

Summakonsistent säsongrensning

Presentation av projektarbete på SCB
av
Suad Elezović

Statistiska institutionen, Stockholms universitet

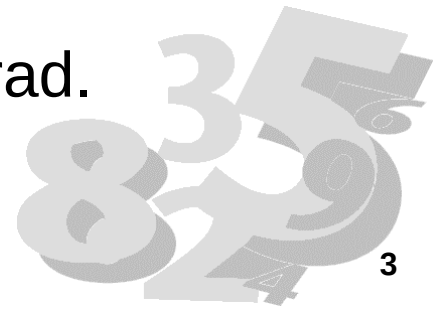
14 Oktober 2009

Säsongrensning : Allmänt

- **Motivering:**
 - ekonomiska tidserier oftast uppvisar periodiska variationer med konstant periodlängd som beror på de växlingar i ekonomiska aktiviteter som har att göra med årstider.
- **Målsättning:**
 - I. Beskrivning av säsongvariation (modell med säsongfaktorn) för att göra prognoser.
 - II. Eliminering av säsongvariation för att uppnå jämförbarhet mellan två på varann följande perioder.

Säsongrensning forts.

- Traditionell säsongrensning:
 - Dekomponering av tidserien i komponenter: **Trend-Cykel** (TC), **Säsong** (S) och **Irreguljär** (I).
 - Additiv eller multiplikativ modell.
- Modern säsongrensning:
 - Utökat antal komponenter: **Trend-Cykel** (TC), **Säsong** (S), **Kalender** (K), **Extremvärden** (E) och **Irreguljär** (I) komponent.
 - Modellbaserad och icke-modellbaserad.



Verktyg och metoder

- Icke-modellbaserade: **X11** (Census Bureau, USA) {och **X11-Arima**}
- Modellbaserade: **Tramo/Seats** (Maravall & Gomez).
- Delvis modellbaserade: **X12-Arima**.
- Ett antal andra metoder används i mindre utsträckning: **BV4, SABL, Dainties, STAMP**, mm.



Säsongrensning på SCB

- Modellbaserad metod: Tramo/Seats (i första hand) och X12-Arima.
- Programvaror: Demetra, SAS, TSW, egna applikationer.
- Stort antal tidserier säsongrensas (ca 700) och det blir ännu mer.



Om Tramo/Seats

- Två stegs metod:
 - Steg 1 (**Tramo**) Regressionsmodell med regressionsvariabler

$$Y_t = X_t' \beta + \varepsilon_t$$

- ...där slumpterm följer en generell ARIMA-process
- ...och vektor av regressionsvariabler består av s.k. deterministiska effekter: **kalendereffekter** och **extremvärdeeffekter**.



Steg 2 (Seats)

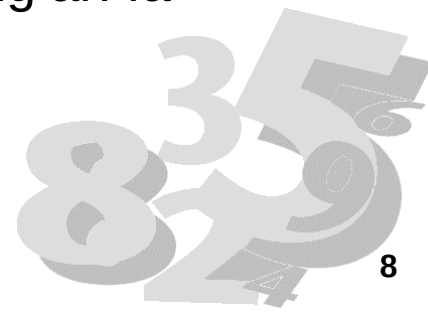
- Den "lineariserade" serien från TRAMO dekomponeras

$$\varepsilon_t = \varepsilon_{1t} + \varepsilon_{2t} + \varepsilon_{3t} + \varepsilon_{4t}$$

- där komponenterna är trend, cykel, säsong och slump. Varje komponent följer en ARIMA modell.
- dekomponering sker m.h.a spektralanalys.

Fördelar och nackdelar

- Modellbaserad tillvägagångssätt:
 - Statistiska egenskaper används (osäkerhetsmått, konfidens- och prediktionsintervall, jämföra modeller med diagnostiska tester osv.)
 - Säsongrensning helt beroende av "rätt" modellval!
- Icke-modellbaserade metoder:
 - Mindre känsliga mot extremvärden och ej känsliga mot modellval.
 - Svårt att bedöma hur optimal säsongrensning är: få diagnostiska verktyg.

