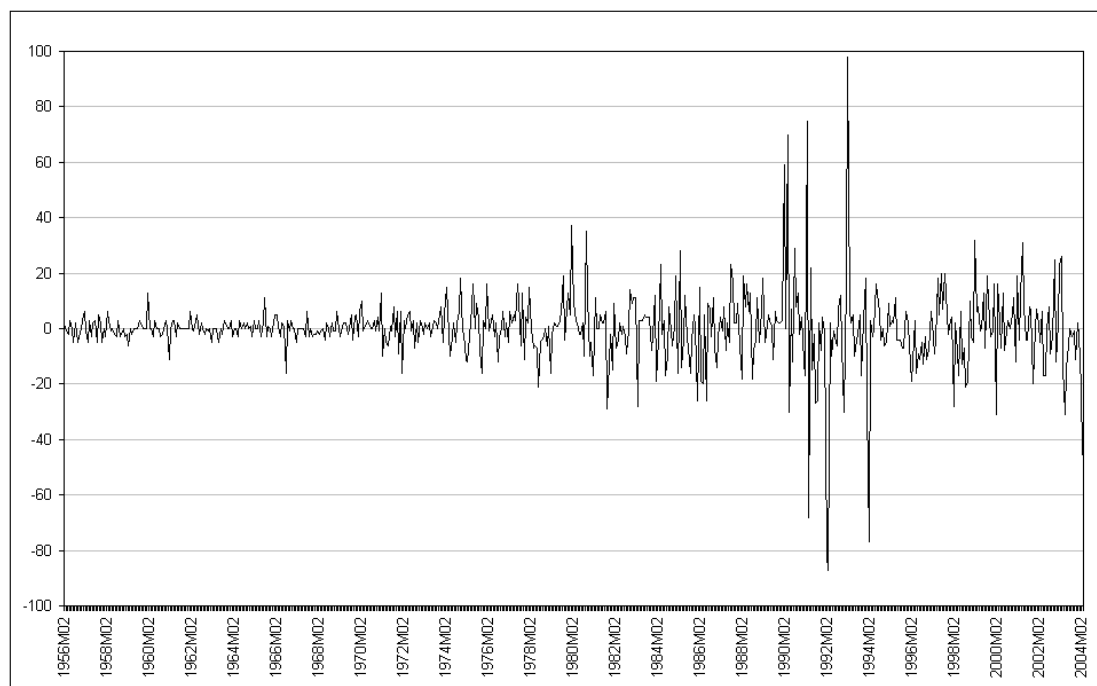


Diagram C.4

2:a differensen och säsongsdifferensen av originalserien

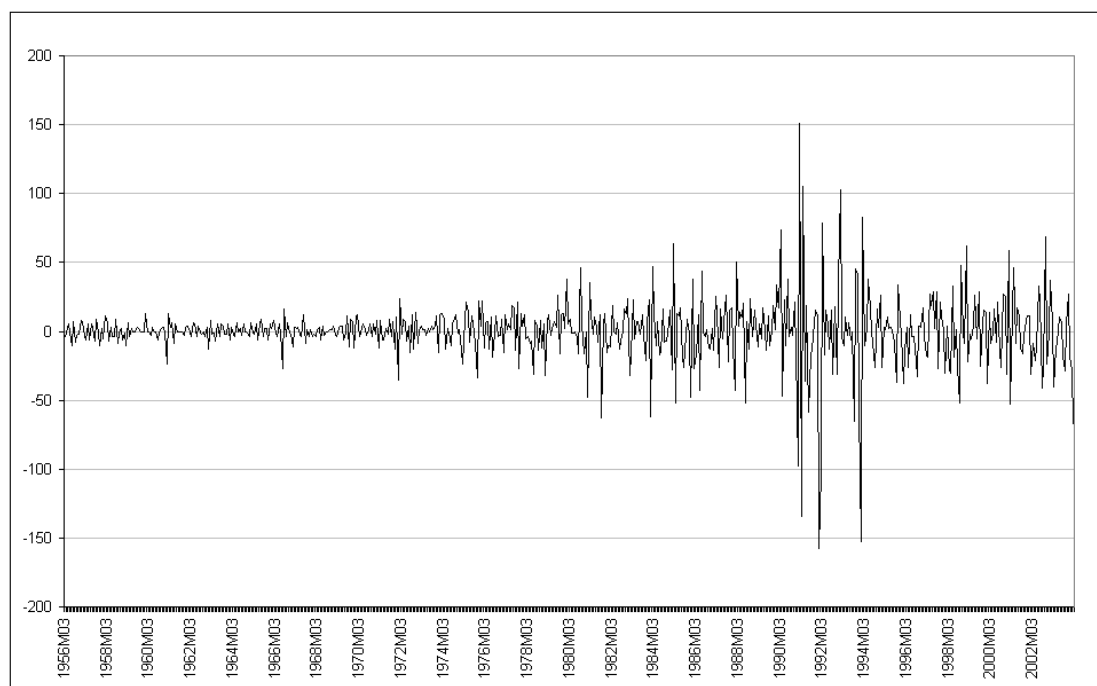


Diagram C.5
1:a differensen av originalserien, logaritmerad

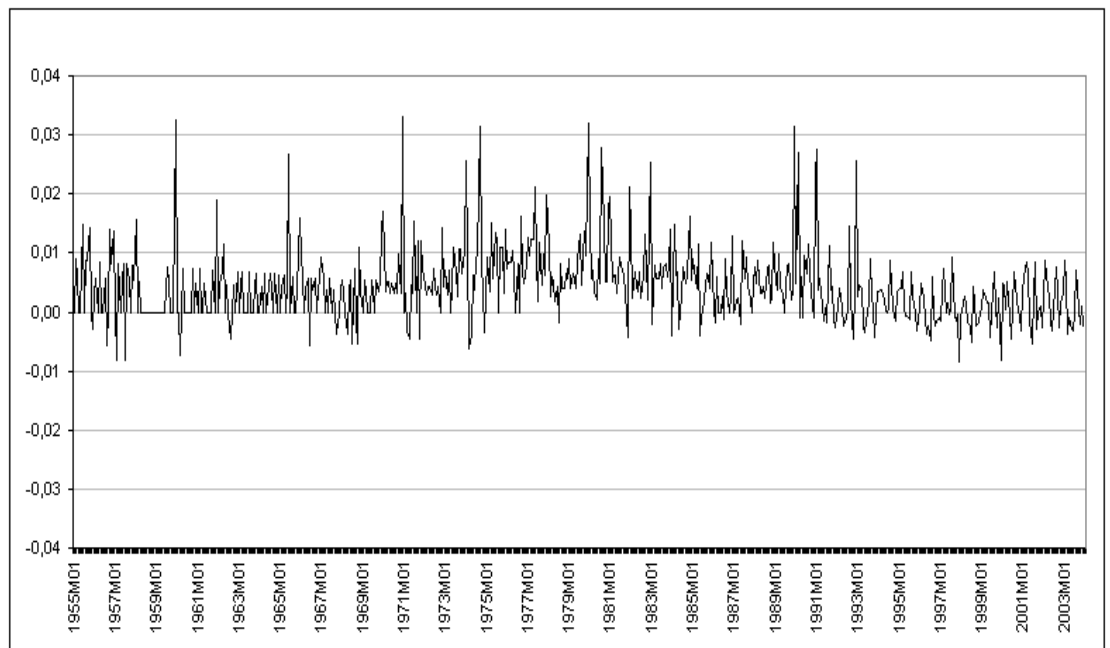


Diagram C.6
2:a differensen av originalserien, logaritmerad

