FÖRDJUPNINGS-PM

Nr 11. 2012

Detaljerad Dekomponering: Alternativa Lösningsmetoder för Invariansproblemet

Av Thomas Andrén

Sammanfattning
Vid detaljerad dekomponering av skillnader i genomsnittlig lön mellan två grupper är den icke-förklarade delen inte entydigt identifierad när grupper av dummyvariabler används i en regressionsmodell. Det beror på att metoden inte är invariant med avseende på vilken referensgrupp som används. Det vill säga, storlek och tecken på den skattade effekten av faktorns betydelse för skillnaden i genomsnittlig lön mellan två grupper kommer att vara beroende av vilken kategori som utesluts. Forskningslitteraturen redovisar ett antal metoder för att hantera detta problem. Dessa metoder bygger huvudsakligen på en ansats där dummyvariablerna normaliseras i regressionsmodellen. I denna PM beskrivs några av dessa metoder i detalj och resultat presenteras för att illustrera skillnader. Resultaten antyder att det för närvarande inte finns någon tillfredställande metod för att lösa invariansproblemet. Vid jämförelse av andra forskares resultat är det därför viktigt att använda samma referenskategori för att jämförelsen ska bli meningsfull.

1. Inledning

I regleringsbrevet framgår att Lönebildningsrapporten bland annat ska innehålla en analys där relevanta jämställdhetsaspekter beaktas. En återkommande del i den analysen berör skillnader i genomsnittlig lön mellan kvinnor och män, där försök görs att beskriva vilka faktorer som ligger bakom storleken på lönegapet. Detta sker huvudsakligen med hjälp av regressionsanalys. I det samanhanget används ofta dekomponeringstekniker, vilket är metoder som är vanligt förekommande i forskningslitteraturen. Problemet med ansatsen är att en detaljerad dekomponering av den icke-förklarade delen med avseende på dummyvariablernas koefficienter som förekommer i regressionsmodellen är oidentifierad i så motto att resultatet blir godtyckligt styrda av vilken referenskategori som används. Eftersom regressionsmodellerna huvudsakligen innehåller dummyvariabler är detta ett problem som ofta löses genom att inte redovisa den oförklarade delen vid detaljerad dekomponering. Eftersom den oförklarade delen i någon mening beskriver betydelsen av att värderingen skiljer sig mellan grupperna, innebär en exkludering att värdefull och intressant information går förlorad. Nedan följer nu en beskrivning av olika metoder som hanterar detta problem samt en diskussion om hur väl metoderna lyckas med uppgiften. Promemorian inleder med en beskrivning av standardansatsen för dekomponering samt en illustration av hur invariansproblemet påverkar resultaten.

2. Blinder-Oaxaca Dekomponering

Dekomponering av ekonomiska variabler förekommer ofta i den nationalekonomiska litteraturen. Det kan till exempel handla om dekomponering av lönegap, löneutveckling mellan två år eller skillnader i arbetskraftdeltagande mellan olika demografiska eller regionala grupper. Forskningslitteraturen innehåller en mängd olika applikationer av detta slag där ekonomiska storheter delas upp i en del som kan förklaras med skillnader i sammansättning med avseende på ett antal observerbara faktorer och i en residual del som beror på andra faktorer som är okända eller icke-observerbara för statistikern.

2.1 STANDARDANSATSEN

Den vanligast förekommande metoden för dekomponering bygger på Oaxacas (1973) och Blinders (1973) forskning som studerar skillnader i genomsnittlig lön mellan kvinnor och män. Utgångspunkten för deras analyser är en grupp av kvinnor och män vars löner observeras vid ett givet tillfälle. Syftet är att undersöka hur mycket de genomsnittliga lönerna skiljer sig mellan grupperna. Storleken på lönerna varierar från individ till individ och mellan kvinnor och män på gruppnivå, och det finns en mängd faktorer som kan förklara varför de genomsnittliga lönerna skiljer sig åt mellan grupperna. En del av dessa faktorer är observerbara medan andra är icke-observerbara (för statistikern). Med hjälp av statistiska lönemodeller kan individens lön delas upp i två delar för respektive grupps individer:

$$Y_M = X_M \beta_M + U_M \qquad \text{(Män)} \qquad (1)$$

$$Y_K = X_K \beta_K + U_K \qquad \text{(Kvinnor)} \qquad (2)$$

Y_K och Y_M avser logaritmerade löner för kvinnor respektive män, vars nivå kan förklaras av ett antal observerbara lönepåverkande faktorer som finns representerade i variabelmatriserna X_K och X_M. Dessa matriser innehåller dels faktorer som är relaterade till individens produktivitet, så som utbildning, yrkeserfarenhet, ålder, men också faktorer som är relaterade till arbetsplatsen så som typ av arbete, arbetsplatsen storlek och region. Koefficienterna för X_K och X_M representerar de observerbara faktorernas marginaleffekter på inkomsten eftersom lönerna är logaritmerade. Det vill säga, när en faktor som ingår i X ökar med en enhet kommer en genomsnittlig individs lön att förändras med motsvarande $\beta \times 100$ procent.

Eftersom individernas löner aldrig är helt deterministiska, och vissa individer har olika lön trots identiska observerbara karaktäristika behöver ekvationerna kompletteras med var sin slumpterm (U_K och U_M) som i ett aggregat representerar de icke-observerbara faktorerna, det vill säga sådana faktorer som påverkar individens lön men som inte kan observeras av statistikern. Det skulle till exempel kunna handla om olika typer av personliga egenskaper som främjar eller reducerar individens lön på grund av att de är olika hårt betonade hos olika individer. Det kan också handla om andra vanligtvis observerbara faktorer som statistikern inte har tillgång till och som av den anledningen inte ingår explicit i modellen.