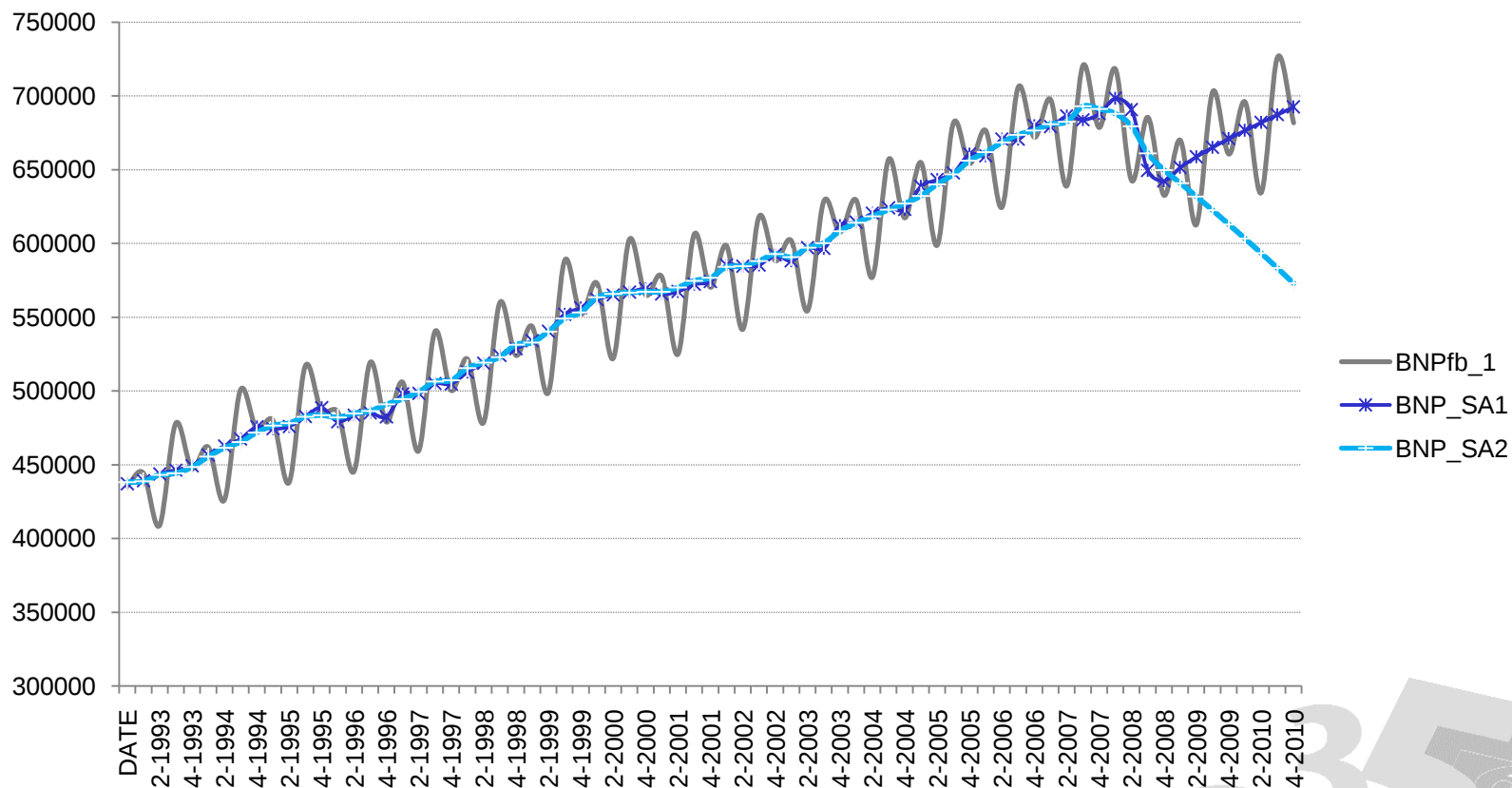


Säsongrensnings problem

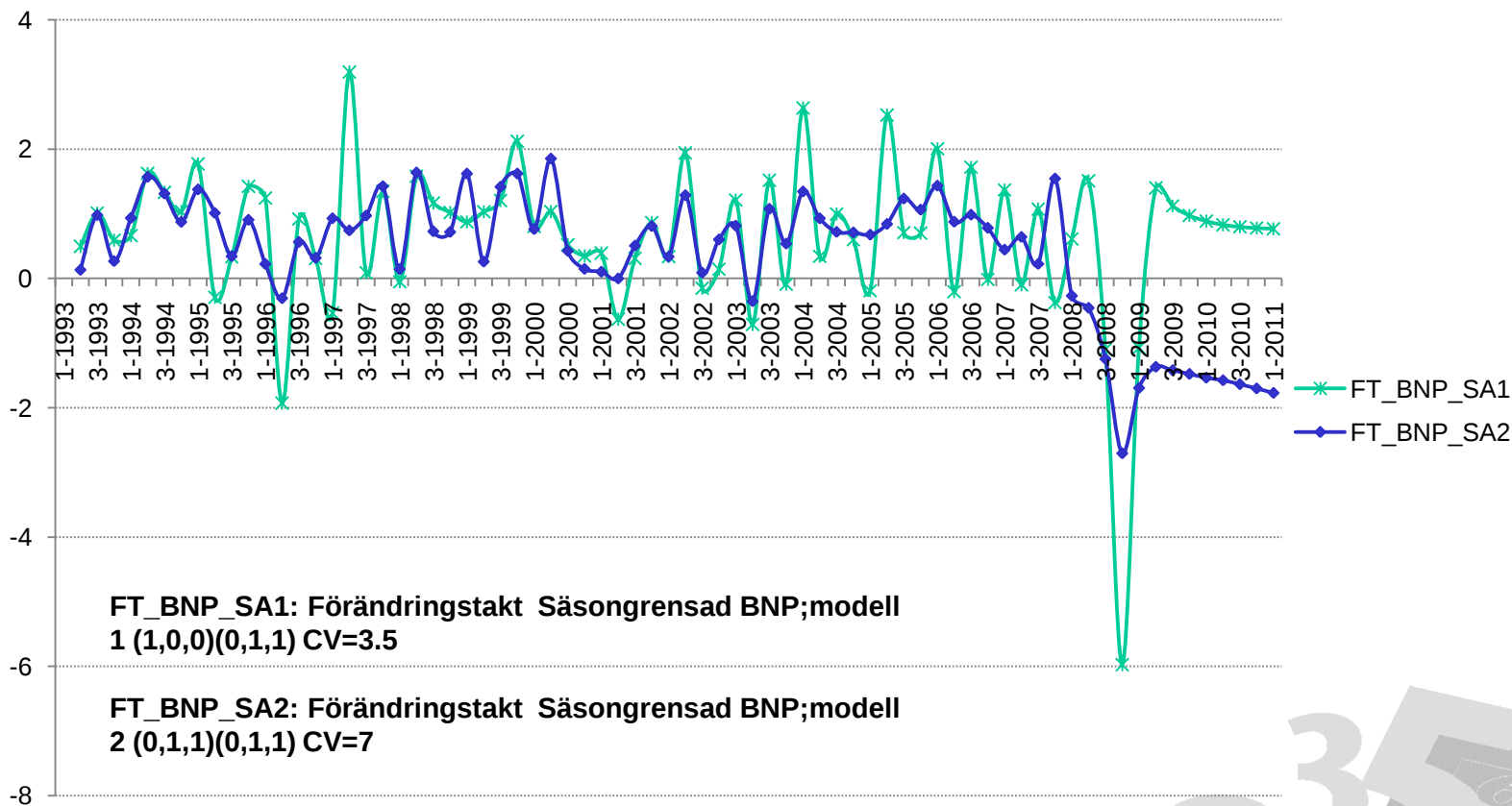
- Samma data men olika resultat:
 - Olika programvaror/verktyg
 - Olika modellval
 - Olika parameterinställningar
 - Kalenderkorrigering och extremvärde hantering
- Säsongrensade data kan resultera i slutsatser som inte stämmer överens med förväntningar (tolkning av ekonomiska utvecklingen)



Ex: BNP säsongrensad med två modeller

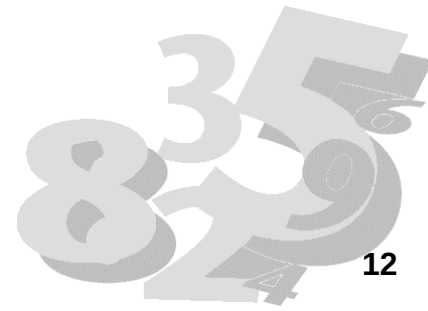


BNP säsongrensad med två modeller: förändringstakt



Projekt: Bakgrund

- Säsongrensade data som SCB redovisar skiljer sig (ibland) från säsongrensade data som vissa användare publicerar.
- Svår att förklara varför en säsongrensad serie har positiv utveckling vid en viss tidpunkt när originalserien uppvisar negativ utveckling vid samma tidpunkt.
- Skillnader mellan statistikernas och ekonomernas synsätt.

