

Ekonomisk statistik 3 Economic statistics 3

Masterkurs
Daniel Thorburn
Höstterminen 2008
Stockholms Universitet

Länkning

Ekonomisk statistik
Höstterminen 2008
Stockholms Universitet

Länkning

När statistiken läggs om uppstår ofta språng. T ex när näringsgrensindelningen moderniseras eller när man byter insamlingssystem från disketter till Internet eller ändrar en cutoff-gräns. Vid länkning försöker man räkna om serien bakåt som om det nya insamlingssystemet hade gällt. Bra för personer som arbetar med tidsserier.

Samma problem vid ändringar i samhället. Följ sjuklighetens utveckling i samhället, när antalet karensdagar ändras. Detta brukar ses som en uppgift för dem som analyserar tidsserier. Medan länkning är något för statistikproducenten.

Vid omläggningar bör statistikern planera för länkning. Vid reformer i samhället bör politikerna planera så att reformen kan utvärderas.

Länkning ?

Tänk er att man har en serie som ser ut så här
(Andel positiva till statliga investeringsgarantier %.)

År	1	2	3	4	5	6	7	8	9	10	11	12	13
	23	24	27	27	26	28	30	42	43	45	44	43	47

År 8 lades undersökningen om och frågan formulerades om. Om man vill analysera utvecklingen vill man ha en serie utan språng. Hur skall man göra?

- Svar 1 o 2: Enklast är att sätta värdena år 7 och 8 lika och ändra värdena innan motsvarande (additivt resp multiplikativt).

År	1	2	3	4	5	6	7	8	9	10	11	12	13
Urspr.	23	24	27	27	26	28	30	42	43	45	44	43	47
add	35	36	39	39	38	40	42	42	43	45	44	43	47
Mult	32.2	33,6	37.8	37.8	36.4	39.3	42	42	43	45	44	43	47

Problem: Detta säger inget om skillnaden mellan år 7 och 8.

- Svar 3: (För enkelhets skull ges lösningen bara additivt)

Mer avancerat gör regression $y_t = a + b \cdot I(t < 8) + c \cdot t$ och justera med b

	32	33	36	36	35	37	39	42	43	45	44	43	47
--	----	----	----	----	----	----	----	----	----	----	----	----	----

Andra modeller kan användas. Regression mot en logistisk linje hade förmodligen varit ännu bättre.

Problem: Man vill kunna göra länkningen redan första året efter tidsseriebrottet (dvs år 8).

- Svar 4: Använd samma modell redan år 8 och håll sedan modellen fix

	33	34	37	37	36	38	40	42	43	45	44	43	47
--	----	----	----	----	----	----	----	----	----	----	----	----	----

Ger lite större osäkerhet än svar 3.

- Svar 5: Ännu mer komplicerad modeller kan användas t ex en ARIMA process eller Kalmanfilter. T.ex

- $X_t = a \cdot X_{t-1} + \varepsilon_t$ X är en latent serie
- $Y_t = X_t + \eta_t$ före brott; $Y_t = X_t + b + \eta_t$ efter brott;
- Predikterat värde av Y redovisas före brottet

- Men i allmänhet försöker man välja så enkla modeller som möjligt

- Detta var mekaniska länkingsmetoder för en enda serie där det antogs att man inte hade mer information.
- En vanlig rekommendation är att om det är möjligt göra undersökningen på båda sätten parallellt vid ett tillfälle eller göra andra andra metodstudier. T ex När ULF (undersökningen om levnadsförhållanden) gick över från besök till telefon-undersökning som bas gjorde man under övergångsåret halva urvalet besök och halva per telefon. Slumpmässiga halvvar.
- När större omläggningar görs t ex av SNI-kod försöker man koda uppgifterna på båda sätten under en period eller gå igenom gamla statistikunderlag (blanketter) och koda om företagen enligt den nya mallen
- Dessa metoder ger en uppskattning av språnget - men inte över hur man skall ändra längre bakåt

- Vid bl a omklassificering, t ex omläggning av SNI, kanske man vet summan av två grupper men det tillkommer en uppdelning. T ex när Dator och IT-konsulter bröts ur kategorin annan konsultverksamhet. Då visste man att den inte under de senaste tio åren haft samma andel av den större kategorin.
- Då gäller det att föra en uppdelning bakåt. Då får man lägga in en subjektiv gissning. T ex andel 0 tio år tidigare och sedan en linjär ökning av andelen till den första observerade nivån. Eller samma andel hela tiden bakåt (Dumt i just detta fall)

- Om man länkat en dataserie, skall den som använder serien alltid kunna läsa om vad som hänt. Orsaken till språnget och hur det eliminerats.
- Dokumentera vad du gör

Modeller för svarsprocessen i företag

Ekonomisk statistik
Höstterminen 2008
Stockholms Universitet

Svarsprocessen

Teorin för hur företagens uppgiftslämnande går till är inte lika väl utvecklad som för insamling från individ

Men speciellt för småföretag (lantbruk, hantverkare, butiksinnehavare etc) gäller till stor del samma regler

Vi skall titta främst på det som är speciellt för datainsamling från företag

- Företag är speciella i vissa avseenden.
- För många av dem är tid pengar - De är vana att kunna ta betalt för sin tid - Undvik att "erbjuda" en löjligt låg timpeng/ ersättning (t ex trisslott). Bättre ingen ersättning alls. Ersättning till uppgiftslämnaren för det han gör på betald arbetstid är också oetiskt/olämpligt.
- Datainsamling via telefon är i allmänhet svårt Företagaren har sällan alla uppgifter i huvudet eller ofta är upptagen. (Men däremot efter granskning)

Kognitiva Modeller

- Torangeaus (1984) modell för individers svarande:
 - Comprehension
 - Försök förstå frågan
 - Retrieval
 - Leta i minnet och försök hitta ett svar
 - Judgment
 - Bedöm om detta är det önskade svaret och om jag vill lämna ut det
 - Formatting
 - Formulera svaret (t ex i en mening eller välj det alternativ som passar bäst)
- Det kan gå fel i alla steg. Konstruera enkäten så att man garderar sig för alla typer av fel.