Parametrarna i ekvationerna (1) och (2) kan skattas med hjälp av minsta kvadratmetoden (OLS), vars egenskaper medför att genomsnittslönerna för de två grupperna kan formuleras på följande sätt: $\overline{Y}_M = \overline{X}_M \beta_M$ och $\overline{Y}_K = \overline{X}_K \beta_K$. Linjerna över variabelmatriserna indikerar att det handlar om genomsnittsvärden för löner och observerbara faktorer.

$$\overline{Y}_M - \overline{Y}_K = \left(\overline{X}_M - \overline{X}_K\right)\beta_M + \overline{X}_K\left(\beta_M - \beta_K\right) \qquad (3)$$

$$\overline{Y}_M - \overline{Y}_K = \left(\overline{X}_M - \overline{X}_K\right)\beta_K + \overline{X}_M\left(\beta_M - \beta_K\right) \qquad (4)$$

$$\overline{Y}_K - \overline{Y}_M = \left(\overline{X}_K - \overline{X}_M\right)\beta_M + \overline{X}_K\left(\beta_K - \beta_M\right) \qquad (5)$$

$$\overline{Y}_K - \overline{Y}_M = \left(\overline{X}_K - \overline{X}_M\right)\beta_K + \overline{X}_X\left(\beta_K - \beta_M\right) \qquad (6)$$

Skillnaden i genomsnittlig lön mellan två grupper kan dekomponeras på ett flertal olika sätt beroende på vad det är som ska mätas. För det första handlar det om huruvida man är intresserad av att veta hur mycket *lägre* kvinnors genomsnittliga löner är jämfört med mäns (ekvation 5 och 6), eller hur mycket *högre* mäns genomsnittliga löner är jämfört med kvinnors (ekvation 3 och 4).

För det andra handlar det om huruvida det är kvinnornas eller männens genomsnittliga löner ska betraktas som norm. Om männens löner betraktas som norm, är frågan hur mycket högre kvinnors genomsnittliga löner kommer att vara om deras lönepåverkande faktorer prissattes (värderades) på samma sätt mäns (jämför $\bar{Y}_K = \bar{X}_K \beta_K$ med $\bar{Y}_K^* = \bar{X}_K \beta_M$). På motsvarande sätt kan också kvinnors löner betraktas som norm, vilket innebär att frågan istället blir hur mycket lägre mäns genomsnittliga löner skulle vara om deras lönepåverkande faktorer prissattes på samma sätt som kvinnors. Den ansats som oftast används utgår från mäns löner som norm och att kvinnor har lägre genomsnittlig lön, vilket motsvaras av ekvation (3) (eller ekvation(5)).

Givet att modellen är korrekt specificerad så handlar valet av norm om huruvida den ena eller andra gruppen är diskriminerad då alla övriga förklaringar är uttömda. Det är emellertid inte självklart att bara en grupp är diskriminerad. Cotton (1988) argumenterar för möjligheten att båda grupper skulle kunna vara diskriminerade i så måtto att den ena gruppens egenskaper är övervärderade medan den andra gruppens egenskaper är undervärderade och att det diskrimineringsneutrala läget skulle finnas någonstans där emellan. Detta resonemang leder till tre möjliga alternativ.

Det första alternativet bygger på det aritmetiska genomsnittet av kvinnors och mäns koefficienter för att på så sätt skatta en neutral effekt på lönen som ligger någonstans mitt emellan kvinnors och mäns priseffekter.

$$\beta^* = 0.5\beta_K + 0.5\beta_M \tag{7}$$

Ett annat alternativ som Cotton (1988) argumenterar för skulle kunna vara att använda ett viktat genomsnitt av de två gruppernas koefficienter. Det skulle i så fall innebära att:

$$\beta^* = \frac{n_K}{n_K + n_M}\beta_K + \frac{n_M}{n_K + n_M}\beta_M$$

Ett tredje alternativ, som dessutom är baserat på en teoretisk härledning (Neumark, 1988) är också baserad på ett viktat genomsnitt av gruppernas koefficienter. I det här fallet skapas vikterna med hjälp av poolad regression där koefficienterna för grupperna kvinnor och män skattas gemensamt.

Oavsett vilket mått på diskrimineringsneutral koefficient som används kräver införandet av denna koefficient att ekvation (3) modifieras. Modifieringen resulterar i följande uttryck:

$$\bar{Y}_M - \bar{Y}_F = \left(\bar{X}_M - \bar{X}_F\right)\beta^* + \bar{X}_K\left(\beta^* - \beta_K\right) + \bar{X}_M\left(\beta_M - \beta^*\right) \tag{3'}$$

Ekvation (3') är ekvivalent med ekvation (3) under villkoret att $\beta^* = \beta_M$ vilket gäller om mäns löner betraktas som norm, och kvinnors löner potentiellt är utsatta för någon typ av diskriminering.

Den ovan beskrivna dekomponeringsmetoden är mycket flexibel och kan på ett enkelt sätt utvidgas för att ta hänsyn till andra faktorer utöver det som finns observerat. Det skulle kunna handla om att tillåta för variation av koefficienternas storlek mellan olika branscher. Det skulle också kunna handla om skillnader i icke-observerbara faktorer mellan olika grupper. För att ta hänsyn till detta kan regressionsmodellen utökas med en selektionsjusteringsterm som i någon mening inkorporerar icke-observerbara faktorer och som därmed gör det möjligt att kontrollera för skillnader i det avseendet. För en beskrivning av dessa utvidgningar hänvisas till forskningslitteraturen (Se till exempel Yun, 2007).

3. Invariansproblemet

Oaxaca och Ransom (1999) diskuterar och redovisar ett viktigt problem vid användning av detaljerad dekomponering. Det visar sig att den oförklarade delen för grupper av dummyvariabler inte är invarianta med avseende på vilken referenskategori som används. Detta innebär till exempel att en viss region eller en viss bransch kommer att förklara olika stora delar av lönegapet beroende på vilken referenskategori som används. Eftersom referenskategorin i många fall kan vara godtyckligt vald är detta ett problem om syftet med analysen är att visa hur stor den oförklarade delen är med avseende på en viss faktor.

