



Flexibelt startår i Raps

- Förutsättningar och vägledning vid modellvalidering och kontrafaktisk analys

Tillväxtverket stärker Sverige genom att stärka företagens konkurrenskraft

Vi skapar bättre förutsättningar för företagande och bidrar till attraktiva regionala miljöer där företag utvecklas. Våra verktyg är kunskap, nätverk och finansiering: Kunskap om företag och regioner. Nätverk för effektiv samverkan. Och finansiering som stärker näringslivet. Tillväxtverket är nationell myndighet med regional närvaro - vi är 430 medarbetare och har kontor på nio orter. Ett Sverige med fler företag som vill, kan och vågar är vår vision.

Regionalt analys- och prognossystem (Raps) är ett verktyg för regional planering. Raps utgår från en databas med regional statistik som täcker ett flertal områden som är viktiga för regional utveckling.

Tillväxtverkets publikationer kan laddas ner på tillvaxtverket.se. Vill du beställa en tryckt publikation eller söker du en publikation som publicerades innan 2015 hänvisar vi till vår webbshop publikationer.tillvaxtverket.se.

© Tillväxtverket

Stockholm, Februari 2018

Version: 1

Har du frågor om denna publikation, kontakta:

Magnus Hägg

Telefon, växel 08-681 91 00

Innehåll

1	Bakgrund	4
1.1	Basår i Raps	4
1.2	Motiv för flexibelt basår.....	4
2	Förutsättningar för flexibelt basår	5
2.1	Datamässiga förutsättningar i Raps 5.0	5
2.2	Förutsättningar för modellvalidering (att tänka på).....	5
2.3	Förutsättningar för kontrafaktisk analys (att tänka på).....	6

1 Bakgrund

1.1 Basår i Raps

I nuvarande version av Raps 5.0 (1:a halvåret 2018) är basåret 2013 vid körning med full modell, och 2014 vid körning med partiell befolkningsmodell¹. Liksom i tidigare versioner av Raps är basåret det *senaste* år för vilket det finns data för *alla* variabler med årsdimension i den databas som används. Normalt är det den ekonomiska statistiken som är eftersläpande och som medför att basåret ligger några år tillbaka i tiden.

Motivet för att låta basår vara senast möjliga år är att analyser med Raps i regel är framåtblickande, t ex vid prognoser/scenarios. Det finns dock frågor där analyser i ett bakåtblickande perspektiv kan vara motiverade. Sådana analyser förutsätter att det är möjligt att välja ett tidigare historiskt år som basår.

1.2 Motiv för flexibelt basår

Bakåtblickande analyser har två huvudsakliga syften, modellvalidering och kontrafaktisk analys.

Modellvalidering innebär t ex en analys av skillnader mellan modellresultat och statistik (faktiskt utfall) över, säg, en tioårsperiod. Analysen kan fokusera på olika aspekter, exempelvis hur avvikelser mellan modellresultat och statistik beror på specifika parameterantaganden. Ett annat exempel på modellvalidering kan vara att analysera hur valet mellan olika historiska basår påverkar skillnaden mellan modellresultat och statistik. Ett tredje, och närliggande, exempel är analyser av skillnaden i modellresultat mellan en körning med ett tidigare historiskt basår, respektive en modellkörning med senast möjliga basår.

Kontrafaktiska analyser kan beskrivas som analyser genomförda i två steg. I det första steget körs modellen med ett historiskt basår, där modellen kalibreras så att resultatet så nära som möjligt överensstämmer med statistik (faktiskt utfall). I det andra steget införs i modellen en händelse (t ex aktivitet) för att belysa det modellberäknade kontrafaktiska utfallet av att denna händelse inte inträffar. Det kan handla om att analysera effekten av att lägga ned en verksamhet som faktiskt har etablerats under perioden, eller effekten av att lägga till en verksamhet som faktiskt har upphört under perioden.

¹ Hösten 2018 kommer en ny uppdaterad version med basår 2015 vid körning med full modell, och basår 2016 vid körning med partiell befolkningsmodell.

2 Förutsättningar för flexibelt basår

2.1 Datamässiga förutsättningar i Raps 5.0

I många modellberäkningar används statistikvariablers värden för föregående år ($t-1$) vid beräkning av värden för innevarande år (t). I några fall används variabelvärden med mer än ett års lag ($t-2$, $t-3$, $t-4$). Det senare fallet rör variabler för bostäder och byggande i delmodell 4, Bostadsmarknaden, och variabeln för förvärvsinkomst som i beräkning av inkomstutjämnning i delmodell 5 använder värden för år $t-2$. Vid beräkningen av kalibreringsfaktorer i projektgenereringen används också variabelvärden med mer än ett års lag. I Raps 5.0 är 2008 det *tidigaste* år för vilket det finns data för *alla* statistikvariabler med årsdimension i databasen. De statistikvariabler som i modellberäkning (och projektgenerering) använder variabelvärden med mer än ett års lag har i databasen värden för motsvarande år. Samtliga *parametrar* med årsdimension har parametervärden från 2008. Därför, med avseende på data för både *statistikvariabler* och *parametrar* kan 2008 tjäna som basår ($t-1$), med 2009 som första år (t) i modellberäkningen,

2.2 Förutsättningar för modellvalidering (att tänka på)

Vid modellvalidering analyseras skillnader mellan modellresultat och statistik. Avvikelser mellan Raps modellresultat och faktisk utveckling beror i princip på två faktorer: fel i nationella förutsättningar (dvs. fel i nationell modell) och fel i regional modell. För att kunna genomföra en modellvalidering med relevanta data är därför parametrarna i Raps för åren 2008-2013 baserade på de nationella förutsättningar som var aktuella 2008. Det gäller både antaganden om ekonomi och arbetsmarknad från Långtidsutredningen 2008 (LU) och SCB:s bakomliggande antaganden om befolkningsutvecklingen.

