

Resursfördelningsmodell

Göteborgs stad hösten 2019

Inledning

Underlag för analysen utgörs av ett register som innehåller elever som gått ut årskurs nio i en skola i Göteborgs stad våren 2015 – 2018. Registret innehåller uppgifter om skolenhet och examensår, information om elevens resultat och ett antal variabler som beskriver den socioekonomiska bakgrunden.

Den variabel som beskriver elevens resultat och används som beroende variabel i analysen är om eleven har uppnått behörighet till gymnasieskolan (yrkesprogram) eller inte. Valet av beroende variabel motiveras med att de elever som inte uppnått behörighet antas vara de elever som har större behov av stöd och stimulans.

De förklarande variabler som ingår i modellen är

- Kön
- Invandring
- Högsta utbildning för vårdnadshavarna
- Ekonomiskt bistånd
- Bor med en eller båda vårdnadshavarna
- Vårdnadshavarnas inkomst
- Utländsk bakgrund
- Human Development Index (HDI)
- Bostadsområdets sociala tyngd

Resultatet från modellen tillämpas därefter på ett datamaterial med elever i grundskola inklusive förskoleklass och förskola inklusive annan typ av verksamhet för barn 0-5 år i Göteborgs stad hösten 2018. Detta ger ett underlag som indikerar förväntade studieresultat för olika skolenheter givet elevernas socioekonomiska egenskaper.

Tillämpningen resulterar i ett index för var och en av skolenheterna som ingår i analysen. En skolenhet med en lägre andel elever som riskerar att inte klara målen än genomsnittet får ett index lägre än 100 och en skolenhet med en högre beräknad andel som riskerar att inte klara målen får ett index som är högre än 100. Indexet kan sedan användas som underlag för att fördela resurser mellan skolenheter i Göteborgs stad.

Variabler

Variabeln som beskriver elevens resultat anger om eleven nått eller ej nått behörighet till gymnasieskolan (yrkesprogram).

Bakgrundsvariablerna beskriver egenskaper hos eleven, elevens hushåll och elevens bostadsområde. Här ingår exempelvis kön, invandring, utbildningsnivå hos vårdnadshavare, uppgifter om ekonomiskt bistånd, inkomst och på vilket sätt eleven bor med vårdnadshavare. Detta är variabler som har visat sig vara viktiga

i tidigare analyser på riksnivå och för andra kommuner. Dessutom ingår bostadsområdets sociala tyngd.

Med vårdnadshavarnas inkomst avses förvärvsinkomst i det hushåll som eleven och minst en vårdnadshavare tillhör enligt folkbokföringen. Om ingen av vårdnadshavarna tillhör elevens hushåll, används hushållsinkomst för den vårdnadshavare som har högst hushållsinkomst. Hushållsinkomst kodas om till en kategorisk variabel (Låg, Medel, Hög) som beror på i vilket intervall hushållsinkomsten ligger. Intervallen har konstruerats så att ungefär 1/3 av eleverna i årskurs nio finns i kategorin ”låg”, 1/3 i kategorin ”medel” och 1/3 i kategorin ”hög”.

Bostadsområdets sociala tyngd: Bostadsområdena delas in i nivåer efter ”andelen elever med förväntad problemtyngd”. En regressionsmodell (bestående av variablerna kön, invandring + HDI hos ursprungsland, utbildning, ekonomiskt bistånd och boende) tillämpas på elever i Göteborgs stad (elever som gått ut årskurs nio i en skola i Göteborgs stad våren 2015 – 2018), så att varje elev får ett predikerat värde för sannolikheten att inte uppnå behörighet till gymnasieskolan. Därefter beräknas den genomsnittliga sannolikheten per bostadsområde.

Bostadsområdena klassas enligt nedan:

Tabell 1 Indelning av områden i det historiska materialet

<i>Andel elever med förväntad problemtyngd</i>	<i>Antal områden</i>	
<=11.0 procent	25	Boendemiljö 1
11.0 – 15.0 procent	25	Boendemiljö 2
15.0 – 22.1 procent	25	Boendemiljö 3
Över 22.1 procent	24	Boendemiljö 4

Eleverna med okända boendeområden antas inte komma från samma område utan ges medelvärdet 19,75. Dessa elever hamnar i boendemiljö 3.