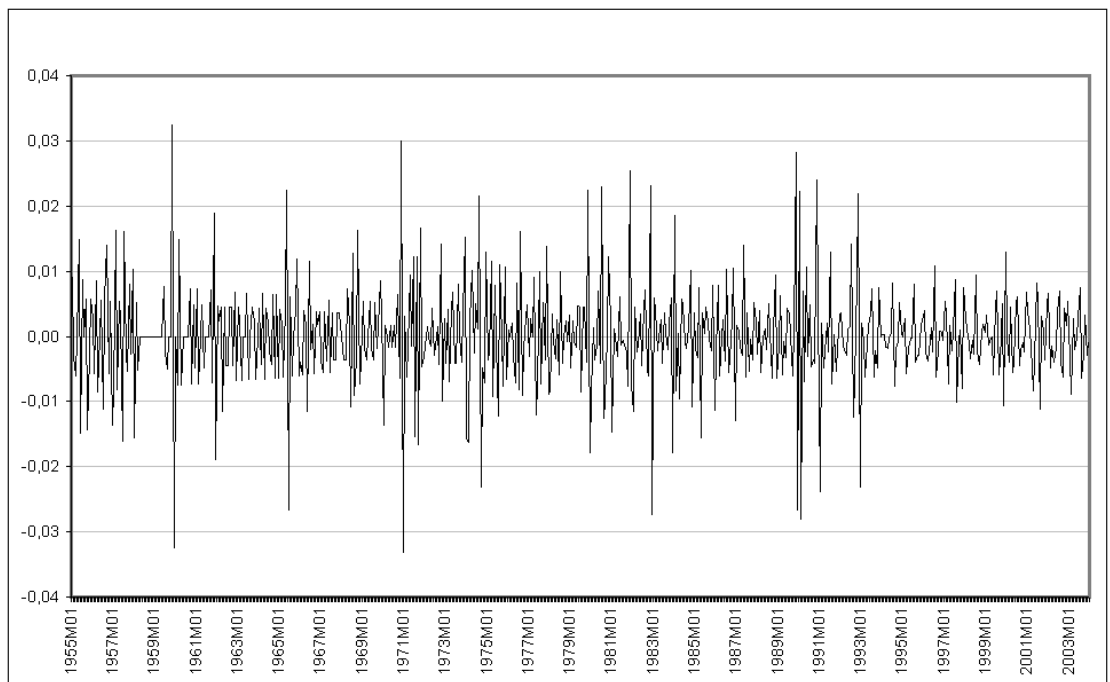


Diagram C.7

1:a differensen och säsongsdifferensen, logaritmerad

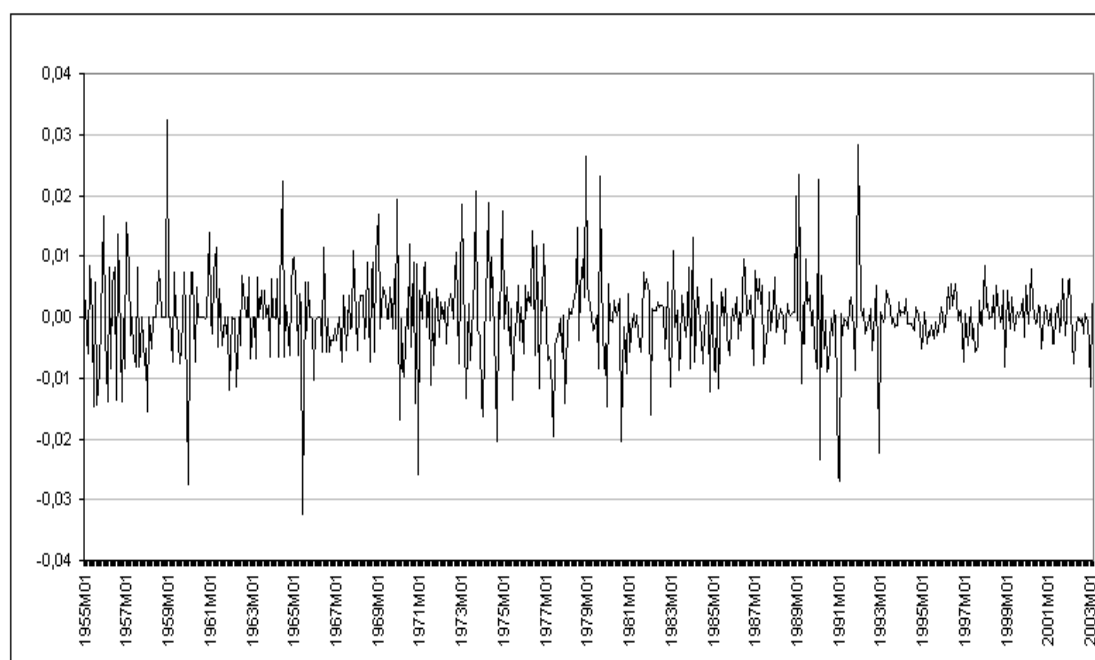


Diagram C.8

2:a differensen och säsongsdifferensen, logaritmerad

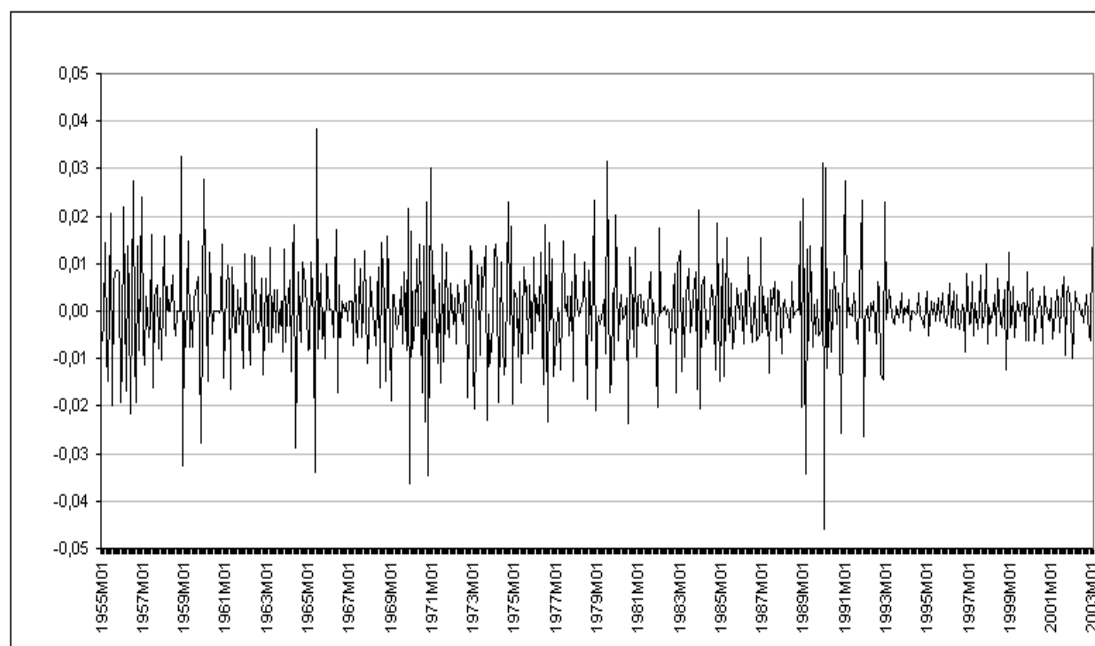


Diagram C.9

1:a differensen av originalserien, Box Cox-transformerad, $\lambda=0,28$