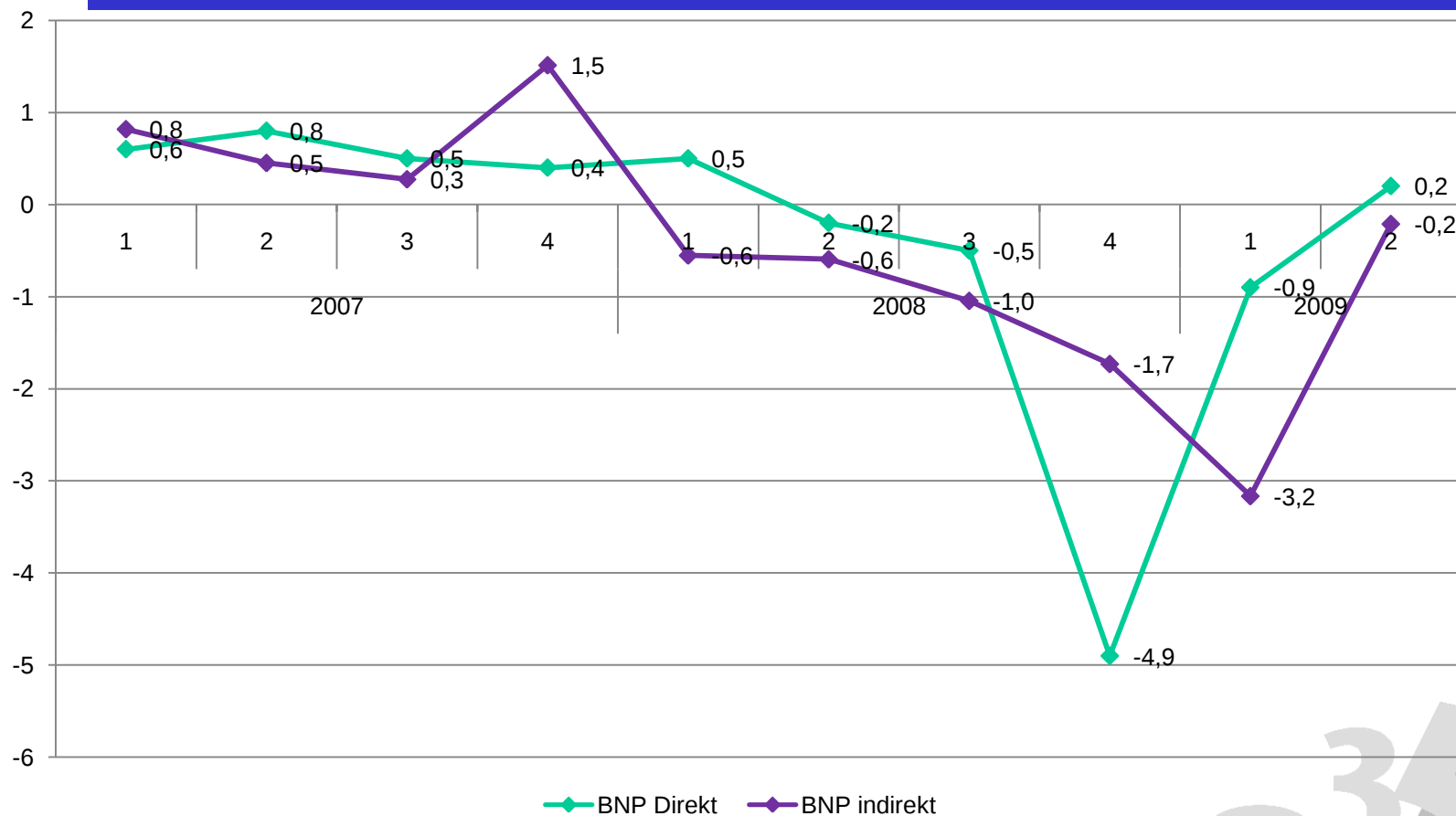


Användarnas (Konj.instit., Riksbanken, m.fl.) önskemål

- Säsongrensade tidserier av delposter ska summera till aggregat
- Ex:
$$\text{BNP} = \text{konsumtion} + \text{investeringar} + \text{Lager} + \text{Export} - \text{Import}$$
- Nu används *direkt* säsongrensning, dvs varje serie rensas för sig med bästa möjliga modell
- Alternativ: *Indirekt* säsongrensning.
Säsongrensade delserier summeras till aggregat.