Identifierings- eller invaransproblemet illustreras bäst med hjälp av ett exempel. För tydlighetens skull används en enkel linjär regressionsmodell för att åskådliggöra var och varför problemet uppstår. Antag till exempel att vi analyserar lönegapet mellan kvinnor och män och att vi vill undersöka hur stor del av gapet som kan förklaras med en viss variabel, som är diskret (dichotom) och därför endast antar två värden (ett och noll).

$$X_1 = \begin{cases} 1 & \text{om ja} \\ 0 & \text{om nej} \end{cases}$$

$$1 - X_1 = X_2 = \begin{cases} 1 & \text{om nej} \\ 0 & \text{om ja} \end{cases}$$

Den diskreta variabeln representerar två utfall som inte specificeras explicit men som i den här framställningen betecknas med ett ”ja” eller ett ”nej”. Variabeln antas vara viktig när det gäller att förklara nivån på lönen för en viss individ. Dessutom skiljer sig variabelns effekt på lönen för kvinnor och män. Följande ekvationer kan därmed formulera för de två grupperna:

$$Y_M = \beta_{M0} + \beta_{M1} X_{M1} + U_M \tag{8}$$

$$Y_K = \beta_{K0} + \beta_{K1} X_{K1} + U_K\,. \tag{9}$$

Anledningen till att ekvationen skattas separat för kvinnor och för män beror på att man vill tillåta för möjligheten att koefficienterna skiljer sig mellan grupperna. Om så ej är fallet skulle det räcka med att skatta en ekvation, med en dummyvariabel för kvinna (eller man). Med hjälp av ekvationerna kan vi dela upp (dekomponera) skillnaden i genomsnittlig lön mellan kvinnor och män i en förklarad del som bygger på sammansättningsskillnader i observerbara lönepåverkande faktorer, och en oförklarad del som bygger på skillnader i koefficienter och konstanter. Detta kan formuleras på följande sätt:

$$\overline{Y}_M - \overline{Y}_K = \underbrace{\left(\overline{X}_{M1} - \overline{X}_{K1}\right)\beta_{M1}}_{\text{förklarad}} + \underbrace{\overline{X}_{K1}\left(\beta_{M1} - \beta_{K1}\right) + \left(\beta_{M0} - \beta_{K0}\right)}_{\text{oförklarad}}. \tag{10}$$

Hur förändras då den förklarade och den oförklarade delen när referensalternativet alterneras? För att åskådliggöra detta måste vi veta hur differensen mellan variablerna ändras, samt hur koefficienterna påverkas. Den förklarade delen består av två komponenter: $\left(\overline{X}_{M1} - \overline{X}_{K1}\right)$ och β_{M1}. Eftersom $X_1 = 1 - X_2$ och dessutom $\overline{X}_1 = 1 - \overline{X}_2$ så visar det sig att

$$\left(\overline{X}_{M1} - \overline{X}_{K1}\right) = -\left(\overline{X}_{M2} - \overline{X}_{K2}\right). \tag{11}$$

För att se vad som händer med koefficienterna när referenskategorin alterneras i modellen är det enklast att utgå från formeln för de skattade parametrarna som ges med hjälp av OLS. I detta fall har vi en ekvation för konstanten (12) och en ekvation för parametern relaterad till dummyvariabeln (13):

$$\beta_0 = \overline{Y} - \beta_1 \overline{X} \tag{12}$$

$$\beta_1 = \frac{\sum\left(X_i - \overline{X}\right)Y_i}{\sum\left(X_i - \overline{X}\right)^2} \tag{13}$$

Med hjälp av (13) tillsammans med lämpliga substitutioner erhålls följande resultat:

$$\beta_1 = \frac{\sum\left(X_{1i} - \overline{X}_1\right)Y_i}{\sum\left(X_{1i} - \overline{X}_1\right)^2} = \frac{\sum\left(1 - X_{2i} - (1 - \overline{X}_2)\right)Y_i}{\sum\left(1 - X_{2i} - (1 - \overline{X}_2)\right)^2} = \frac{-\sum\left(X_{2i} - \overline{X}_2\right)Y_i}{\sum\left(X_{2i} - \overline{X}_2\right)^2} = -\beta_2,$$

vilket innebär att

$$\left(\overline{X}_{M1} - \overline{X}_{K1}\right)\beta_{M1} = \left(\overline{X}_{M2} - \overline{X}_{K2}\right)\beta_{M2}. \tag{14}$$

Eftersom de två minustecknen tar ut varandra framkommer att det vänstra uttrycket är exakt lika med det högra. Det innebär att den förklarade delen är opåverkad av valet av referenskategori. Vi säger att den är invariant med avseende på val av referenskategori. Detta resonemang kan generaliseras till det multipla regressionsfallet med fler grupper av dummyvariabler.

Tyvärr gäller inte denna egenskap för den oförklarade delen som består av variabler, koefficienter och konstanter. Vad som händer när referensalternativ alterneras är att de inbördes andelsrelationerna ändras, men att summan av de oförklarade delarna är oförändrad. Eftersom det är de enskilda komponenterna i den oförklarade delen som ska redovisas vid detaljerad dekomponering blir detta ett problem. Med hjälp av beteckningarna ovan kan problemet illustreras på följande sätt:

För konstanten har vi: $\beta_0 = \overline{Y} - \beta_1 \overline{X}_1 = \overline{Y} - \beta_1(1 - \overline{X}_2) = \overline{Y} - \beta_2 \overline{X}_2 - \beta_1,$

vilket innebär att: $\overline{Y} - \beta_2 \overline{X}_2 = \beta_0 + \beta_1.$

Konstanten innan bytet av referensalternativ, ändras i det här fallet med storleken på koefficienten $(\beta_0 + \beta_1)$. Storleken på ändringen är beroende av hur och vilka variabler som finns med i modellen, men det faktum att konstanten ändras av att byta referensalternativ är viktigt att inse. Detta innebär nämligen att den oförklarade delen relaterad till skillnader i konstanter kommer att ändra sig godtyckligt beroende på val av referenskategori.