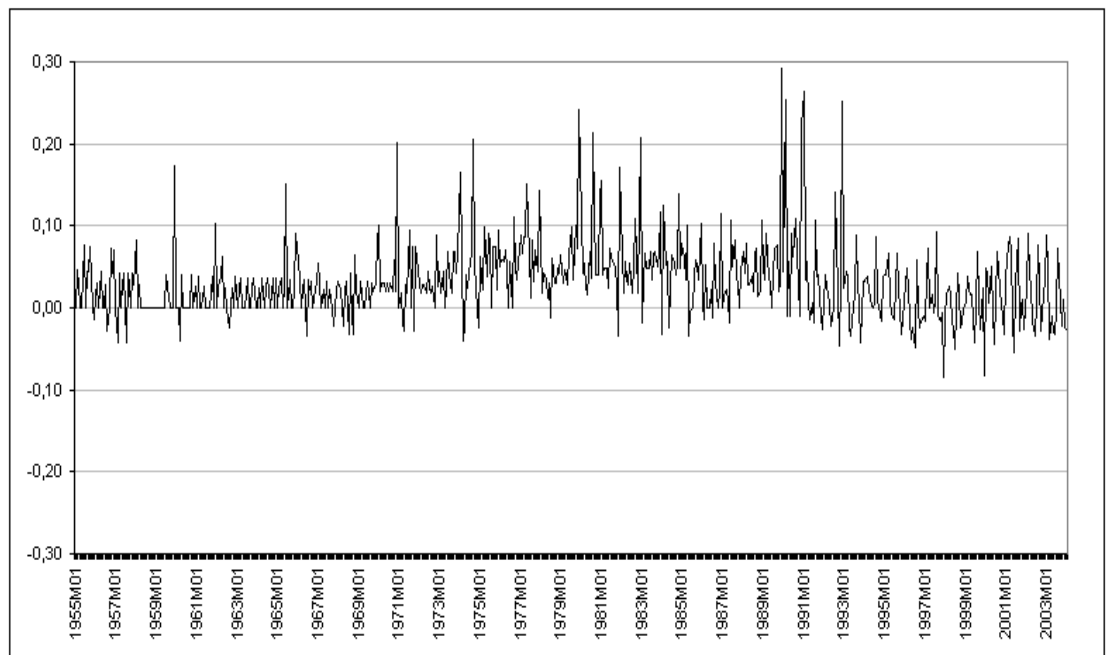


Diagram C.10

2:a differensen av originalserien, Box Cox-transformerad, $\lambda=0,28$

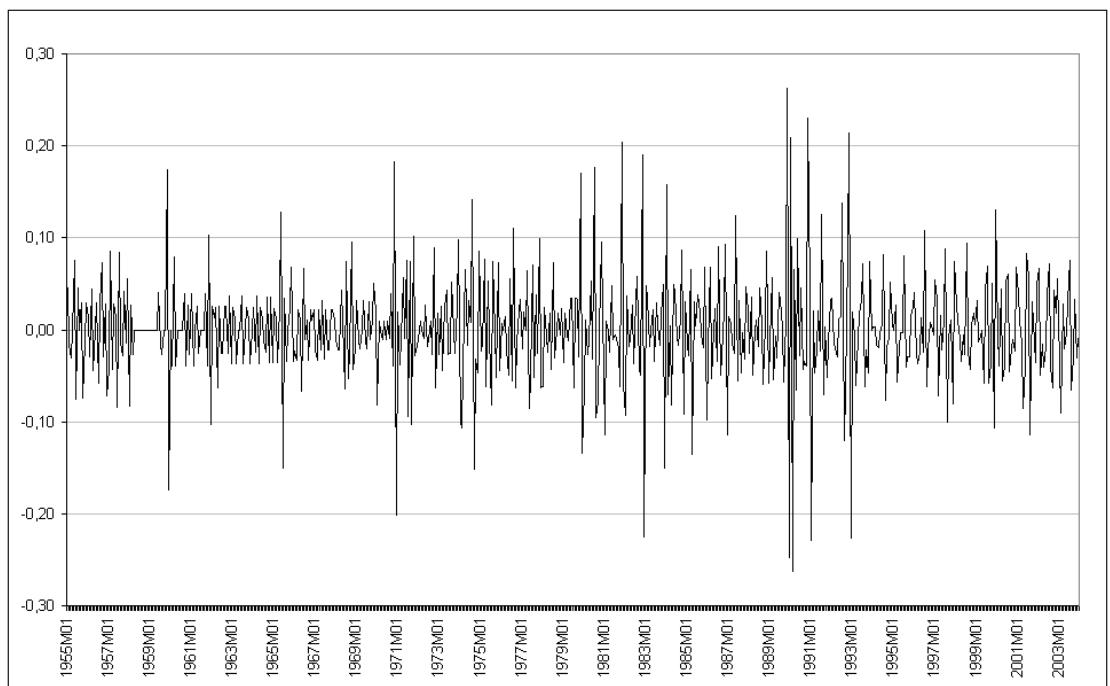


Diagram C.11

1:a differensen och säsongsdifferensen, Box Cox-transformerad, $\lambda=0,28$.

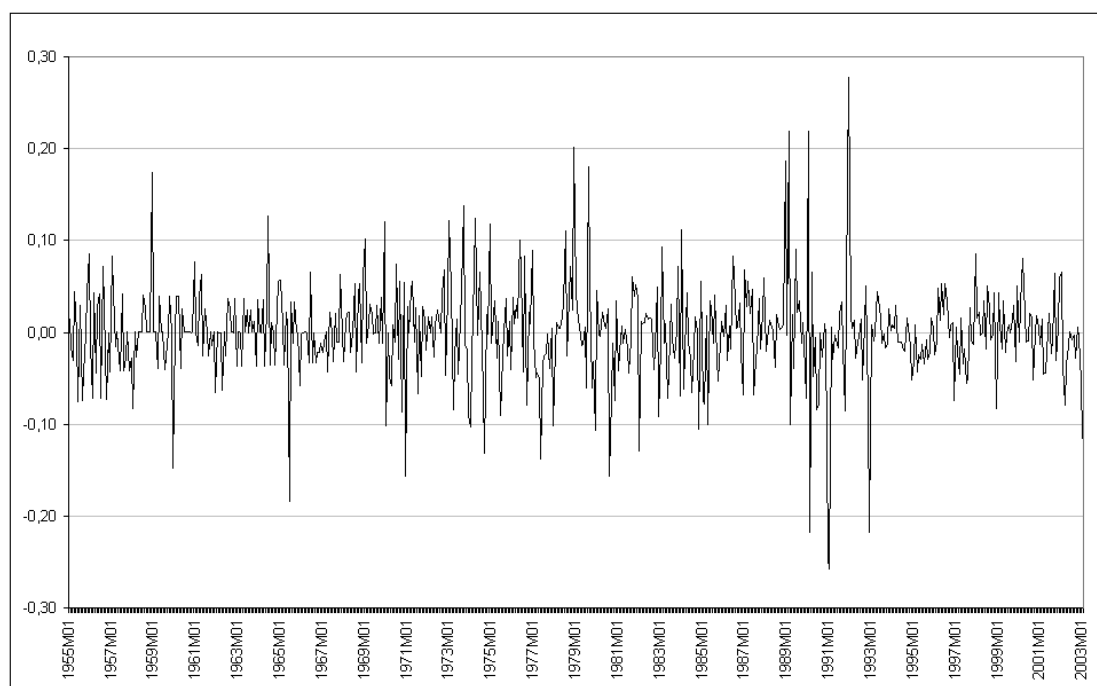
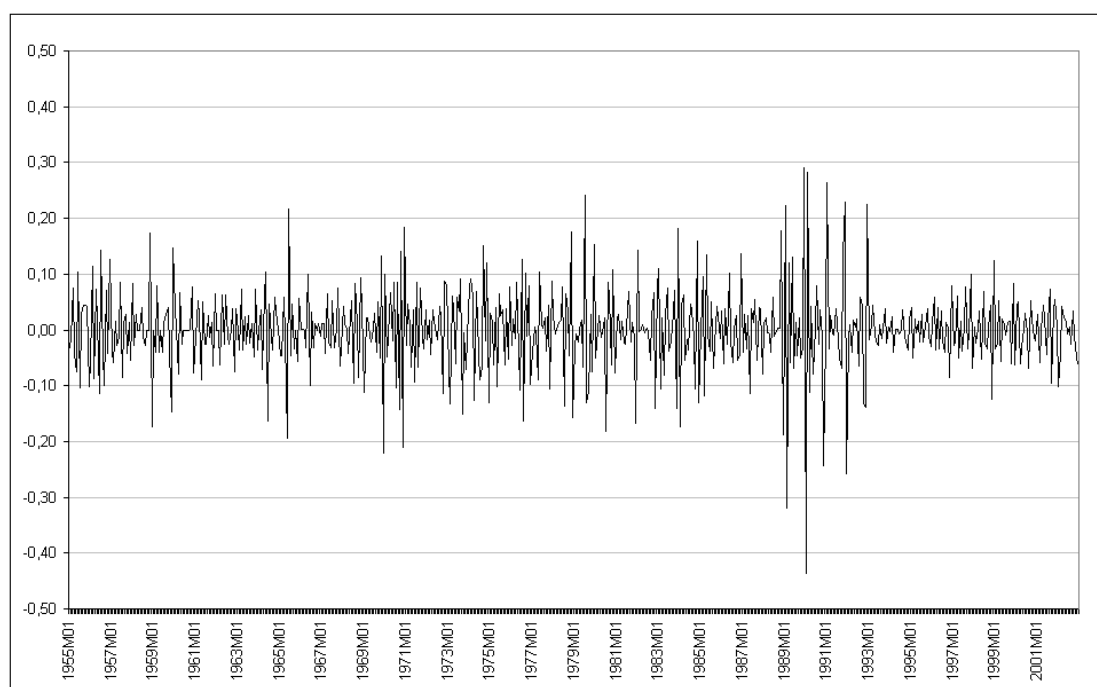