Direkt vs. indirekt-säsongrensad BNP

Statistics Sweden
Statistiska centralbyrån



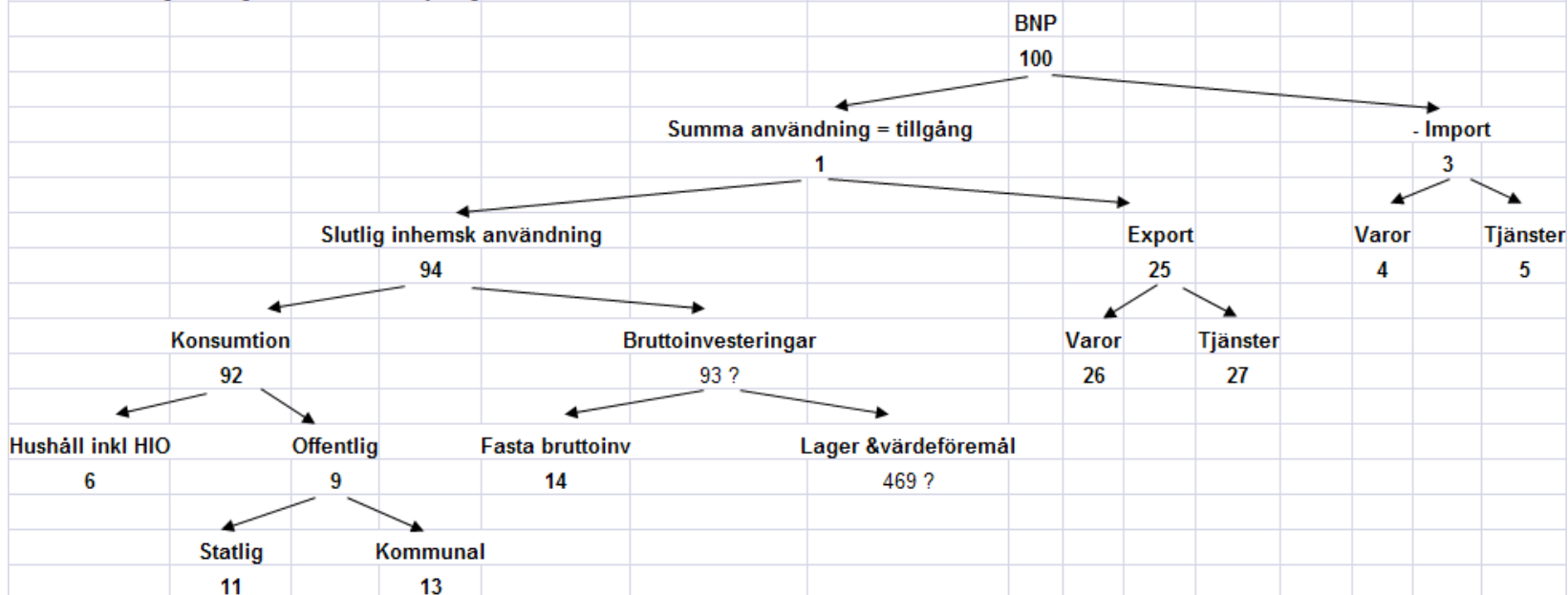
Orsaker till icke-summerbarhet

- Olika modellval för olika delserier (komponenter).
- Olika typer av kalenderkorrigering och outlier-hantering för olika komponenter.
- Olika verktyg.
- Originalserier (ej säsongrensade) ej summerbara p.g.a kedjningsmetoder.
- **Grundorsaken: Säsongrensning är en icke-linjär operation!**

NR data och summerbarhet

Försörjningsbalansen

Utkast till säsongrensning för adderbar nedbrytning



Siffrorna under respektive benämning är seriens aggrid



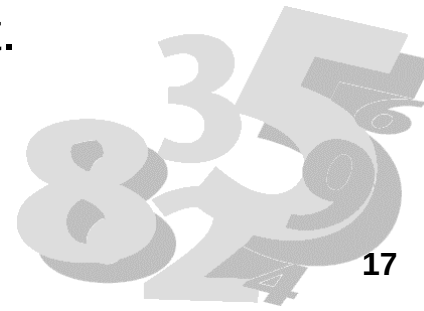
Projekt Summakonsistent säsongrensning

■ Bakgrund

- De viktiga användarna (främst Konjunktursinstitutet (KI) och Riksbanken (RB)) uttrycker önskemål (och krav) på summakonsistent säsongrensning.