Här är det viktigt att tänka på att parametrarna från LU är antaganden om *genomsnittlig tillväxt per år*, baserade på ett nationellt scenario för utvecklingen på lång sikt; i LU 2008 med slutår 2030. Den modell som används på nationell nivå är en statisk allmän jämviktsmodell som inte beskriver anpassningsförloppet mellan startår och slutår. Det innebär bland annat att man för perioden 2009-2013 kan förvänta mycket stora avvikelser mellan modell och faktiskt utfall, till följd av den globala nedgång i konjunkturen som utlöstes av krisen på den amerikanska finansmarknaden hösten 2008. Man kan generellt förvänta relativt stora avvikelser mellan modell och faktiskt utfall, även för en period med mer måttliga konjunkturvariationer. Det är helt enkelt inte meningsfullt att validera modellen med avseende på den ekonomiska utvecklingen på kort sikt (ca 5 år), eftersom parametrarna (tillväxttakterna) är antaganden om utvecklingen på längre sikt (>25 år).

Till skillnad från LU är SCB:s framskrivning på lång sikt baserad *årliga prognosdata* (parametrar för födelsetalens och dödstalens årliga förändring, samt årlig in- och utvandring). Därmed kan parametrarna för befolkningsutvecklingen på kort sikt användas för att validera befolkningsmodellen i Raps. Med 2008 som basår är det både möjligt och meningsfullt att analysera i vilken grad fel i den regionala befolkningsframskrivningen beror på fel i den regionala modellen, respektive fel i nationella antaganden, t ex antaganden om in- och utvandring.

2.3 Förutsättningar för kontrafaktisk analys (att tänka på)

Om syftet med ett tidigt basår är att genomföra en kontrafaktisk analys spelar det inte så stor roll vilka parameterantaganden som tillämpas. Modellen ska hur som helst kalibreras för att så nära som möjligt replikera den faktiska utvecklingen, som underlag för den kontrafaktiska analysen. Tekniskt underlättas kalibreringen av att modellen gör det möjligt att autokalibrera antalet födda och antalet döda. Regionens faktiska befolkningsutveckling kan därmed relativt enkelt replikeras genom att i modellen använda statistik för in- och utflyttning, in- och utvandring. På motsvarande sätt, med val av exogent bostadsbyggande, kan statistik för bostadsbyggande läggas in i modellen. I modellen är det även möjligt att autokalibrera antalet sysselsatta per bransch.

Vid kontrafaktisk analys är det viktigt att tänka på vilka konsekvenser det medför att använda autokalibrering. Vi kan belysa med ett exempel. Anta en kontrafaktisk analys som avser effekterna av att lägga ned en verksamhet som faktiskt har etablerats under perioden. I det första steget (1) kalibreras modellen för att så nära som möjligt återge den faktiska utvecklingen, enligt statistik. Här används autokalibrering av antal sysselsatta per bransch.

I det andra steget (2) körs modellen med en aktivitet = det bortfall av produktion, antal sysselsatta och inkomster som verksamheten svarar för. Aktiviteten är specificerad med avseende på mest närliggande bransch². Om autokalibrering används även i detta steg blir den beräknade sysselsättningen per bransch densamma som i det första steget. Dvs., den beräknade effekten på regionens sysselsättning består endast av aktivitetens direkta effekt. Om autokalibreringen inte används i det andra steget kommer sysselsättningen i respektive bransch inte vara densamma som i det första steget. Men, det går inte att avgöra om skillnaden beror på aktivitetens beräknade indirekta effekt, eller om skillnaden beror på att sysselsättningen per bransch inte har kalibrerats.

För att hantera detta problem kan analysen läggas upp enligt följande. Låt det andra steget bestå av tre körningar, (2a), (2b) och (2c). I körning (2a) genomförs en beräkning utan kalibrering och utan aktiviteten. Körning (2b) utgår från (2a) med aktiviteten införd (dvs. nedläggning av verksamhet). Skillnaden mellan (2b) och (2a) fångar modellberäknade direkta och indirekta effekter av aktiviteten. Beräkna den relativa förändringen av sysselsatt dagbefolkning (SYS) och beskattningsbar förvärvsinkomst (FINK). Tillämpa den relativa förändringen på resultatet för SYS och FINK i det första steget (1), för att beräkna skillnaden i absoluta tal. Körning 2c utgår från körningen i det första steget (1) med en aktivitet som *inte* är specificerad med avseende på bransch. Antal sysselsatta i aktiviteten är den enligt ovan beräknade skillnaden för SYS; lön per sysselsatt beräknas så att aktivitetens lönesumma stämmer med den enligt ovan beräknade skillnaden för FINK.

Målet är att i relativa tal ska skillnaden i resultat mellan (2c) och (1) bli densamma som mellan (2b) och (2a). För att uppnå denna överensstämmelse behöver (2c) köras om, med justerad lön per sysselsatt för aktiviteten. Skälet är att aktivitetens lönesumma får effekt på den privata konsumtionen, som i sin tur påverkar regionens inkomst och sysselsättning.

² Det innebär att indirekta effekter på branscher beräknas.