Diagram C.12

2: a differensen och säsongsdifferensen, Box Cox-transformerad, $\lambda=0,2$



Appendix D

I Appendix D redovisas Anderson Darlings normalitetstest för de fem modellerna.

Diagram D.1

Anderson Darlings normalitetstest, Modell 1

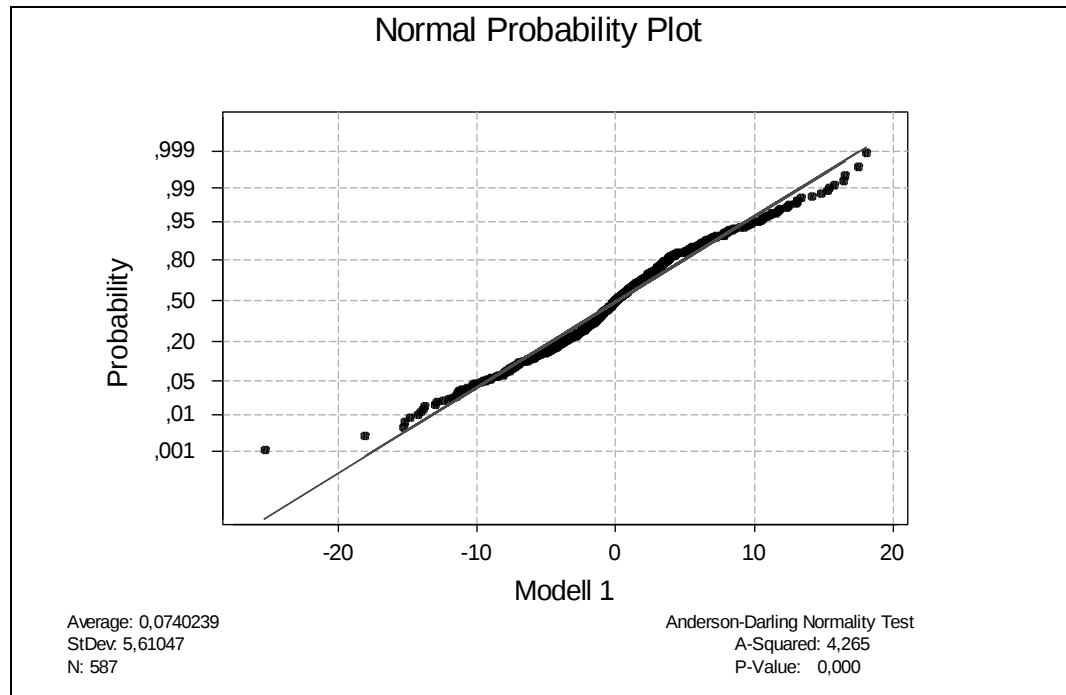


Diagram D.2

Anderson Darlings normalitetstest, Modell 2

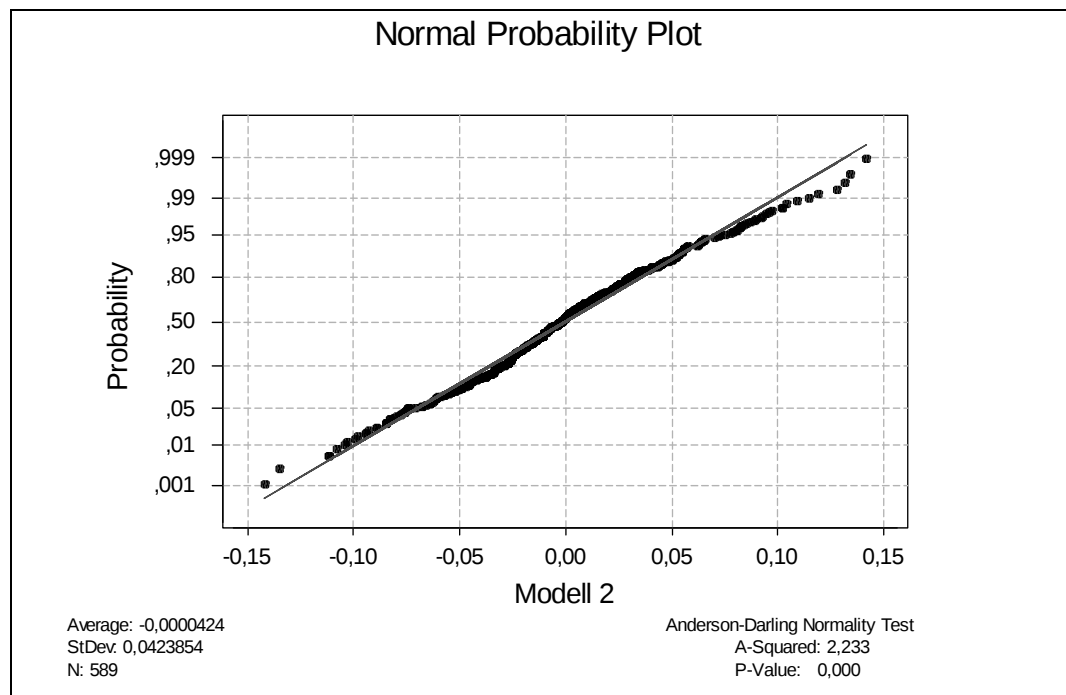


Diagram D.3
Anderson Darlings normalitetstest, Modell 3

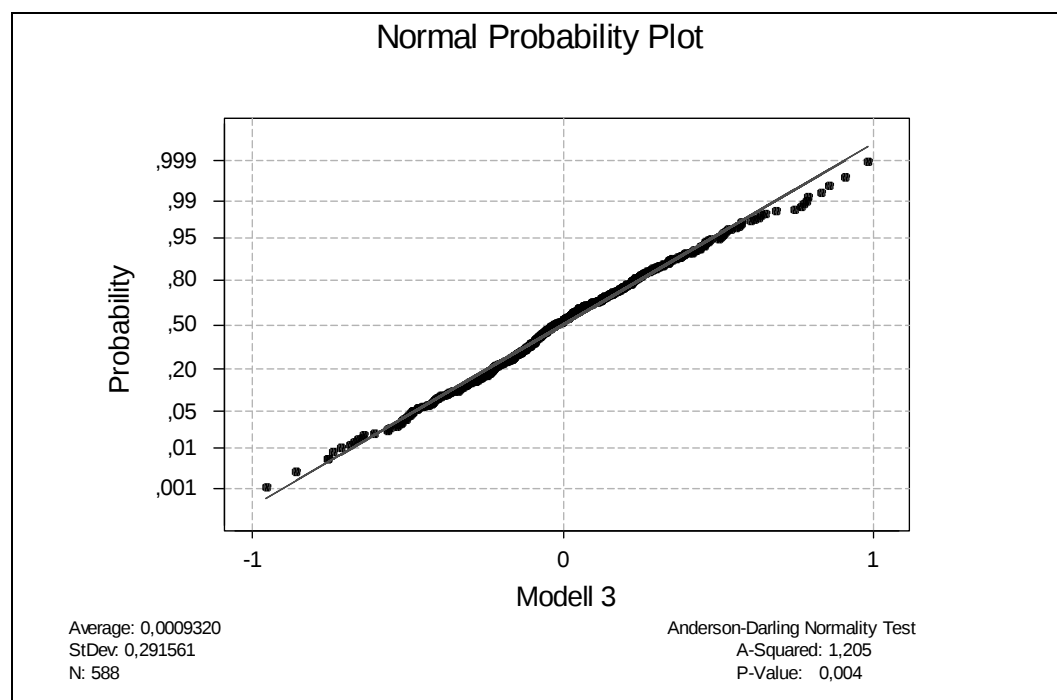