Analysmetod

Regressionsanalys är en metod för att analysera samband mellan en beroende variabel och olika bakgrundsvariabler. En sådan analys ger underlag för att bedöma vilka variabler eller faktorer som är viktiga för att förklara den beroende variabeln och hur betydelsefulla faktorerna är i sig och i förhållande till varandra. Eftersom den beroende variabeln här endast kan anta två värden, används en variant av regressionsanalys - logistisk regression.

Vid logistisk regression kan regressionsmodellen beskrivas som

$$\log\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

där p är sannolikheten att den beroende variabeln antar ett visst värde och x_1, x_2, \dots, x_k utgör k stycken förklarande variabler. $\beta_1, \beta_2, \dots, \beta_k$ utgör regressionskoefficienterna och α utgör interceptet.

Koefficienterna anger betydelsen av de olika bakgrunds faktorerna. Dessa bakgrunds faktorer har kodats så att olika grupper lätt kan jämföras med varandra. En grupp är referensalternativet för varje variabel eller faktor. För kön, till exempel, är flickor referensgruppen med en koefficient som är 0. Koefficienten för pojkar anger hur pojkar skiljer sig från flickor i genomsnitt, när vi kontrollerar för de övriga variablerna som ingår i regressionsmodellen. En positiv koefficient innebär att pojkar inte lyckats lika bra som flickor.

I logistisk regression räknas regressionskoefficienterna ofta om till så kallade oddskvoter. En oddskvot för könen är, till exempel, kvoten mellan oddset för att inte klara målen för pojkar och oddset att inte klara målen för flickor. Oddset motsvarar sannolikheten att inte klara målen dividerad med sannolikheten att klara målen ($p/(1-p)$ där p är den skattade sannolikheten). En oddskvot större än 1 i exemplet anger då att pojkarna jämfört med flickorna har ett sämre utfall, en förhöjd risk att inte klara målen.

Anpassad modell för Göteborgs stad

Totalt ingår ca 20 300 elever i beräkningarna.

Tabell 2 Regressionskoefficienter, oddskvoter och p-värde¹

<i>Bakgrundsvariabel</i>	<i>Regr. koefficient</i>	<i>Oddskvot</i>	<i>p-värde</i>
Kön			
Pojkar	0,29	1,33	0,0001
Flickor	Referens	1	-
Invandring²			
För 0-6 år sedan, HDI grp1	2,63	13,87	<0,0001
För 0-6 år sedan, HDI grp2	2,03	7,63	<0,0001
För 0-6 år sedan, HDI grp3	1,39	4,03	<0,0001
För 0-6 år sedan, HDI grp4	1,57	4,79	<0,0001
Svensk bakgrund eller för mer än 6 år sedan	Referens	1	-
Utbildningsnivå			
Vårdnadshavare med grundskola eller lägre	1,19	3,29	<0,0001
Gymnasieskola	0,78	2,18	<0,0001
Eftergymnasial	Referens	1	-
Ekonomiskt bistånd			
Någon av vårdnadshavarna har ekonomiskt bistånd	0,30	1,36	<0,0001
Ingen av vårdnadshavarna har ekonomiskt bistånd	Referens	1	-
Boende			
Bor inte med båda vårdnadshavarna	0,30	1,35	<0,0001
Bor med båda vårdnadshavarna	Referens	1	-
Inkomst			
Låg	0,84	2,30	<0,0001
Medel	0,47	1,60	<0,0001
Hög	Referens	1	-
Utländsk bakgrund			
Inrikes född med minst en inrikes född förälder	Referens	1	-
Inrikes född med två utrikes födda föräldrar	0,32	1,38	<0,0001
Utrikes född	0,49	1,63	<0,0001
Boendemiljö			
Predikterat andel obehöriga i bostadsområdet	0,01	1,02	<0,0001

Percent concordant: 84,8³

¹ Med p-värde avses sannolikheten att observera ett minst lika extremt värde som det värde som observerats, givet att variabeln inte har någon effekt på den beroende variabeln. Ett lågt p-värde indikerar att det är låg sannolikhet att variabeln inte påverkar den beroende variabeln.