■ Projektets mål och avgränsningar:

- Att föreslå en metod som tar hänsyn till KI- och RB:s önskemål.
- Fokusera på nationalräkenskapernas (NR) data.
- Frågor som berör typiska säsongrensningsfrågor (t.ex. outliershantering) ligger utanför projektet.



Projektgrupp

- Projektperiod 1 maj – 30 september
- Projektdeltagare:
 - Anna Hallqvist (ledning; projektledare, administrationsfrågor)
 - Martin Ribe (ledning; övergripande ansvar)
 - Suad Elezović (metodstatistiker; huvudansvarig för metod)
 - Xie Yingfu (metodstatistiker; delansvar för metod)
 - Martin Odenkrants (metodstatistiker; ansvar för tester)
 - Anders Jäder (utredare; kontakt med NR och ansvar för data mm)

Översikt

□ Användarnas rekommendationer:

- I. Utgå ifrån direkt säsongrensad aggregat (BNP).
- II. Säsongrensa alla delserier separat (direkt).
- III. Justera delarna så att summan av delarna överensstämmer med direkt säsongrensad BNP
- IV. Utgå ifrån detaljeringsnivå enligt bilagan.



Försörjningsbalans

Grov

Detaljerad

- BNP
- Import
- Hushållens konsumtion
- Offentlig konsumtion
- Fasta bruttoinvesteringar
- Investering i lager och värdeföremål
- Export

- BNP
- Import av varor
- Import av tjänster
- Hushållens konsumtion
- Offentlig konsumtion kommuner
- Offentlig konsumtion, staten
- Fasta bruttoinvesteringar i näringslivet exkl bostäder
- Fasta bruttoinvesteringar i bostäder
- Fasta bruttoinvesteringar i offentliga myndigheter
- Investering i lager och värdeföremål
- Export av varor
- Export av tjänster

Metod

□ Strategi:

- BNP säsongrensas direkt (**slutgiltigt!**).
- Delserierna enl. detaljeringsnivåer säsongrensas direkt.
- Diskrepanser fördelas på delserier under summakonsistens restriktionen

$$\sum_{i=1}^k \tilde{f}_{FSA}(X_{it}) = f_{SA}(X_t)$$



Projekt Summakonsistent säsongrensning: Presentation av valda ansatser

■ Strategi forts.

- Välj lämplig minimeringsfunktion utifrån användarens behov:
 - Fokusera på nivåer (absolut eller proportionell förändring).
 - Fokusera på förändringstakter.
 - Fokusera på volatilitet (större justering för serier med starkare volatilitet).
 - Kombinationer av de tre aspekterna.
- För att undvika optimerings- (numeriska) metoder:
 - Minimeringsfunktioner som har en analytisk lösning väljs ut!
 - Ett antal diagnostiska mått används för utvärdering av målfunktionerna.

Minimeringsfunktioner

- Exempel på lämpliga minimeringsfunktioner:
 - i. Minimera differensen mellan säsongjusterade serier och ex-post justerade serier m.a.p. nivåer:

$$\min \left[\sum_{t=1}^T \sum_{i=1}^k \left(f_{SA}(X_{it}) - \tilde{f}_{FSA}(X_{it}) \right)^2 \right]$$

- ii. Minimera proportionell differens (procentuell förändring):

$$\min \left[\sum_{t=1}^T \sum_{i=1}^k \left(\frac{f_{SA}(X_{it}) - \tilde{f}_{FSA}(X_{it})}{f_{SA}(X_{it})} \right)^2 \right]$$

Minimeringsfunktioner, forts.