Diagram D.4
Anderson Darlings normalitetstest, Modell 4

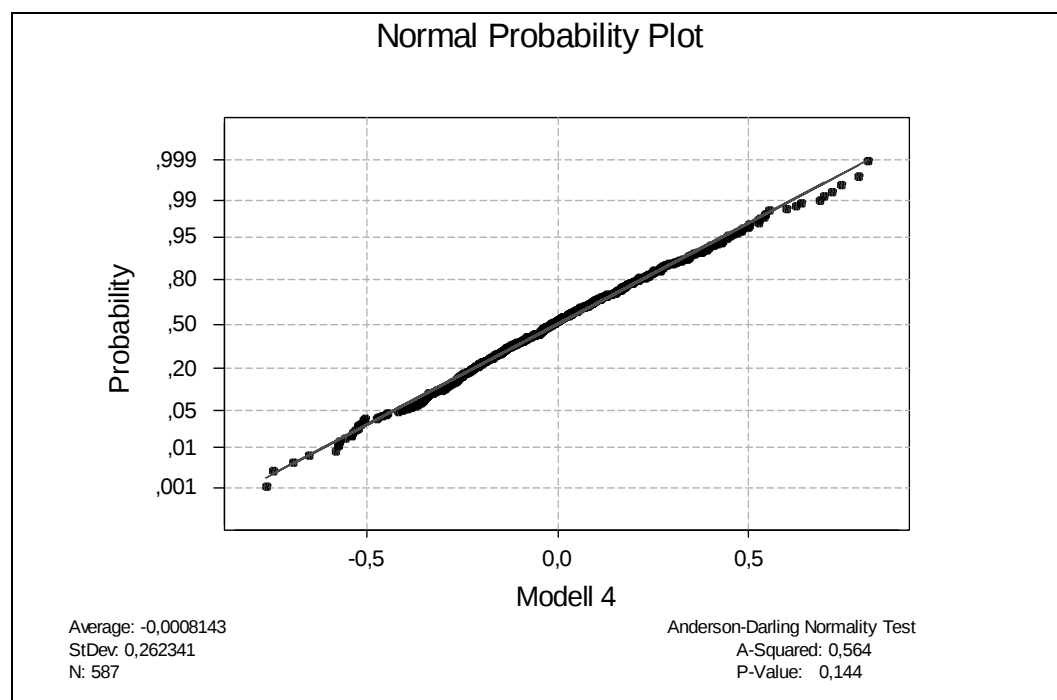
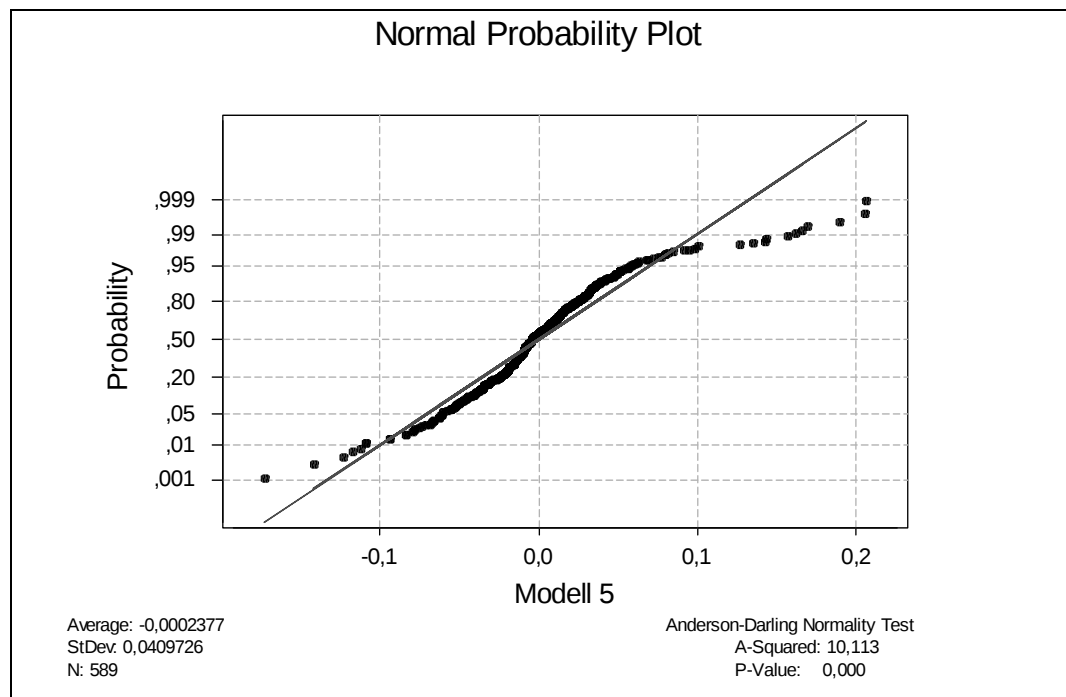


Diagram D.5
Anderson Darlings normalitetstest, Modell 5



Appendix E

I Appendix E redovisas de olika modeller som har testats i TRAMO på originalserien, den logaritmerade serien samt den Box Cox-transformerade serien. Tabellerna inkluderar modellspecifikationer samt en summering av modellernas diagnostik.