Den andra delen som bygger på skillnader i koefficienter är också påverkad. När referenskategorin alterneras kommer andelskomponenten att öka eller minska beroende på om andelen är större eller mindre än 0,5 (om $\overline{X}_{K1}$ är lika med 0,5 så sker av naturliga skäl ingen förändring). I det avseendet kommer $\overline{X}_{K1}(\beta_{M1} - \beta_{K1})$ att minska om andelen är större än 0,5. Differensen mellan koefficienterna kommer också att påverkas genom att byta tecken. Faktorns oförklarade delkomponenter kommer därför att vara beroende av vilken referenskategori som utesluts, vilket är ett problem om detta val är godtyckligt. Däremot kommer summan av samtliga oförklarade delkomponenter att vara opåverkad av valet av referenskategori. När den kategoriska variabeln innehåller flera alternativ kommer koefficienterna även att byta värde, vilket ytterligare bidrar till den godtyckliga förändringen i den oförklarade delen.

4. Lösningsmetoder

Inom regressionsanalysen används dummyvariabler när ekonomiska variabler är uttryckta i kvalitativa termer. Det skulle kunna handla om regionindelningar, branschindelningar, utbildningsinriktningar, eller utbildningsnivåer. När dummyvariabler ska inkluderas i en regressionsmodell kan samtliga kategorier inte inkluderas eftersom det skapar perfekt multikollinearitet mellan gruppen dummyvariabler och konstanten (interceptet) som normalt ingår i modellen. Detta medför att en normalisering måste införas för att kunna skatta regressionsmodellens koefficienter. Det kan ske på ett antal olika sätt:

Låt oss utgå från en situation där vi vill kontrollera för regional variation i löner. För detta ändamål skapas fyra regiondummyvariabler: norr (X_1), öst (X_2), syd (X_3) och väst (X_4) och följande modellspecifikation kan formuleras:

$$Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4 + U \tag{15}$$

För enkelhetens skull antas Y avse lön uttryckt i nivå. Koefficienterna i modellen anger hur mycket den genomsnittliga lönen i regionen avviker från populationens medelvärde. Eftersom $X_1 + X_2 + X_3 + X_4 = 1$ kan modellens koefficienter inte skattas med hjälp av OLS utan ytterligare restriktioner. Ett alternativ skulle kunna vara att utesluta konstanten, eftersom den är perfekt korrelerad med dummyvariablerna. Det ger följande specifikation:

$$Y = B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4 + U \tag{16}$$

Detta förfaringssätt är likställt med att ha fyra konstanter i modellen. När inga andra variabler ingår kommer parametrarna att var lika med den genomsnittliga lönen i respektive region. Om modellen utökas med fler variabler och ytterligare grupper av dummyvariabler så får konstanten en mer traditionell tolkning med den skillnaden att modellen har en konstant för respektive region. Dessutom kan inte motsvarande justering göras för ytterligare införda grupper av dummyvariabler eftersom de då skulle vara perfekt korrelerade med den första uppsättningen dummyvariabler. Av den anledningen måste andra metoder övervägas. Det vanligaste förfaringssättet är att normalisera en av koefficienterna till noll, vilket är ekvivalent med att utesluta en dummyvariabel.

$$Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + U \tag{17}$$

Med $B_4 = 0$ kommer de andra koefficienterna att vara relaterade till X_4 på så sätt att B_1 kommer att vara ett mått på hur mycket större eller mindre lönen är i region norr (X_1) i förhållande till region väst (X_4). På motsvarande sätt kommer också de övriga dummyvariablernas parametrar att vara relaterade till region väst. Observera att tolkningen av konstanten i (17) utgör medelvärdet för den uteslutna kategorin. Vid införande av fler variabler går det dock inte att göra den tolkningen av konstanten. Det som är viktigt att inse med denna ansats är att tolkningen av koefficienterna kommer att ändra sig beroende på vilken koefficient som normaliseras (eller variabel som utesluts). Det är detta som är kärnan till problem vid detaljerad dekomponeringsanalys, vilket innebär att andra former av normaliseringar måste övervägas.

En annan typ av normalisering som hanterar invaransproblemet sätter istället summan av dummyvariablernas koefficienter till noll: $B_1 + B_2 + B_3 + B_4 = 0$. Det kan enkelt åstadkommas med de flesta statistiska program. Ett alternativ till att använda OLS under bivillkor skulle kunna vara att substituera in restriktionen i regressionsmodellen och på så sätt frigöra modellen från en variabel (referenskategorin). Detta gör att OLS utan bivillkor kan användas varefter den tappade koefficienten härleds med de skattade parametrarna. Detta går till på följande sätt:

Metod 1: Normalisering av koefficienter

Steg 1: Välj referenskategori och bryt ut dess parameter från restriktionen. Om kategori 4 väljs erhålls:

$$B_4 = -B_1 - B_2 - B_3. \tag{18}$$

Steg 2: Substituera in B_4 i regressionsmodellen och samla ihop termerna.

$$Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4 + U$$

$$Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + (-B_1 - B_2 - B_3)X_4 + U$$

$$Y = B_0 + B_1(X_1 - X_4) + B_2(X_2 - X_4) + B_3(X_3 - X_4) + U \tag{19}$$

Steg 3: Skatta B_1, B_2 och B_3 med (19) och härled B_4 med hjälp av (18).

Det som är viktigt vid dekomponeringen är att inte använda de transformerade variablerna utan att X_1, X_2, X_3 och X_4 används tillsammans med de skattade parametrarna från de transformerade variablerna. Denna metod kan enkelt generaliseras till fall där fler grupper av dummyvariabler finns med i modellspecifikationen.

