

## FÖRDJUPNING

### Dekomponering av förändringar i lönespridning

Förändringar i lönespridning mellan två tidpunkter kan vara en följd av sammansättningseffekter eller av priseffekter. Sammansättningseffekterna följer av att fördelningen av lönepåverkande faktorer – till exempel ålder, utbildning och yrke – bland lönta-garna har förändras. Preiseffekterna följer av att de lönepåverkande faktorernas pris, det vill säga hur faktorerna påverkar lön, har förändrats.

Ett sätt att analysera vad som påverkat lönespridningen är att genomföra en så kallad dekomponering. Vid en dekomponering delas förändringen upp i en förklarad del och en oförklarad del. Den förklarade delen är den del av förändringen som kan förklaras av att fördelningen av lönepåverkande faktorer har förändrats. Om alla lönepåverkande faktorer inkluderas i dekomponeringen skulle den oförklarade delen enbart bestå av priseffekter. Faktorerens observerbarhet och tillgänglig statistik begränsar emellertid vilka faktorer som kan inkluderas i analysen. Detta gör att resultaten bör tolkas med försiktighet.

Dekomponering bygger på att ett kontrafaktiskt utfall beräknas. Här utgörs det kontrafaktiska utfallet av den lönespridning som hade observerats om faktorerna varit fördelade som 2012 men fått ersättning (pris) som 2003.<sup>104</sup> Givet det kontrafaktiska utfallet kan differensen av lönespridningen mellan 2012 och 2003 skrivas som

$$V_{12}^{12} - V_{03}^{03} = \underbrace{V_{12}^{12} - V_{12}^{03}}_{\text{oförklarad}} + \underbrace{V_{12}^{03} - V_{03}^{03}}_{\text{förklarad}}$$

där  $V_j^i$  är lönespridningen med faktorfördelning som år  $j$  och priser som år  $i$ .  $V_{12}^{12}$  och  $V_{03}^{03}$  är således den uppmätta lönespridningen 2012 respektive 2003.  $V_{12}^{03}$  är den kontrafaktiska lönespridningen. Differensen mellan den kontrafaktiska lönespridningen och 2003 års lönespridning utgör den förklarade delen av förändringen i lönespridning. Den förklarade delen anger således vad förändringen hade varit om endast faktorfördelningen ändrats mellan åren. Den oförklarade delen, som utgörs av differensen av 2012 års lönespridning och den kontrafaktiska lönesprid-

<sup>104</sup> En alternativ dekomponering är att använda det kontrafaktiska utfallet där de lönepåverkande faktorerna är fördelade som 2003 och fått ersättning (pris) som 2012. Valet av kontrafaktiskt utfall är en fråga om att välja referensår för priserna, det vill säga ett indexproblem. Resultaten av en dekomponering påverkas av vilket referensår som väljs. Huvuddragen för resultaten i det här kapitlet pekar åt samma håll oavsett val av referensår.

ningen, anger hur mycket mindre lönespridningen hade varit 2012 om inte priserna förändrats.

Det finns olika sätt att beräkna det kontrafaktiska utfallet.<sup>105</sup> Här används en metod där observationerna för 2003 viktas om för att återspegla fördelningen av observerbara faktorer för 2012.<sup>106</sup> Metoden består av tre steg. Baserat på data för de olika tidpunkterna skattas en logit-modell för sannolikheten att en observation är från 2012, det vill säga  $\Pr(D_{12} = 1|X_i)$ . Oberoende variabler i logit-modellen är de lönepåverkande faktorer som ingår i analysen. I det andra steget beräknas vikterna med hjälp av den skattade logit-modellen som

$$\hat{\Psi}(X_i) = \frac{\widehat{\Pr}(D_{12} = 1|X_i)/\widehat{\Pr}(D_{12} = 1)}{\widehat{\Pr}(D_{12} = 0|X_i)/\widehat{\Pr}(D_{12} = 0)}$$

där  $\widehat{\Pr}(D_{12} = 1|X_i)$  och  $\widehat{\Pr}(D_{12} = 0|X_i)$  är de predikterade sannolikheterna att observation  $i$  är från 2012 respektive 2003 givet de oberoende variablerna  $X_i$ .  $\widehat{\Pr}(D_{12} = 1)$  och  $\widehat{\Pr}(D_{12} = 0)$  är andelen av totala antalet observationer från respektive tidpunkt. I det tredje steget beräknas den kontrafaktiska utfallet genom att observationerna för 2003 viktas om med  $\hat{\Psi}(X_i)$ , varefter lönespridningen beräknas.

### Lönepåverkande faktorer i dekomponeringarna

Vid skattningen av logit-modellerna ingår ett antal förklarande variabler som samtliga är indikatorvariabler (dummyvariabler). Det är viss variation i variablerna beroende på sektor. För samtliga dekomponeringar indikerar variablerna vilken *åldersgrupp* (18–24, 25–29, 30–34, ..., 60–64) individen tillhör samt individens *utbildningsnivå* och *utbildningsinriktning* (SUN 2000). För dekomponeringarna på sektornivå ingår individens *köns* som lönepåverkande faktor. För privatanställda tjänstemän och arbetare ingår också inom vilken *näringsgren* (SNI2002 bokstavs nivå med A+B) individen arbetar. Den sista lönepåverkande faktorn är vilket *yrke* som individen har. Utgångspunkten för denna variabel är 2-ställig SSK, men för landsting och kommuner indikeras vissa yrken på högre nivå. För både landsting och kommuner är huvudgrupp 51 (Service, omsorg och säkerhetsarbete) uppdelad på 4-ställig nivå. Det samma gäller huvudgrupp 24 (Lärare inom universitet, gymnasie- och grundskola) för kommuner. Vidare är huvudgrupp 24 (Annat arbete som kräver teoretisk specialkompetens) uppdelad för kommuner, men på 3-ställig

<sup>105</sup> Se Fortin m.fl. (2011).

<sup>106</sup> DiNardo m.fl. (1996).

nivå. För landsting gäller att barnmorskor och sjuksköterskor med särskild kompetens (SSYK 223), sjuksköterskor (SSYK 323) och läkare (SSYK 2221) indikeras separat.