Tabell E.1

Modellspecifikationer och diagnostik, originalserie, modell 1 – 12

TRAMO						
ARIMA-modell:	(3,1,1)(0,1,1)	(2,1,1)(0,1,1)	(1,1,1)(0,1,1)	(0,1,1)(0,1,1)	(3,1,1)(0,1,0)	(3,1,1)(1,1,0)
Specifikationen ändras till:	(Automatiska)	(1,2,1)(0,1,1)				
Modell:	1	2	3	4	5	6
1. BIC	3,807	4,163	4,183	3,989	4,221	3,988
2. Normalitet	30,7	215,0	216,0	108,0	43,3	26,3
3. Ljung-Box Q (24 autok)	27,03	29,03	55,20	41,50	74,94	33,71
4. Ljung-Box Q ² (24 autok)	171,98	128,10	126,70	137,10	144,40	112,60
5. Parametervärden (t-värden)	-0,976 (-10,0)	-0,013 (-0,3)	0,927 (19,0)	0,162 (3,9)	0,051 (-100,0)	-0,133 (-100,0)
	0,196 (3,2)	-0,949 (-68,0)	0,961 (26,0)	-0,620 (-18,0)	-0,210 (-5,1)	-0,048 (-1,1)
	-0,126 (-2,6)	-0,718 (-23,0)	-0,643 (-18,0)		-0,159 (-100,0)	-0,138 (-100,0)
	-0,761 (-8,5)				0,051 (-100,0)	0,464 (12,0)
	-0,636 (-18,0)					-0,040 (-100,0)
6. Antal outliers	26	7	7	14	20	19
TRAMO						
ARIMA-modell:	(2,1,1)(0,1,0)	(2,1,1)(1,1,0)	(1,1,1)(0,1,0)	(1,1,1)(1,1,0)	(0,1,2)(0,1,1)	(0,1,3)(1,1,1)
Specifikationen ändras till:						
Modell:	7	8	9	10	11	12
1. BIC	4,255	4,01764	4,284	4,021	4,158	4,081
2. Normalitet	68,1	41,5	48,4	46	168	103
3. Ljung-Box Q (24 autok)	91,92	37,41	97,71	45,43	44,01	36,61
4. Ljung-Box Q ² (24 autok)	117,4	150,1	159,2	157,6	154,6	173,9
5. Parametervärden (t-värden)	0,040 (14,0)	-0,758 (-100,0)	-0,254 (-100,0)	-0,599 (-18,0)	0,038 (0,8)	-0,038 (-0,5)
	-0,857 (-36,0)	-0,057 (-14,0)	-0,144 (-100,0)	0,457 (-100,0)	0,088 (2,1)	0,206 (4,8)
		0,506 (-100,0)		-0,479 (-100,0)	-0,650 (-20,0)	0,102 (2,4)
		-0,697 (-26,0)				0,106 (2,6)
						-0,650 (-12,0)
6. Antal outliers	21	17	18	17	8	13

Tabell E.2
Modellspecifikationer och diagnostik, originalserie, modell 13 – 30

TRAMO						
ARIMA-modell:	(0,2,3)(1,1,1)	(2,2,1)(0,1,1)	(3,2,1)(1,1,1)	(2,1,2)(1,1,0)	(3,1,0)(1,1,1)	(3,1,1)(1,1,1)
Specifikationen ändras till:						
Modell:	13	14	15	16	17	18
1. BIC	3,98	4,11	4,136	4,027	3,924	3,97
2. Normalitet	131	232	226	43,4	41,3	95,3
3. Ljung-Box Q (24 autok)	25,76	36,41	38,52	50,19	31,92	30,11
4. Ljung-Box Q² (24 autok)	115,3	172,9	164,2	141,3	159,9	120,9
5. Parametervärden (t-värden)	0,009 (0,14)	-0,213 (-4,3)	-0,113 (-2,1)	-1,234 (100,0)	-0,165 (-3,8)	0,0212 (0,1)
	-0,826 (-19,0)	0,0601 (1,3)	0,0891 (1,8)	0,453 (100,0)	-0,014 (-0,3)	-0,049 (-0,8)
	-0,144 (-2,6)	-0,889 (-34,0)	-0,116 (-2,3)	0,468 (100,0)	-0,145 (-3,4)	-0,147 (-3,5)
	0,035 (0,8)	-0,711 (-22,0)	0,048 (0,8)	-1,165 (100,0)	-0,044 (-0,6)	0,023 (0,3)
	-0,667 (-13,0)		-0,891 (-29,0)	0,458 (100,0)	-0,640 (-11,0)	0,188 (0,6)
			-0,681 (-15,0)			-0,610 (-11,0)
6. Antal outliers	16	15	12	18	19	17
TRAMO						
ARIMA-modell:	(3,0,1)(0,1,1)	(3,1,0)(0,1,1)	(3,1,0)(1,1,0)	(2,1,0)(1,1,0)	(2,1,0)(1,1,1)	(2,1,1)(1,1,1)
Specifikationen ändras till:						(1,2,1)(1,1,1)
Modell:	19	20	21	22	23	24
1. BIC	4,287	4,151	4,014	4,025	3,992	3,985
2. Normalitet	365	172	34	42,9	104	145
3. Ljung-Box Q (24 autok)	60,54	30,65	34,45	46,33	36,87	25,37
4. Ljung-Box Q² (24 autok)	154,4	153,3	153,8	162,8	147,8	129,5
5. Parametervärden (t-värden)	-0,042 (0,0)	-0,046 (-1,1)	-0,069 (100,0)	-0,080 (-1,9)	-0,167 (-3,9)	-0,122 (-2,7)
	-0,074 (0,0)	-0,078 (-1,9)	-0,079 (-1,9)	-0,097 (-2,2)	-0,0497 (-1,2)	-0,014 (-0,2)
	0,253 (0,0)	-0,126 (-3,0)	-0,147 (100,0)	0,460 (100,0)	-0,007 (-0,1)	-0,957 (-75,0)
	-0,687 (-21,0)	-0,657 (-20,0)	0,494 (13,0)		-0,625 (-11,)	-0,670 (-13,0)
6. Antal outliers	8	8	17	17	15	15
TRAMO						
ARIMA-modell:	(1,1,0)(1,1,0)	(1,1,0)(1,1,1)	(0,2,2)(0,1,1)	(0,2,1)(0,1,1)	(2,1,0)(0,1,1)	(1,1,0)(0,1,1)
Specifikationen ändras till:						
Modell:	25	26	27	28	29	30
1. BIC	3,96	3,988	4,33	4,255	4,178	3,998
2. Normalitet	34,6	106	291	424	217	94,1
3. Ljung-Box Q (24 autok)	54,408	42,59	26,87	45,18	46,42	38,78
4. Ljung-Box Q² (24 autok)	98,64	145,5	252,4	114,6	122,9	155,1
5. Parametervärden (t-värden)	-0,129 (100,0)	-0,180 (-4,2)	-0,634 (-16,0)	-0,888 (-44,0)	-0,072 (-1,7)	-0,171 (-4,0)
	0,415 (100,0)	-0,046 (-0,6)	-0,285 (-7,1)	-0,744 (-24,0)	-0,079 (-1,9)	-0,622 (-18,0)
		-0,645 (-11,0)	-0,746 (-25,0)		-0,664 (-21,0)	
6. Antal outliers	20	15	10	12	7	14