- Minimera differensen mellan två seriers förändringstakter (samma utveckling)

$$\min \left[\sum_{t=2}^T \sum_{i=1}^k \left(\frac{\tilde{f}_{FSA}(X_{it})}{f_{SA}(X_{it})} - \frac{\tilde{f}_{FSA}(X_{it-1})}{f_{SA}(X_{it-1})} \right)^2 \right]$$

- Kombinera nivåer och förändringstakt

$$\min \left[\lambda \left\{ \sum_{t=2}^T \sum_{i=1}^k \left(\frac{\tilde{f}_{FSA}(X_{it})}{f_{SA}(X_{it})} - \frac{\tilde{f}_{FSA}(X_{it-1})}{f_{SA}(X_{it-1})} \right)^2 \right\} + (1 - \lambda) \sum_{t=2}^T \sum_{i=1}^k \left(\frac{\tilde{f}_{FSA}(X_{it}) - f_{SA}(X_{it})}{f_{SA}(X_{it})} \right)^2 \right]$$

För utvärdering används olika avståndsmått...

$$A_1 = \sqrt{\frac{1}{kT} \sum_{t=1}^T \sum_{i=1}^k \{f_{SA}(X_{it}) - \tilde{f}_{FSA}(X_{it})\}^2}$$

$$A_2 = \sqrt{\frac{1}{kT} \sum_{t=1}^T \sum_{i=1}^k \{\tilde{f}_{FSA}(X_{it}) - f_{SA}(\tilde{f}_{FSA}(X_{it}))\}^2}$$

$$A_3 = 100 \times \sqrt{\frac{1}{kT} \sum_{t=1}^T \sum_{i=1}^k \left\{ \frac{\tilde{f}_{FSA}(X_{it})}{f_{SA}(X_{it})} - 1 \right\}^2}$$

$$A_4 = 100 \times \sqrt{\frac{1}{T-1} \frac{1}{k} \sum_{t=2}^T \sum_{i=1}^k \left\{ \frac{\tilde{f}_{FSA}(X_{it})}{f_{SA}(X_{it})} - \frac{\tilde{f}_{FSA}(X_{it-1})}{f_{SA}(X_{it-1})} \right\}^2}$$

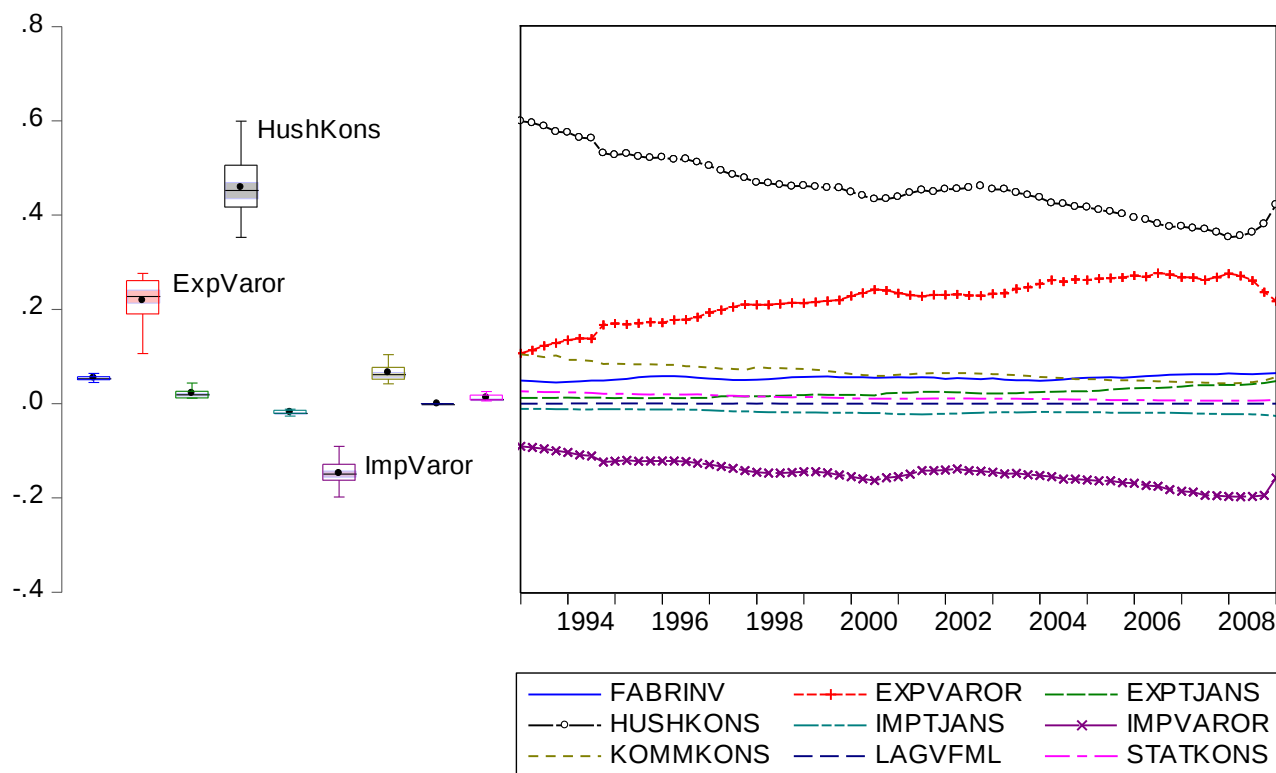
$$A_5 = 100 \times \sqrt{\frac{1}{T-1} \frac{1}{k} \sum_{t=2}^T \sum_{i=1}^k \left\{ \frac{\tilde{f}_{FSA}(X_{it})}{\tilde{f}_{FSA}(X_{it-1})} - \frac{f_{SA}(X_{it})}{f_{SA}(X_{it-1})} \right\}^2}$$