Ytterligare en typ av normalisering diskuteras i Suits (1984) och Kennedy (1986) och som gör att samtliga koefficienterna i (15) kan skattas. Metoden tillåter också för möjligheten att tillföra ytterligare variabler. Anta följande regressionsmodell:

$$Y = B_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + U \qquad (20)$$

Genom att välja en lämplig konstant K och använda egenskapen att $X_1 + X_2 + X_3 + X_4 = 1$, och normaliseringen att $B_4 = 0$, så kan (20) skrivas om på följande sätt:

$$\begin{aligned} Y &= B_0 - K + (B_1 + K)X_1 + (B_2 + K)X_2 + (B_3 + K)X_3 + (B_4 + K)X_4 + U \qquad (20') \\ Y &= B_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + K(X_1 + X_2 + X_3 + X_4) - K + U \\ Y &= B_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + U \end{aligned}$$

Med hjälp av lämpligt val av K kan samtliga koefficienter i regressionsmodellen skattas samtidigt som modellen har formen enligt ekvation (20). Frågan är hur K ska väljas? Två förlag finns i forskningslitteraturen.

Metod 2: Justeringsfaktormetod I, Yun (2005)

Steg 1: Skatta parametrarna i regressionsmodellen med vald referenskategori:

$$Y = B_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + U$$

Steg 2: Skapa justeringsfaktor K med hjälp av de skattade parametrarna:

$$K = -(B_1 + B_2 + B_3)/4$$

Steg 3: Härled invarianta koefficienter för respektive kategori:

$$\begin{aligned} \hat{B}_1 &= B_1 + K \\ \hat{B}_2 &= B_2 + K \\ \hat{B}_3 &= B_3 + K \\ \hat{B}_4 &= K \\ \hat{B}_0 &= B_0 - K \end{aligned}$$

Vid fler variabler i modellen kan tranformationen göras för respektive grupp av dummyvariabler med justering av konstanten med respektive justeringsfaktor från varje grupp av dummyvariabler.

Metod 3: Justeringsfaktormetod II, Kennedy (1986)

Den tredje och avslutande metoden som diskuteras i denna PM är en modifiering av metod 2. Kennedy (1986) noterar att den population som åsyftas i metod 2 är restriktiv och avviker från den som avses i normalfallet eftersom parametrarna i konstruktionen av K har fått lika stora vikter. För att råda bot på denna brist föreslås en justering av justeringsfaktorn under steg 2 på så sätt att vikterna för respektive parameter sätts till kategoriandelarna som gäller i populationen man vill göra inferens på. Metod 2 kommer att sammanfalla med metod 3 endast om kategoriandelarna är lika stora, vilket generellt sett torde vara en ovanlighet. I praktiken innebär detta att justeringsfaktorn definieras på följande sätt:

$$K = -\left(\beta_1 \overline{X}_1 + \beta_2 \overline{X}_2 + \beta_3 \overline{X}_3\right)$$

där vi vet att $\overline{X}_1 + \overline{X}_2 + \overline{X}_3 + \overline{X}_4 = 1$, och respektive andel representerar populationsandelarna. För att lättare förstå skillnaden mellan metod 2 och metod 3 följer ett numeriskt exempel. Låt oss anta att vi har en regressionsmodell och en kategorisk variabel med tre kategorier, vilket resulterar i tre dummyvariabler. Respektive kategori har följande medelvärden: $E[Y \mid X_1 = 1] = 1100$, $E[Y \mid X_2 = 1] = 1400$, och $E[Y \mid X_3 = 1] = 1800$ där andelarna för respektive kategori är ¼, ¼ och ½, vilket innebär ett totalt medelvärde på 1350. Resultaten från fyra olika specifikationer ser ut på följande sätt:

$$Y = 1100 + 300X_2 + 700X_3 \qquad \text{(Standard metod)}$$

$$Y = 1433 - 333X_1 - 33X_2 + 366X_3 \qquad \text{(Metod 1 \& Metod 2)}$$

$$Y = 1350 - 250X_1 + 50X_2 + 450X_3 \qquad \text{(Metod 3)}$$

Skattningarna med standardmetoden har en enkel tolkning. Konstanten representerar medelvärdet för den uteslutna kategorin, och koefficienterna representerar avvikelsen från den kategorin. Motsvarande situation har vi när metod 3 tillämpas, men där representerar konstanten populationens medelvärde och koefficienterna representerar avvikelser från populationens medelvärde. För metod 1 och 2 är situationen något mer komplicerad. Det är fortfarande så att koefficienterna representerar avvikelser från *ett* medelvärde vars nivå styrs av hur andelarna för de olika kategorierna ser ut. Av den anledningen blir tolkningen något svår, även om skattningarna är unika för respektive kategori. Det är detta problem som Kennedy (1986) framför och föreslår lösning på enligt metod 3. Givetvis kommer konstanten enligt metod 3 att ändra sig när fler variabler tillförs, men tolkningen av parametrarna förblir oförändrad.

Med hjälp av de ovan beskrivna metoderna för att hantera invariansproblemet kan nu dekomponeringsmetoden användas på vanligt sätt. Den skattade oförklarade delen för respektive kategori kommer att vara invariant med avseende på referensgrupp, och därmed vara identifierad på det sätt som beskrivs i forskningslitteraturen.

4.1 APPLIKATION OCH JÄMFÖRELSE AV RESULTAT

Metoderna som beskrevs i det föregående avsnittet skiljer sig något i sitt utförande. Det skulle kunna innebära att resultaten kommer att variera beroende på vilken metod som används. För att få en uppfattning om hur och i vilken omfattning resultaten skiljer sig mellan alternativen genomförs en övning med data från lönestrukturstatistiken. Övningen genomförs på lönegapet mellan kvinnor och män och variabelspecifikationen är identisk med den som användes i Lönebildningsrapporten 2008. Eftersom parametervärdena ändras kommer metoderna potentiellt även att få effekt på den förklarade delen, vilket innebär att det är intressant att även jämföra metodernas resultat med standardansatsen där referenskategorin inte ingår.