Tabell E.3

Modellspecifikationer och diagnostik, logaritmerad serie, modell 1 – 18

TRAMO						
ARIMA-modell:	(0,2,1)(0,1,1)	(0,2,1)(0,0,0)	(0,2,2)(0,1,1)	(0,2,1)(1,1,1)	(0,2,1)(0,1,0)	(0,2,1)(0,0,1)
Modell:	1	2	3	4	5	6
1. BIC	-10,71	-10,67	-10,54	-10,62	-10,30	-10,62
2. Normalitet	25,60	24,50	173,00	134,00	22,50	26,17
3. Ljung-Box Q (24 autok)	19,80	77,29	17,64	21,46	129,2	92,98
4. Ljung-Box Q ² (24 autok)	25,12	16,40	37,33	30,94	39,17	30,52
5. Parametervärden (t-värden)	-0,92 (-56,2)	-0,94 (-70,0)	-0,96 (-23,0)	-0,11 (-2,0)	-1,0 (100,0)	-0,85 (-100,0)
	-0,84 (-36,7)		-0,03 (-0,6)	-0,94 (-63,0)		0,14 (-100,0)
			-0,84 (-33,0)	-0,85 (-26,0)		
6. Antal outliers	16	14	6	12	13	12
TRAMO						
ARIMA-modell:	(3,2,0)(0,1,0)	(3,2,3)(0,1,0)	(0,2,1)(1,0,0)	(0,2,2)(1,0,0)	(0,2,2)(1,1,0)	(1,2,0)(0,1,0)
Modell:	7	8	9	10	11	12
1. BIC	-10,26	-10,16	-10,69	-10,68	-10,44	-9,94
2. Normalitet	17,10	32,30	28,40	27,70	57,30	15,40
3. Ljung-Box Q (24 autok)	115,40	119,60	49,64	48,39	95,95	169,80
4. Ljung-Box Q ² (24 autok)	48,61	50,84	32,41	31,75	49,40	71,62
5. Parametervärden (t-värden)	0,90 (-100,0)	0,79 (3,1)	-0,25 (-6,30)	-0,27 (-6,7)	0,53 (15,0)	0,57 (17,0)
	0,65 (-100,0)	0,35 (1,3)	-0,95 (-77,00)	-1,00 (-24,0)	-0,94 (-22,0)	
	0,23 (-100,0)	-0,13 (-2,9)		0,05 (1,1)	0,05 (1,3)	
		-0,21 (-0,8)				
		-0,34 (-1,4)				
		-0,42 (-1,7)				
6. Antal outliers	19	10	12	12	13	12
TRAMO						
ARIMA-modell:	(0,1,1)(0,1,1)	(1,1,1)(0,1,1)	(1,1,1)(1,1,1)	(0,1,2)(0,1,1)	(0,1,3)(0,1,1)	(1,1,3)(0,1,1)
Modell:	13	14	15	16	17	18
1. BIC	-10,53	-10,51	-10,66	-10,48	-10,48	-10,47
2. Normalitet	232,40	218,70	32,68	246,00	260,00	261,00
3. Ljung-Box Q (24 autok)	49,41	35,35	61,11	45,00	37,32	33,05
4. Ljung-Box Q ² (24 autok)	55,56	53,41	45,98	51,06	50,45	50,44
5. Parametervärden (t-värden)	0,05 (-1,1)	-0,54 (-1,3)	0,14 (0,1)	0,03 (0,7)	0,03 (0,7)	-0,44 (-0,8)
	-0,74 (-23,0)	-0,47 (-1,1)	-0,03 (-0,5)	0,10 (2,4)	0,09 (2,2)	-0,41 (-0,7)
		-0,75 (-25,0)	0,10 (0,1)	-0,76 (-24,0)	0,07 (1,8)	0,08 (1,9)
			-0,73 (-18,1)		-0,77 (-27,0)	0,03 (0,5)
						-0,77 (-27,0)
6. Antal outliers	6	6	13	5	5	5

Tabell E.4
Modellspecifikationer och diagnostik, logaritmerad serie, modell 19 – 36

TRAMO						
ARIMA-modell:	(0,1,1)(1,1,1)	(0,1,1)(1,1,0)	(0,1,3)(1,1,1)	(0,1,0)(1,0,1)	(1,1,0)(1,0,1)	(0,1,1)(1,0,1)
Modell:	19	20	21	22	23	24
1. BIC	-10,67	-10,53	-10,62	-10,73	-10,71	-10,71
2. Normalitet	32,40	47,00	67,60	39,70	39,60	39,30
3. Ljung-Box Q (24 autok)	60,06	79,60	50,91	72,64	75,36	72,55
4. Ljung-Box Q ² (24 autok)	46,20	45,01	32,11	75,11	73,70	74,32
5. Parametervärden (t-värden)	-0,04 (-0,6)	0,47 (100,0)	-0,05 (-1,0)	-0,90 (-27,0)	0,02 (0,6)	-0,90 (-34,0)
	-0,03 (-0,7)	-0,03 (-100,0)	-0,01 (-0,2)	-0,66 (-12,0)	-0,90 (-32,0)	-0,02 (-0,5)
	-0,74 (-16,0)		0,08 (2,0)		-0,65 (-12,0)	-0,64 (-13,0)
			0,13 (3,2)			
			-0,77 (-21,0)			
6. Antal outliers	13	14	11	13	13	13
TRAMO						
ARIMA-modell:	(0,1,2)(1,0,1)	(2,1,1)(0,1,1)	(2,2,0)(0,1,0)	(3,2,1)(0,1,1)	(3,2,0)(0,1,1)	(1,2,1)(0,1,1)
Modell:	25	26	27	28	29	30
1. BIC	-10,54	-10,52	-10,23	-10,55	-10,45	-10,60
2. Normalitet	283,00	154,00	8,04	144,00	62,10	180,00
3. Ljung-Box Q (24 autok)	45,66	51,19	116,20	15,25	58,23	10,99
4. Ljung-Box Q ² (24 autok)	52,74	60,42	40,58	45,32	36,72	20,89
5. Parametervärden (t-värden)	-0,90 (-29,0)	-0,18 (-0,1)	0,79 (-100,0)	0,02 (0,5)	0,84 (21,0)	-0,02 (-0,5)
	0,03 (0,6)	0,01 (0,0)	0,48 (-100,0)	-0,01 (-0,2)	0,57 (12,0)	-0,95 (-67,0)
	0,10 (2,4)	-0,12 (0,0)		-0,08 (-1,7)	0,21 (5,1)	-0,85 (-32,0)
	-0,68 (-13,0)	-0,76 (-23,0)		-0,95 (-61,0)	-0,82 (-30,0)	
				-0,82 (-29,0)		
6. Antal outliers	6	7	19	7	9	10
TRAMO						
ARIMA-modell:	(2,2,1)(0,1,1)	(0,2,0)(0,1,1)	(2,2,2)(0,1,1)	(3,2,1)(1,0,0)	(2,2,1)(0,1,0)	(3,2,1)(0,1,0)
Modell:	31	32	33	42	35	36
1. BIC	-10,56	-10,01	-10,55	-10,66	-10,28	-10,22
2. Normalitet	143,00	7,83	141,00	31,50	6,34	30,50
3. Ljung-Box Q (24 autok)	17,65	134,90	16,00	50,43	118,30	115,80
4. Ljung-Box Q ² (24 autok)	44,00	118,20	47,14	29,68	50,62	41,16
5. Parametervärden (t-värden)	0,04 (0,9)	-0,81 (-33,0)	0,44 (1,0)	0,04 (1,0)	0,30 (9,8)	-0,04 (-0,9)
	0,01 (0,1)		0,06 (1,4)	0,00 (0,0)	0,13 (7,5)	-0,15 (-3,2)
	-0,93 (-53,0)		-0,54 (-1,3)	-0,04 (-1,0)	-0,77 (-100,0)	-0,22 (-5,1)
	-0,83 (-29,0)		-0,36 (-0,9)	-0,27 (-6,5)		-0,95 (-36,0)
			-0,83 (-30,0)	-0,96 (-82,0)		
6. Antal outliers	7	10	7	12	16	13