² HDI grupp 1 utgörs av invandring från länder med låg HDI, grupp 2 från länder med medel HDI, grupp 3 från länder med hög HDI och grupp 4 från länder med mycket hög HDI

³ Utifrån parvisa jämförelser beräknas hur ofta som modellen lyckas prediktera vilken person som inte uppnår behörighet till gymnasieskolan. Parvisa jämförelser görs mellan elever som uppnår behörighet till gymnasieskolan och elever som inte uppnår behörighet till gymnasieskolan.

Tillämpning av modeller för att skatta variationen i behov

Modellerna har tillämpats på elever som går i skolan i Göteborgs stad hösten 2019 (ca 89 500 elever), för att få en uppfattning om elevernas förväntade risk att inte nå målen. Vid tillämpningen används de skattade regressionskoefficienterna samt uppgifter om den socioekonomiska bakgrunden för en elev. Utifrån denna information beräknas den skattade sannolikheten för att en elev inte ska uppnå målen.

De skattade sannolikheterna på elevnivå summeras till skolnivå. För varje skolenhet skattas antal och andel elever som inte förväntas nå målen.

De skattade andelarna räknas om till ett index för varje skolenhet. Indexet kan sedan användas för att få fram hur behovet av stöd och stimulans varierar mellan kommunens skolenheter.

I materialet finns det 576 elever med okända (tillfälliga) personnummer. För dessa elever finns ingen eller få uppgifter om socioekonomisk bakgrund. Eleverna med okända personnummer har därför tilldelats följande värden på de förklarande variablerna

- Invandrad 0-6 år från land med låg HDI
- Gymnasial utbildning som högsta utbildning hos vårdnadshavarna
- Vårdnadshavarna får ej ekonomiskt bistånd
- Bor med båda vårdnadshavarna
- Vårdnadshavarna har låg inkomst
- Utrikes född
- Från ett bostadsområde med förväntad problemtyngd enligt medelvärde

Dessa antaganden innebär att elever med tillfälliga personnummer får relativt höga skattade sannolikheter att inte uppnå målen. Skolenheter med hög andel elever med tillfälliga personnummer får antagligen högre index.

För elever i Internationella skolenheter, 1 350 elever, antas följande där uppgifter om föräldrarnas utbildningsnivå är okänd (231 elever):

- Vårdnadshavare ges den vanligaste förekommande utbildningsnivån på enheten där eleven går.
- Vårdnadshavarna har inte ekonomiskt bistånd.
- Eleven bor med båda sina vårdnadshavare.
- Vårdnadshavare har inkomst i Kvartil 3.

Variabeln *bostadsområdets sociala tyngd* togs fram enligt samma förfarande som för elever som gått ut årskurs 9 våren 2015 - våren 2018. En standardmodell tillämpas på elever i Göteborgs stad hösten 2019, så att varje elev får ett predikerat värde för sannolikheten att inte uppnå behörighet till gymnasieskolan. Därefter beräknas den genomsnittliga sannolikheten per bostadsområde.

<i>Andel elever med förväntad problemtygnd</i>	<i>Antal områden i materialet avseende 2019</i>	
<=11.0 procent	74	Boendemiljö 1
11.0 – 15.0 procent	32	Boendemiljö 2
15.0 – 22.1 procent	27	Boendemiljö 3
Över 22.1 procent	26	Boendemiljö 4

Index tas fram enligt modellen som redovisas i tabell 2 i denna rapport.

När man använder indexen bör man beakta att de är modellskattningar och att elever i själva verket kan prestera annorlunda än vad modellen säger. Modellen säger hur en elev i genomsnitt borde prestera. Indexen ska ses som en vägledning om hur det kommer gå för eleverna i en skola. Särskild försiktighet bör iaktas om det handlar om små skolenheter.

Tabeller som levereras

Tabellerna innehåller följande variabler:

- Enhetsnamn
- Regi
- Skolform – Förskola eller Grundskola
- Antal - antal elever på skolenheten
- Skattat antal ej behöriga
- Skattad andel ej behöriga
- Index – index beräknat utifrån Skattad andel ej behöriga

Resultat för skolenheter med färre än 9 elever visas inte i tabellerna då resultaten anses alltför osäkra. (Det är många enheter som har just 9 elever, de får förbli oprickade men tänk på att modellen är osäkrare ju färre elever det är per enhet.)