Tabell 1 Detaljerad dekomponering av könslönegap med olika metoder för justering av invariansproblemet för tjänstemän 2007

	Standard		Metod 1		Metod 2		Metod 3	
Variabler	För.	Oför.	För.	Oför.	För.	Oför.	För.	Oför.
Konstant	0	-0,0358	0	0,0447	0	0,0447	0	0,2266
Region	-0,0055	0,0186	-0,0055	0,0038	-0,0055	0,0038	-0,0055	-0,0055
Ålder	0,0124	0,0835	0,0124	0,0036	0,0124	0,0036	0,0124	-0,0124
Näringsgren	0,0238	-0,0098	0,0238	0,0077	0,0238	0,0077	0,0238	-0,0238
Födelseland	0,0013	-0,0005	0,0013	0,0028	0,0013	0,0028	0,0013	-0,0013
Utbildningsnivå	0,0013	0,0565	0,0013	0,0095	0,0013	0,0095	0,0013	-0,0013
Utbildningsinriktning	0,0056	-0,0252	0,0056	0,0157	0,0056	0,0157	0,0056	-0,0056
Arbetsplatsstorlek	-0,0031	0,0998	-0,0031	0,0105	-0,0031	0,0105	-0,0031	0,0031
Anställd med timlön	0,0024	-0,0013	0,0024	0,0052	0,0024	0,0052	0,0024	-0,0024
Yrke	0,0808	-0,0951	0,0808	-0,0134	0,0808	-0,0134	0,0808	-0,0808
Total	**0,1192**	**0,0904**	**0,1192**	**0,0904**	**0,1192**	**0,0904**	**0,1192**	**0,0904**

Anm. För. avser förklarad del av könslönegapet, medan Oför. avser oförklarad del av könslönegapet.

Källa: Lönestrukturstatistiken 2007 (Konjunkturinstitutet).

Tabell 1 presenterar resultat från fyra olika skattningar. Den sista raden i tabellen visar de aggregerade måtten för förklarad och oförklarad del av könslönegapet. Som framgår är dessa resultat identiska och därmed oberoende av vilken metod som används. De resultat som presenteras i tabellen är fortfrarande i viss grad aggregerad men med avseende på respektive faktor. Detta är intressant eftersom det gör det möjligt att se hur skattningarna skiljer sig mellan metoderna. På denna aggregeringsnivå är också den förklarade delen av den detaljerade dekomponeringen identisk mellan metoderna. Det som skiljer sig mellan metoderna ligger i den oförklarade delen. Att standardansatsen avviker från de övriga metoderna främgår med stor tydlighet, vilket också var förväntat. Dessutom skiljer sig också tecknet på skattningarna mellan standardansatsen och de övriga metoderna, vilket ytterligare illustrerar bristerna med standardansatsen.

Metod 1 och 2 genererar identiska resultat medan metod 3 genererar resultat som skiljer sig från de övriga två metoderna. Här visar det sig att införandet av viktade medelvärden av koefficienterna vid skapandet av justeringsfaktorn medför att den oförkarade delen blir spegelbilden av den förklarade delen, vilket innebär att de är lika stora i absoluta termer. Det är priset som kommer av att koefficienterna för respektive dummyvariabel representerar avvikelsen från populationsmedelvärdet snarare än ett medelvärde som avviker från detta.

Tabell 2 dissagregerar resultaten ytterligare och presenterar en dekomponeringen för respektive kategori och respektive faktor för de fyra metoderna. Skattningarna för metod 1 och 2 är fortfarande identiska. När det gäller metod 3 så finns det nu en viss heterogenitet mellan alternativen för respektive faktor. Men från Tabell 1 vet vi att dessa resultat har en inneboende normalisering som gör att summan av den förklarade delen och summan av den oförklarade delen för respektive faktor är lika stora i absoluta termer. Det förefaller därför som om samtliga metoder som presenterats här har brister, och att det av den anledningen är svårt att direkt säga vilken metod som är att föredra.

Fördelen med metod 1 och metod 2 är att de är mycket lättare att använda eftersom ansatsen i princip bygger på att man normaliserar summan av koefficienterna till noll. Däremot kommer dessa koefficienter att ha en otydlig tolkning eftersom de avser avvikelsen från något som inte direkt är definierat och som inte är lika med populationsmedelvärdet annat än om samtliga andelar är lika stora för kategorierna. Det sistnämnda problemet löses med metod 3, men resulterar i att storleken för den oförklarade och förklarade delen blir lika stora i absoluta termer när summering sker över kategorierna för respektive faktor.

Vi behöver veta på vilket sätt detta är ett problem för tolkningen av resultaten vi dekomponeringsanalysen. Svaret på denna fråga kommer att skilja sig beroende på vilken applikation som avses och vilka antaganden som har införts. När det gäller dekomponering av lönegap mellan kvinnor och män så skulle man kunna tänka sig några olika ansatser. För det första måste man ta ställning till vilket populationsmedelvärde som är relevant. När vi har två grupper skulle det kunna handla om ett populationsmedelvärde för kvinnor och ett för män, vilket innebär att koefficienterna inte blir jämförbara om populationsmedelvärdena skiljer sig mellan grupperna. Det är därför önskvärt att båda grupper refererar till samma populationsmedelvärde. Det skulle kunna handla om ett poolat populationsmedelvärde eller ett populationsmedelvärde för en av grupperna. Om man utgår från att en av grupperna fungerar som norm skulle det kunna vara motiverat att även ha denna grupps medelvärde som referenspopulation och analysera hur grupperna avviker från detta medelvärde. Ett mer neutralt alternativ skulle kunna vara att använda det poolade medelvärdet som referens för båda. Oavsett vilket populationsmedelvärde som används så är poängen att referenspopulationen måste vara samma för båda grupper för att det ska vara möjligt att ge skattningen på den oförklarade delen en rimlig tolkning. Om andelarna skiljer sig åt mellan grupperna och gruppernas populationsmedelvärde skiljer sig måste detta problem beaktas explicit om den beroende variabeln är uttryckt på nivå. Om den är logaritmerad mildras problemet något på grund av att det då handlar om hur stor den procentuella avvikelsen är från respektive grupps medelvärde. Det förefaller dock som att det fortfarande handlar om ett visst mått av godtycke och att det därför fortfarande finns utrymme för förbättringar.