Tabell E.5

Modellspecifikationer och diagnostik, logaritmerad serie, modell 37 – 42

TRAMO						
ARIMA-modell:	(2,2,1)(0,0,1)	(2,1,1)(1,1,0)	(1,2,1)(0,1,0)	(2,2,1)(0,0,0)	(2,2,2)(0,0,0)	(2,2,0)(0,1,1)
Modell:	37	38	39	40	41	42
1. BIC	-10,65	-10,52	-10,31	-10,64	-10,27	-10,56
2. Normalitet	29,70	38,80	26,60	25,20	916,00	143,00
3. Ljung-Box Q (24 autok)	59,41	72,93	116,10	84,21	77,56	17,65
4. Ljung-Box Q ² (24 autok)	27,44	73,17	46,11	34,20	22,83	44,00
5. Parametervärden (t-värden)	0,04 (1,0)	-0,32 (-100,0)	0,10 (2,3)	0,01 (0,2)	0,12 (0,0)	0,04 (0,9)
	0,02 (0,4)	-0,09 (-2,4)	-0,99 (100,0)	0,04 (1,0)	0,00 (0,0)	0,01 (0,1)
	-0,95 (-83,0)	0,49 (14,0)		-0,94 (-80,0)	-0,82 (-0,2)	-0,93 (-53,0)
	0,20 (-4,8)	-0,33 (100,0)			-0,13 (0,0)	-0,83 (-29,0)
6. Antal outliers	12	14	14	13	13	7

Tabell E.6

Modellspecifikationer och diagnostik, Box Cox-transf. serie, modell 1 – 12

TRAMO	Automatisk					
ARIMA-modell:	(1,2,1)(0,1,1)	(0,2,1)(0,1,1)	(0,2,1)(0,0,1)	(3,2,0)(0,1,0)	(2,2,1)(0,1,1)	(0,2,1)(1,0,0)
Modell:	1	2	3	4	5	6
1. BIC	-6,83	-6,42	-6,66	-6,21	-6,67	-6,68
2. Normalitet	8,19	456,00	7,53	9,04	90,40	12,10
3. Ljung-Box Q (24 autok)	19,63	13,50	97,72	115,40	15,41	49,58
4. Ljung-Box Q ² (24 autok)	20,36	55,58	43,08	60,97	26,12	51,79
5. Parametervärden (t-värden)	0,07 (1,7)	-0,94 (-65,0)	-0,95 (-71,0)	0,90 (100,0)	-0,04 (-0,8)	-0,37 (-9,3)
	-0,93 (-58,7)	-0,82 (-34,0)	0,24 (5,9)	0,57 (19,0)	0,01 (0,31)	-0,96 (-82,0)
	-0,78 (-27,9)			0,14 (100,0)	-0,90 (-37)	
					-0,78 (-28)	
6. Antal outliers	17	2	14	18	13	12
TRAMO						
ARIMA-modell:	(0,2,1)(0,1,0)	(2,2,0)(0,1,0)	(0,2,0)(0,1,1)	(1,2,1)(1,1,0)	(2,2,1)(0,1,0)	(1,2,1)(0,1,0)
Modell:	7	8	9	10	11	12
1. BIC	-6,34	-6,16	-6,06	-6,60	-6,32	-6,36
2. Normalitet	5,10	2,89	17,60	13,80	2,02	0,94
3. Ljung-Box Q (24 autok)	113,40	117,30	118,20	64,24	110,60	113,80
4. Ljung-Box Q ² (24 autok)	59,84	66,83	108,80	46,67	59,42	63,74
5. Parametervärden (t-värden)	-0,99 (100,0)	0,79 (100,0)	-0,76 (-28,0)	0,00 (0,0)	0,22 (6,3)	0,14 (3,2)
		0,46 (100,0)		0,54 (15,0)	0,04 (2,8)	-0,96 (-76,0)
				-0,92 (-53,0)	-0,88 (100,0)	
6. Antal outliers	13	16	11	16	16	15

Tabell E.7
Modellspecifikationer och diagnostik, Box Cox-transf. serie, modell 13 – 18

TRAMO						
ARIMA-modell:	(1,2,1)(1,1,1)	(2,2,1)(0,0,0)	(2,2,1)(0,0,1)	(1,2,2)(0,1,1)	(0,2,2)(0,1,1)	(0,2,2)(0,1,0)
Modell:	13	14	15	16	17	18
1. BIC	-6,71	-6,67	-6,62	-6,78	-6,39	-6,27
2. Normalitet	16,34	90,40	20,60	7,59	411,00	20,90
3. Ljung-Box Q (24 autok)	27,50	15,41	81,62	18,81	18,50	111,20
4. Ljung-Box Q ² (24 autok)	20,89	26,12	27,60	33,05	53,77	67,49
5. Parametervärden (t-värden)	-0,02 (-0,5)	-0,04 (-0,8)	-0,03 (-0,8)	0,80 (0,0)	-0,92 (-22,0)	-1,01 (-25,0)
	0,03 (0,5)	0,01 (0,3)	0,07 (1,5)	-0,14 (0,0)	-0,02 (-0,5)	0,03 (0,8)
	-0,91 (-50,0)	-0,90 (-37,0)	-0,95 (-62,0)	-0,75 (0,0)	-0,83 (-31,0)	
	-0,77 (-20,0)	-0,78 (-28,0)	0,24 (5,8)	-0,79 (-28,0)		
6. Antal outliers	14	13	12	16	2	11

- 2004:01 Hjälpverksamhet. Avrapportering av projektet Systematisk hantering av hjälpverksamhet
- 2004:02 Report from the Swedish Task Force on Time Series Analysis
- 2004:03 Minskad detaljeringsgrad i Sveriges officiella utrikeshandelsstatistik
- 2004:04 Finansiellt sparande i den svenska ekonomin. Utredning av skillnaderna i finansiellt sparande Nationalräkenskaper, NR – Finansräkenskaper, FiR
Bakgrund – jämförelser – analys
- 2004:05 Designutredning för KPI: Effektiv allokering av urvalet för prismätningarna i butiker och tjänsteställen. Examensarbete inom Matematisk statistik utfört på Statistiska centralbyrån i Stockholm
- 2004:06 Tidsserieanalys av svenska BNP-revideringar 1980–1999
- 2004:07 Labor Quality and Productivity: Does Talent Make Capital Dance?
- 2004:08 Slutrapport från projektet Uppsnabbning av den ekonomiska korttidsstatistiken
- 2004:09 Bilagor till slutrapporten från projektet Uppsnabbning av den ekonomiska korttidsstatistiken
- 2004:10 Förbättring av bortfallsprocessen i Intrastat
- 2004:11 PLÖS. Samordning av produktion, löner och sysselsättning
- 2004:12 Net lending in the Swedish economy. Analysis of differences in net lending National accounts (NA) – Financial accounts (FA). Background – comparisons - analysis
- 2004:13 Testing for Normality and ARCH. An Empirical Study of Swedish GDP Revisions 1980–1999
- 2004:14 Combining leading indicators and a flash estimate
- 2004:15 Comparing welfare of nations
- 2004:16 ES-avdelningens utvecklingsplan 2004

ISSN 1650-9447

Statistikpublikationer kan beställas från SCB, Publikationstjänsten, 701 89 ÖREBRO, e-post: publ@scb.se, telefon: 019-17 68 00, fax: 019-17 64 44. De kan också köpas genom bokhandeln eller direkt hos SCB, Karlavägen 100 i Stockholm. Aktuell publicering redovisas på vår webbplats (www.scb.se). Ytterligare hjälp ges av Bibliotek och information, e-post: information@scb.se, telefon: 08-506 948 01, fax: 08-506 948 99.

www.scb.se