Statistiska Centralbyrån
Avdelningen för social statistik och analys

Utvärdering av resursfördelningsmodell för fördelning av statsbidrag till huvudmän

Innehåll

1. Inledning

1.1. Bakgrund

På uppdrag av 2015 års Skolkommission tog Statistiska Centralbyrån (SCB) fram en statistisk modell som underlag till fördelning av statsbidrag till huvudmän som bedriver förskoleklass- och grundskoleverksamhet. Den nationella resursfördelningsmodellen har som syfte att stödja i resursfördelningen och ta hänsyn till strukturella skillnader i skolors förutsättningar och behov. Information om arbetet bakom framtagandet av modellen beskrivs i skolkommissionens slutbetänkande under avsnittet *Resursfördelningsmodell för huvudmän*.

Den grundläggande tanken bakom modellen är att elever som har hög sannolikhet (risk) att inte uppnå behörighet till studier på gymnasieskolan har ett större behov av stöd än elever som har en låg sannolikhet. Utifrån modellering baserat på historiska data kan dessa sannolikheter skattas, vilket i sin tur ger möjlighet till att beräkna ett index som speglar resursbehovet på huvudmannanivå. En huvudman med en lägre (större) andel elever som riskerar att inte uppnå gymnasiebehörighet än genomsnittet i riket får ett indexvärde mindre (större) än 100. Indexet kan därmed användas som underlag för att bedöma skillnader i huvudmäns behov av resurser.

1.2 Syfte och rapportens fokus

Regeringskansliet har gett SCB i uppdrag att utvärdera den nuvarande statistiska modellen i termer av hur väl den presterar som underlag till resursfördelningen. Den nuvarande modellansatsen baseras på ett antal, relativt grova socioekonomiska variabler. Syftet med ansatsen är att indikera ett resursbehov utifrån elevernas socioekonomiska egenskaper. Målet när modellen togs fram var inte enbart att producera träffsäkra prediktioner. Det finns antagligen flera faktorer som påverkar elevprestationer, vilka av olika skäl inte bör vara med i en resursfördelningsmodell. Se diskussionsavsnittet för mer om valet av förklarande variabler.

För att utvärdera modellen är det dock värdefullt att jämföra prediktioner med faktiska utfall. Det ger ett underlag för att belysa modellens styrkor och svagheter. Utvärderingen fokuserar därför på resursfördelningsmodellens prediktiva förmåga. En god prediktiv förmåga innebär att resursbehov för ett kommande år predikteras med hög träffsäkerhet, vilket är positivt för modellens tillförlighet som underlag i Regeringskansliets arbete.

1.3 Disposition

Denna rapport är disponerad på följande vis. I Avsnitt 2 ges en beskrivning av de dataunderlag och den statistiska modellen som

utvärderats i denna rapport. I Avsnitt 3 beskrivs de utvärderingsmått som tillämpats. I Avsnitt 4 presenteras resultaten av utvärderingen. Rapporten avslutas med en diskussion i Avsnitt 5.

2. Beskrivning av modell och data

2.1 Modell

Prediktioner av resursbehov tas fram på elevnivå via en flernivå-regressionsmodell. Den variant av flernivåmodell som används tar hänsyn till att datastrukturen är hierarkisk, där elever är grupperade (klustrade) i skolenheter.

Flernivåmodellen i fråga är en *random intercept model*. Om vi låter y_i vara en binär variabel som visar om en individ *i* har en viss egenskap ($y_i = 1$) eller inte ($y_i = 0$) kan sannolikheten att $y_i = 1$ givet en uppsättning förklarande variabler (samlade i listan $\boldsymbol{X}_i$) uttryckas som $\Pr(y_i = 1|\boldsymbol{X_i})$. Flernivåmodellen har formen

$$\text{logit}(\Pr(y_i = 1|\boldsymbol{X_i})) = \alpha_j + \beta_1 x_{1ij} + \beta_{2ij} + \cdots + \beta_k x_{kij}. \quad (1)$$