Bedömning av minimeringsfunktionerna

- 6 avståndsmått beräknas för samtliga minimeringsfunktioner.
- "Rankning" av ett och samma kvalitetsmått i förhållande till olika metoder (minimeringsfunktioner).
- Ett genomsnittsvärde för de 6 'ranks' beräknas för varje minimeringsfunktion.
- Minimeringsfunktionen med den lägsta medel-rank väljs ut som den bästa!
- *Obs! Användarna kan bedöma vad som prioriteras och välja en annan minimeringsfunktion.*



Diagram: Diskrepans mellan direkt och indirekt säsongrensad BNP fördelad mellan komponenter enligt PQ metod; värden på lodrätta axeln avser respektive komponenters andel av den totala diskrepansen.



Resultat

Tabell 1. *Rankning av diagnostiska mått för målfunktioner för försörjningsbalansen; alternativ 1.*

Metod \ Rankning	A1	A2	A3	A4	A5	A6
Q	1	2	7	7	11	8
PQ	7	7	1	1	1	2.5
MQ	11	11	6	6	6	5.5
MLQ01	2	4	8	8	10	8
MLQ10	3	3	9	9	9	8
MLQ50	4	1	10	10	8	10
MLQ99	9	5	11	11	7	11
MLPQ01	6	6	2	2	2	2.5
MLPQ10	5	8	3	3	3	1
MLPQ50	8	9	4	4	4	4
MLPQ99	10	10	5	5	5	5.5

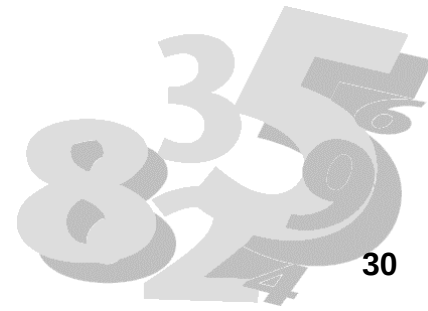
A6: Antal ggr. med olika utveckling i förändringstakterna

Tabell 2. Mått A_6 för försörjningsbalansen fördelat på serier; alternativ 1.

	Import	HushKons	OffKons	FaBrinv	LagVfml	Export	Totalt
Q	0	0	6	2	8	0	16
PQ	0	5	3	0	0	0	8
MQ	1	4	7	2	1	0	15
MLQ01	0	0	6	2	8	0	16
MLQ10	0	0	6	2	8	0	16
MLQ50	0	0	7	4	8	0	19
MLQ99	0	1	8	4	8	0	21
MLPQ01	0	5	3	0	0	0	8
MLPQ10	0	4	3	0	0	0	7
MLPQ50	0	5	7	0	1	0	13
MLPQ99	1	4	7	2	1	0	15
Totalt	1	24	56	16	42	0	139

Slutsatser

- ❑ **PQ** metod fungerar bäst i genomsnitt!
 - *PQ alltid bäst när det gäller mått som tar hänsyn till förändringstakter !*
- ❑ **Viktigt att påpeka är** att vissa andra målfunktioner kan fungera bättre än PQ när enskilda mått studeras.



Fortsatt arbete

- Test på ytterligare material.
- Test då flera aggregeringsnivåer ska vara konsistenta.
- IT-implementering.
- Testa hur outliers, kalendereffekter och olika modeller påverkar summerbarheten?