Tabell 2 Detaljerad dekomponering för region, ålder och utbildningsinriktning

	Standard		Metod1		Metod2		Metod3	
Variabler	För.	Oför.	För.	Oför.	För.	Oför.	För.	Oför.
Regioner								
Stockholms län	-0,0070	0,0103	-0,0048	0,0049	-0,0048	0,0049	0,0008	-0,0351
Östra och mellansverige	0,0003	0,0018	-0,0000	-0,0000	-0,0001	-0,0000	-0,0010	-0,0142
Småland och öarna	0,0005	0,0016	-0,0002	0,0007	-0,0003	0,0007	-0,0023	-0,0062
Sydsverige	0,0001	0,0009	-0,0000	-0,0009	-0,0000	-0,0009	-0,0005	-0,0149
Västsverige	0,0006	0,0030	0,0000	0,0003	0,0001	0,0003	-0,0014	-0,0200
Norra mellansverige	-0,0001	0,0006	0,0002	-0,0003	0,0002	-0,0003	0,0011	-0,0079
Övre och mellersta Sverige	Ref	Ref	-0,0006	-0,0009	-0,0006	-0,0008	-0,0022	-0,0072
Ålder								
18-24	Ref	Ref	0,0071	-0,0051	0,0071	-0,0051	0,0077	-0,0061
25-29	-0,0011	0,0022	0,0023	-0,0059	0,0024	-0,0059	0,00267	-0,0076
30-34	0,0012	0,0084	-0,0002	-0,0019	-0,0002	-0,0018	-0,0004	-0,0039
35-39	-0,0000	0,0126	-0,0000	0,0009	-0,0001	0,0009	-0,0000	-0,0014
40-44	-0,0002	0,0160	-0,0000	0,0034	-0,0000	0,0034	-0,0000	0,0009
45-49	0,0020	0,0132	0,00058	0,0038	0,0006	0,0038	0,0004	0,0020
50-54	0,0026	0,0115	0,00067	0,0029	0,0007	0,0029	0,0005	0,0012
55-59	0,0038	0,0107	0,0009	0,0025	0,0009	0,0025	0,0007	0,0009
60-64	0,0041	0,0086	0,0011	0,0028	0,0010	0,0028	0,0008	0,0016
Utbildningsinriktning								
Allmän utbildning	Ref	Ref	-0,0022	0,0079	-0,0021	0,0079	-0,0012	0,00380
Pedagogisk och lärarutbildning	0,0047	-0,0029	0,0024	-0,0006	0,0024	-0,0006	0,0036	-0,0018
Humaniora och konst	0,0034	-0,0046	0,0018	-0,0023	0,0018	-0,0022	0,0026	-0,0035
Samhäll. jur. handel, adm.	0,0017	-0,0063	-0,0077	0,0080	-0,0077	0,0080	-0,0030	0,0006
Natur, matematik, data	-0,0007	-0,0021	0,0001	-0,0007	0,0001	-0,0007	-0,0004	-0,0014
Teknik och tillverkning	-0,0110	-0,0030	0,0118	0,0003	0,0118	0,0002	0,0003	-0,0014
Lant-o skogsbruk, djursjukvård	-0,0004	-0,0008	-0,0002	-0,0005	-0,0002	-0,0005	-0,0003	-0,0007
Hälsa, sjukvård, social omsorg	0,0073	-0,0043	-0,0002	0,0017	-0,0002	0,0017	0,0036	-0,0014
Tjänster	0,0008	-0,0004	-0,0003	0,0016	-0,0003	0,0017	0,0003	0,0006
Okänd	-0,0002	-0,0004	0,0001	0,0002	0,0000	0,0002	-0,0001	-0,0001

Anm. Värden är uttryckta i decimalform.

Källa: Lönestrukturstatistiken, 2007 (Konjunkturinstitutet).

5. Slutsatser

I denna PM presenteras tre ansatser för att hantera invariansproblemet vid detaljerad dekomponering. Metod 1 och 2 är förhållandevis lätta att tillämpa samtidigt som metoderna medför att dummyvariablernas koefficienter får en besvärlig tolkning. Detta kan jämföras med metod 3 som är något mer komplicerad men där koefficienterna får en mer begriplig tolkning som handlar om hur mycket varje kategori avviker från populationsmedelvärdet. Vid analys av lönegap mellan kvinnor och män delas urvalet upp i två delar, vilket gör att populationsmedelvärdet som erhålls kommer att avse respektive grupp, vilket innebär en jämförelsen med avseende på två olika medelvärden. Det har inte gjorts några försök till att anpassa ansatserna för att lösa detta problem. Att hitta en generell ansats för viktjustering förefaller dock inte vara helt trivialt. Resultaten antyder därmed att det för närvarande inte finns någon tillfredställande metod för att lösa invariansproblemet.

Referenser

Blinder, A. S. (1973) "Wage Discrimination: Reduced Form and Structural Estimates", Journal of Human Resources, 8(4), pp. 436-55.

Cotton, J. (1988), "On the Decomposition of Wage Differentials", The Review of Economics and Statistics, 70: 236-243.

Gardeazabal, J., and A. Ugidos (2005), "More on Identification in Detaile Wage Decomposition", Review of Economics and Statistics, 86(4), pp. 1034-36.

Gelbach, J. B. (2002), "Identified Heterogeneity in Detailed Wage Decomposition", Unpublished manuscript, Department of Economics, University of Maryland at College Park.

Kennedy, P. (1986), "Interpreting Dummy Variables", The Review of Economics and Statistics, 68: 174-175.

Oaxaca, R. L. (1973), "Male-female Wage Differentials in Urban Labor Markets", International Economic Review, 14(3), pp. 693-709.

Oaxaca, R. L., and M. R. Ransom (1999), "Identification in Detailed Wage Decomposition", Review of Economics and Statistics, 81(1), pp. 154-57.

Suits, D. B. (1984), "Dummy Variables: Mechanics vs. Interpretation", Review of Economics and statistics, 66(1), pp. 177-80.