Koefficienten α_j är här ett skolspecifikt intercept, vilket tillåter interceptvärden att variera mellan skolenheter. Koefficienterna $\beta_1, \beta_2, \ldots, \beta_k$ utgör globala (samma över alla skolenheter) regressionskoefficienter som beskriver sambandet mellan de *k* förklarande variablerna $x_{1ij}, x_{2ij}, \ldots, x_{kij}$ i listan $\boldsymbol{X}_i$ för individ *i* på skolenhet *j* och utfallet $\text{logit}(\Pr(y_i = 1|\boldsymbol{X_i}))$. Utfallet kan även skrivas som $\log\left(\frac{Pr(y_i=1|\boldsymbol{X_i})}{1-Pr(y_i=1|\boldsymbol{X_i})}\right)$, vilket visar att tolkningar av samband görs utifrån log-oddskvoter.

Ekvation (1) kan transformeras till att i stället beskriva sannolikheten att $y_i = 1$ givet en uppsättning förklarande variabler:

$$Pr(y_i = 1|\boldsymbol{X}_i) = \frac{1}{1+e^{-(\alpha_j+\beta_1 x_{1ij}+\beta_{2ij}+\cdots+\beta_k x_{kij})}}. \quad (2)$$

Det är sannolikheten i Ekvation (2) som skattas (predikteras) och jämförs mot faktiska utfall i denna rapport. Det skolspecifika interceptet i Ekvation (1) kan delas upp i två delar som

$$\alpha_j = \alpha_g + \epsilon_j, \quad (3)$$

där α_g är ett genomsnittligt intercept avseende alla skolor och ϵ_j beskriver de skolspecifika skillnaderna i utfallsvariabeln som inte kan förklaras av $\boldsymbol{X}$-variablerna. Dessa skillnader kan bland annat bero på att vissa skolor är högpresterande och har högre kvalitet på sin undervisning. När prediktioner av sannolikheter görs används enbart det genomsnittliga interceptet α_g. Alltså utesluts skolspecifika skillnader som inte de förklarande variablerna kan förklara.

2.2 Prediktioner

Utifrån regressionsmodellen tas predikterade sannolikheter $Pr(\widehat{y_i = 1}|\boldsymbol{x}_i)$ fram för nya observationer. Prediktioner tas fram genom att först skatta (träna) den statistiska modellen på ett träningsdataset och erhålla skattade regressionskoefficienter, och sedan utgå från skattade regressionskoefficienter samt det genomsnittliga interceptet α_g i Ekvation (2).

De skattade sannolikheterna på elevnivå summeras upp till skolenhets- och huvudmannanivå och divideras med skolenhetens/huvudmannens elevantal för att få fram den skattade andelen elever som inte förväntas uppnå behörighet till gymnasieskolan. Prediktionerna utgör underlag för det behovsindex som beräknas till resursfördelningen.

2.3 Index

Behovsindexet beräknas genom att dividera den skattade andelen obehöriga elever med den skattade andelen obehöriga elever totalt sett i det aktuella dataunderlaget. Behovsindexet är med andra ord ett relativt index, vilket relaterar andelen obehöriga elever på en skolenhet eller huvudman till den genomsnittliga andelen över alla skolenheter/huvudmän för det aktuella året för statistikproduktion.

Ett indexvärde på 100 innebär att andelen obehöriga elever på skolenheten/huvudmannen är lika med den genomsnittliga andelen obehöriga elever. Har skolenheten/huvudmannen en högre (lägre) andel elever som inte förväntas få gymnasiebehörighet än den genomsnittliga andelen obehöriga elever tilldelas skolenheten/huvudmannen ett indexvärde större (mindre) än 100.

Behovsindexets konstruktion innebär att två variabler kan införa skillnader mellan skattade och faktiska andelar obehöriga elever. Dels kan skolenhetens/huvudmannens enskilda andel under- eller överskattas, dels kan nämnaren i indexberäkningen (den genomsnittliga andelen bland skolenheter/huvudmän) under- eller överskattas. Då den sistnämnda variabeln är ett genomsnitt är det troligt att den är mer robust än de enskilda andelsskattningarna.