Motsvarigheten för företag

- Encoding the information in memory or company records
 - Selection of the respondent
 - Assessment of priorities
 - Comprehension of data request
 - Retrieval of relevant information from memory or records
 - Judgment
 - Communication of the response
 - Release of the data
- Sudman, Willimack, Nicholls, Meusenbourg (2000)

Socialt distribuerad kognition

- Kunskaperna finns inte hos en enda person eller på ett ställe i företaget
- Mycket av vad som i Torangeaus modell skedde i respondentens huvud sker här öppet och kan alltså observeras
- Vem frågar han, vilka filer öppnas, vem godkänner utlämnande etc
- Propagering av information och fel. Lite som viskleken. Vilka fel uppstår när information flyttas från ett medium/person till en annan
- Hutchins (1995) dock ej vid datainsamling utan allmänt vid informationsspridning inom organisationer. Blivit populärt mest inom data- och systemvetenskap.

Encoding the information in memory or company records

På vad sätt kan man påverka att rätt data sparas och är åtkomligt inom företag?

Encoding the information in memory or company records

På vad sätt kan man påverka att rätt data sparas och är åtkomligt inom företag?

- Fråga bara om det som sparats. Välj t ex definitioner som väl ansluter till vanliga kontoplaner. Fråga dig alltid vid frågekonstruktionen om det finns chans att få rätt svar.
- Stöd utvecklingen av lämpliga redovisningssystem
- Be om sådant som händer framöver.

Selection of the respondent

Hur ser man till att formuläret når de som kan svara på frågan ? (undvika att blanketten går till t ex en semestervikarie eller prao-elev)

Selection of the respondent

Hur ser man till att formuläret når de som kan svara på frågan ? (undvika att blanketten går till t ex en semestervikarie eller prao-elev)

- Fasta kontaktpersoner, som känner företaget och ansvar. De kan vidarebefordra eller söka sig till rätt person
- En postöppnarfunktion kan ofta vara ett problem
- Vända sig till någon med chefsställning som har överblick och kan ge någon i uppdrag
- Adressera till en speciell funktion (t ex personalansvarig, inköpschef, etc)
- Fråga vem som ansvarar för ..., vem som svarade på blanketten. (Bra att veta nästa gång)

Assessment of priorities

Hur får man företagen att svara och svara med viss omsorg och svara inom rimlig tid? Att få företaget och den svarande att prioritera arbetsuppgiften

Assessment of priorities

Hur får man företagen att svara och svara med viss omsorg och svara inom rimlig tid? Att få företaget och den svarande att prioritera arbetsuppgiften

- Lagstadgat krav, viteshot
- Motivera respondenten t ex
 - I följebrev
 - Få stöd av branschorganisationer eller andra med högt förtroende
 - Få in text i branschtidningar
 - Visa att man själv tar uppgiften på allvar. Slarviga eller löjliga undersökningar inger ej förtroende. Enbart WEB-enkäter ger inte förtroende.
- Uppgiftslämnarvård
 - Återrapportering, Tackkort, Julkort

Torangeau

- Comprehension of data request
- Retrieval of relevant information from memory or records
- Judgment
- Communication of the response

Dessa punkter skiljer sig inte mycket från datainsamling från individer (t ex blanketttest, frågelabb, fälttest, intervjuarutbildning, etc) .

Men ibland:

Comprehension, Retrieval, Communication

Hur ser man till att företagen får fram rätt information ur registren? Och att rätt information lämnas?

Comprehension, Retrieval, Communication

Hur ser man till att företagen får fram rätt information ur registren? Och att rätt information lämnas?

- Hjälp till att programmera bokförings-systemet/uppgiftslämnandet, så att uppgifter kan tappas direkt från systemet (av person vid företaget eller av statistikmyndigheten direkt, "Electronic Data Interchange")
- Ta emot information i olika form (t ex de flesta filformat/programspråk ASCII-fil, Excelark, via Internetformulär, e-mail-bilaga, papper, CD, etc. Lägg formateringsbekymret hos SCB
- Granskning och återkontakter. Få uppgiftslämnarens namn och anknytning på svaret.

Release of the data

Hur får man företaget att lämna ut uppgifter som potentiellt kan vara affärshemligheter eller motverka företaget i andra avseenden

Release of the data

Hur får man företaget att lämna ut uppgifter som potentiellt kan vara affärshemligheter eller motverka företaget i andra avseenden

- Strikt sekretessprövning. Det kan vara bra att ibland vägra lämna ut uppgifter så att vägran kommer till potentiella respondenters kännedom
- Bygga upp ett förtroende. Inte göra tveksamma undersökningar eller göra politiskt opportuna analyser/undersökningar.
- Skapa ansvarskänsla hos de svarande. Tillsammans medverkar de till ett bättre Sverige/en bättre värld.

Inte mycket till modell

- Mycket sämre teoretiskt underbyggt än insamling från individer och hushåll.
- Mest en samling goda rekommendationer
- Företag är mer olika sinsemellan än individer. Det betyder att frågekonstruktion och kontaktstrategier är mer varierande.
- Frågetekniken beror också mer av ämnesvalet.