Yun, M. (2007), "An extension of the Oaxaca decomposition using generalized residuals", Journal of Economic and Social Measurement, Vol 32, pp. 15-22.

Yun, M. (2005), "A Simple Solution to the Identification Problem in Detailed Wage Decomposition", Economic Inquiry, Vol. 43, No. 4, pp. 766-772.

Bilaga

SAMPLING VARIATION OCH STATISTISK INFERENS MED STANDARDANSATSEN

Dekomponering av ekonomiska variabler är intressant, men mer kan sägas om de kan relateras till populationsvärden. För att kunna göra det krävs ett mått på hur urvalsskattningarna varierar mellan olika urval. För detta ändamål behöver vi standardfel (medelfel) för skattningarna av den förklarade och icke-förklarade komponenten. För att härleda ett sådant mått behöver ställning tas till huruvida regressorerna kan betraktas som konstanter eller som stokastiska variabler i variansberäkningen. Vid antagande om konstanta regressorer i upprepade urval förenklas beräkningarna av variansen avsevärt. Men det sker på bekostnad av att samplingfelet blir större eftersom relevant variation inte finns med i beräkningen. Anta följande regressionmodell:

$$Y_i = X_i\beta_i + U_i \text{ där } E[U_i] = 0,\ i \in \{1,2\} \text{ och}$$
$$\overline{Y}_i = \overline{X}_i\beta_i\,,$$

där de två kategorierna 1 och 2 kan vara kvinnor och män. Om regressorernas medelvärden betraktas som konstanter kan variansen runt genomsnittsprediktionen skrivs på följande sätt:

$$V\left(\overline{X}\beta\right) = \overline{X}'V(\beta)\overline{X}\ .$$

Om däremot regressorernas medelvärden betraktas som stokastiska variabler kommer variansuttrycket att bli något mer komplicerat.

För att illustrera det utgår vi ifrån två godtyckliga stokastiska vektorer X_1 och X_2 som har varsin fördelning med ändliga medelvärden och varianser $(X_i \sim (\mu_i, \Sigma_i))$. Egenskaperna hos väntevärdesoperatorn medför att följande gäller: $E[X+Y]=E[X]+E[Y]$ och $E[XY]=E[X]E[Y]+Cov(X,Y)$ för två godtyckliga slumpvariabler. Motsvarande förfaringssätt kan användas för de två oberoende stokastiska vektorerna.

$$E[X_1^{'}X_2]=\mu_1^{'}\mu_2$$
$$E[X_iX_i^{'}]=\mu_i\mu_i^{'}+\Sigma_i, \quad i\in\{1,2\}$$

och

$$E\left[\left(X_1^{'}X_2\right)^2\right]=E\left[X_1^{'}X_2X_2^{'}X_1\right]=tr\left[E\left[X_1X_1^{'}X_2X_2\right]\right]=tr\left[E\left[X_1X_1^{'}\right]E\left[X_2X_2^{'}\right]\right]=$$

$$tr\left[\left(\mu_1\mu_1^{'}+\Sigma_1\right)\left(\mu_2\mu_2^{'}+\Sigma_2\right)\right]=tr\left[\mu_1\mu_1^{'}\mu_2\mu_2^{'}\right]+tr\left[\mu_1\mu_1^{'}\Sigma_2\right]+tr\left[\Sigma_1\mu_2\mu_2^{'}\right]+tr\left[\Sigma_1\Sigma_2\right]=$$

$$\left(\mu_1^{'}\mu_2\right)^2+\mu_1^{'}\Sigma_2\mu_1+\mu_2^{'}\Sigma_1\mu_2+tr\left[\Sigma_1\Sigma_2\right]$$

Detta innebär att variansen för produkten av två stokastiska vektorer kan formuleras på följande sätt:

$$V\left(X_1^{'}X_2\right)=E\left[\left(X_1^{'}X_2\right)^2\right]-\left[E\left(X_1^{'}{}'X_2\right)\right]^2=\mu_1^{'}\Sigma_2\mu_1+\mu_2^{'}\Sigma_1\mu_2+tr\left[\Sigma_1\Sigma_2\right]$$

Vi har nu det som behövs för att bestämma varianserna och motsvarande standardfel för total lönedifferens, förklarad lönedifferens och för oförklarade lönedifferensen. Utgå från följande uttryck: $Q=\overline{X}_1\beta_1-\overline{X}_2\beta_2$:

Varians för total lönedifferens:

$$V(Q)=V\left[\overline{X}_1\beta_1-\overline{X}_2\beta_2\right]=V\left[\overline{X}_1\beta_1\right]+V\left[\overline{X}_2\beta_2\right]=$$
$$\overline{X}_1V(\beta_1)\overline{X}_1+\beta_1V\left(\overline{X}_1\right)\beta_1+tr\left[V\left(\overline{X}_1\right)V(\beta_1)\right]+\overline{X}_2V(\beta_2)\overline{X}_2+\beta_2V\left(\overline{X}_2\right)\beta_2+tr\left[V\left(\overline{X}_2\right)V(\beta_2)\right]$$

Varians för förklarad lönedifferens:

$$V\left[\left(\overline{X}_1-\overline{X}_2\right)\beta_2\right]=$$
$$\left(\overline{X}_1-\overline{X}_2\right)V(\beta_2)\left(\overline{X}_1-\overline{X}_2\right)+\beta_1V\left(\overline{X}_1-\overline{X}_2\right)\beta_1+tr\left[V\left(\overline{X}_1-\overline{X}_2\right)V(\beta_1)\right]$$

Varians för oförklarad lönedifferens:

$$V\left[\overline{X}_2(\beta_1-\beta_2)\right]=$$
$$\left(\overline{X}_2\right)V(\beta_2)\left(\overline{X}_2\right)+(\beta_1-\beta_2)V\left(\overline{X}_2\right)(\beta_1-\beta_2)+tr\left[V(\beta_1-\beta_2)V\left(\overline{X}_2\right)\right]$$