2.4 Tränings- och testdata

Utvärderingen utgår från prediktioner på elever i testdata från perioden 2018–2021. Träningen (skattningen) av regressionsmodellen har följt samma upplägg som när modellen används i faktisk statistikproduktion.

Tabell 1 visar de dataunderlag som ingår i utvärderingen. För att träna (skatta) regressionsmodellen används data om elever som avgått grundskolan under vårterminen under en femårsperiod, exempelvis VT17-VT21. Elever som inte får betyg enligt det svenska betygssystemet ingår inte i tränings- eller testdata. Elever med

tillfälliga personnummer ingår inte i modelleringen utan hanteras separat vid indexberäkningar. Av detta skäl har elever med tillfälliga personnummer exkluderats från denna utvärdering som fokuserar på regressionsmodellens prediktionsförmåga.

För att utvärdera skattningar (testa) regressionsmodellen beräknas prediktioner på data om elever som går i grundskolan (årskurs 0 – årskurs 9) under höstterminen under träningsdatas sista år (HT21 i exemplet ovan). Endast elever som vid ett senare tillfälle (senast VT23) avgått grundskolan finns med i testdata så att data finns om huruvida eleven fått behörighet till gymnasiet. Behörighet i testdata har bestämts som den behörighet som först registrerats när eleven gått ut grundskolan.

Tabell 1. Dataset som ingår i utvärderingen.

Dataset	Period	Antal elever
Träningsdata	VT14 – VT18	504 916
Testdata	HT18	446 919
Träningsdata	VT15 – VT19	519 728
Testdata	HT19	341 501
Träningsdata	VT16 – VT20	536 922
Testdata	HT20	231 550
Träningsdata	VT17 – VT21	551 673
Testdata	HT21	118 348

Det bör påpekas att elever kan förekomma i flera av dataunderlagen. Totalt sett ingår 841 485 unika elever i träningsdata fördelat på 844 974 unika observationer (en elev kan exempelvis ha ändrade värden på bakgrundsvariabler mellan år). Motsvarande värde för testdata är 455 576 unika elever fördelat på 1 137 826 unika observationer. Dessa värden visar på den överlappning som finns i datamaterialet mellan de olika tidsperioderna.

2.5 Variabler i modellen

Antaluppgifter per bakgrundsvariabel

I Tabell 2 redovisas antaluppgifter för de ingående variablerna i modelleringen av behörighet. Från Tabell 2 kan det bland annat utläsas att utfallsvariabeln är obalanserad: det är flera elever i tränings- och testdata som fått behörighet till gymnasieskolan (cirka 85 procent i tränings- respektive testdata totalt sett, med variation mellan elevkategorier).

Datamaterialet består av följande uppgifter:

Behörighet till gymnasieskolans yrkesprogram (utfallsvariabel)
Denna uppgift avser om eleven uppnått behörighet till gymnasieskolan eller inte (behörighet till yrkesprogram). För att uppnå behörighet krävs godkända betyg i svenska eller svenska som andraspråk, engelska och matematik och i minst fem andra ämnen från grundskolan.

Kön
Elevens kön har hämtats från Skolverkets elevregister.

Vårdnadshavarnas utbildningsnivå
Uppgift om vårdnadshavarnas utbildningsnivå (vid tidpunkten då eleven avslutade grundskolan) har hämtats från Utbildningsregistret. Utbildningsnivån kodas i fyra kategorier (okänd, förgymnasial, gymnasial, eftergymnasial). En elev klassificeras därefter utifrån den vårdnadshavare som har högst utbildningsnivå.

Invandringsår
Det år då eleven invandrade till Sverige. Variabeln avser senaste invandringsår om eleven invandrat flera gånger. Variabeln har hämtats från Befolkningsregistret.

Ekonomiskt bistånd
Denna variabel identifierar om någon av vårdnadshavarna hade ekonomiskt bistånd under året före eleven gick ut årskurs 9. För elever som gick ut årskurs 9 år 2022 är det således bistånd under 2021 som räknas. Uppgiften har hämtats från Inkomst- och Taxeringsregistret.

Vårdnadshavares inkomstnivå
Uppgifter om vårdnadshavarnas disponibla inkomst hämtas från Inkomst- och Taxeringsregistret. Den disponibla inkomsten består av summan av faktorinkomster (löneinkomst, inkomst av näringsverksamhet och kapitalinkomst) samt skattepliktiga och skattefria transfereringar, minus skatt och övriga negativa transfereringar. Uppgiften avser året före eleven gick ut årskurs 9.

Eleven klassificeras utifrån den sammanlagda disponibla inkomsten i det hushåll där eleven och minst en av vårdnadshavarna är folkbokförda. Om ingen av vårdnadshavarna är folkbokförda i det hushåll där eleven är folkbokförd används inkomsten i det hushåll där inkomsten är störst.

Inkomsten klassificeras sedan i sex kategorier:

- Upp till 60 % av medianen
- Över 60 och upp till 80 % av medianen
- Över 80 % och upp till medianen
- Över medianen och upp till 120 % av medianen
- Över 120 % och upp till 150 % av medianen
- Över 150 % av medianen

Familj
Familj avser huruvida eleven är folkbokförd på samma adress som båda sina vårdnadshavare. Variabeln har hämtats från Befolkningsregistret.

Syskon
Uppgifter om antal syskon hämtas från befolkningsregistret. Uppgiften avser hemmaboende syskon och omfattar hel-, halv- och adoptivsyskon. Variabeln har hämtats från Befolkningsregistret.

Bostadsområdets socioekonomiska status
Det bostadsområde (DeSO-område) som eleven var folkbokförd i, det år eleven gick ut årskurs nio har klassats socioekonomiskt.

Vid klassningen används följande variabler:

- Andelen av befolkningen i bostadsområdet med utländsk bakgrund
- Befolkningens utbildningsnivå (genomsnittlig utbildningsnivå per person, 25–64 år)
- Befolkningens inkomst (median av disponibel inkomst enligt tidigare, 25–64 år)

Vilka variabler som ska användas för att klassa bostadsområdet socioekonomiskt är delvis subjektivt, men de variabler som valts bör ge en god bild av hur bostadsområdet ser ut socioekonomiskt.

En principalkomponentanalys har sedan tillämpats för att väga samman variablerna till ett socioekonomiskt mått. Principalkomponentanalysen resulterar i en linjärkombination av de ingående variablerna.

Bostadsområdena indelas sedan i fyra lika stora grupper efter vilka egenskaper de har på ovanstående variabler:

- Kategori 1: Den svagaste fjärdedelen av områdena
- Kategori 2: Den näst svagaste fjärdedelen av områdena
- Kategori 3: Den näst starkaste fjärdedelen av områdena
- Kategori 4: Den starkaste fjärdedelen av områdena
- Kategori 5: Uppgift saknas

Elevens skolkommun och skolenhetskod
Skolkommun avser den kommun som skolenheten är belägen i. Skolenheter är identifierade i datamaterialet via skolenhetskoder.

Tabell 2. Antaluppgifter per variabelkategori, tränings- och testdata.

Variabel	Träningsdata	Behöriga	Testdata	Behöriga
Behörighet				
Ja	1 813 183	1 813 183	986 101	986 101
Nej	300 056	0	151 725	0
Kön				
Pojke	1 091 495	920 794	584 581	501 291
Flicka	1 021 744	892 389	553 245	484 810
Vårdnadshavares utbildningsnivå				
Okänd	30 217	9 357	17 944	7 750
Förgymnasial	132 917	70 209	66 814	38 609
Gymnasial	773 427	637 421	377 171	307 475
Eftergymnasial	1 176 678	1 096 196	675 897	632 267
Invandringsår				
Inom tre år från grundskoleavgång	96 280	26 088	53 280	24 628
4–6 år	65 391	38 097	40 553	28 767
Mer än 6 år eller född i Sverige	1 951 568	1 748 998	1 043 993	932 706
Ekonomiskt bistånd				
Ja	160 779	85 349	80 587	46 870
Nej	1 952 460	1 727 834	1 057 239	939 231
Vårdnadshavares inkomstnivå				
Upp till 60 % av medianen	446 506	134 182	214 060	154 722
Över 60 och upp till 80 % av medianen	294 377	235 990	154 872	124 584
Över 80 % och upp till medianen	315 746	271 695	160 649	137 039
Över medianen och upp till 120 % av medianen	363 886	332 099	197 436	178 181
Över 120 % och upp till 150 % av medianen	357 152	336 961	209 210	196 885
Över 150 % av medianen	335 572	324 114	201 599	194 690
Familj				
Folkbokförd på samma adress som båda vårdnadshavarna	812 760	642 312	391 647	314 007
Ej folkbokförd på samma adress som båda vårdnadshavarna	1 300 479	1 180 871	746 179	672 094
Syskon				
0	128 253	94 109	62 024	52 503
1	799 787	733 395	451 339	414 269
2	625 211	554 313	339 133	300 285
3	284 014	233 158	148 125	120 714
4	135 522	103 805	68 546	51 980
5 eller fler	140 452	94 403	68 659	46 800
Områdeskategori				
Den svagaste fjärdedelen av områdena	474 374	348 122	247 947	187 906
Den näst svagaste fjärdedelen av områdena	489 737	411 839	253 953	213 239
Den näst starkaste fjärdedelen av områdena	525 904	468 721	282 093	250 604
Den starkaste fjärdedelen av områdena	617 081	583 482	353 572	334 186
Uppgift saknas	6 143	1 019	261	166

3. Utvärderingsstrategi

3.1 Elevnivå

På elevnivå ligger fokus på modellens förmåga att generera predikterade sannolikheter för obehörighet som är låga (höga) för elever som senare blir behöriga (obehöriga). Producerar modellen predikterade sannolikheter som inte överensstämmer med det faktiska utfallet är detta indikationer på att modellen inte fångar upp mönster tillräckligt (är osäker) i dataunderlaget.

För att utvärdera den prediktiva förmågan på elevnivå analyseras följande:

1. De ingående förklaringsvariablernas statistiska signifikans och regressionsmodellens robusthet över tid (utifrån värdet på skattade regressionskoefficienter)
2. Fördelningen av predikterade sannolikheter
3. C-indexvärden
4. Brier Score-värden

C-index

C-indexet anger sannolikheten att en slumpmässigt utvald obehörig elev har en högre predikterad sannolikhet för obehörighet än en slumpmässigt utvald behörig elev. C-indexet visar därmed på modellens förmåga att skilja mellan de två kategorierna hos utfallsvariabeln. C-indexet antar värden mellan 0,50 (modellen är inte bättre än slumpen på att skilja mellan obehöriga och behöriga elever) och 1,00 (modellen skiljer mellan obehöriga och behöriga elever perfekt). Värden över 0,70 indikerar tillfredsställande prediktiv förmåga i denna rapport.

Brier Score

Brier Score-värden mäter avståndet mellan predikterade sannolikheter och faktiska utfall. Brier Score-värdet beräknas som

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - o_i)^2,$$

där N är antalet observationer (totalt eller för en delgrupp), $\hat{p}_i$ är den predikterade sannolikheten för obehörighet för elev i och o_i är det faktiska utfallet (0=behörighet eller 1=obehörighet). Brier Score-värdet antar värden mellan 0 och 1, där ett värde som understiger 0,50 visar på en högre prediktiv förmåga än slumpen och där ett värde på 0 innebär att modellen tilldelar 100 procents sannolikhet för obehörighet till alla elever som sedan visar sig bli obehöriga i datamaterialet, och 0 procents sannolikhet för obehörighet till alla elever som sedan får behörighet.

När utfallsvariabler är obalanserade är det av intresse att undersöka Brier Score-värden i relation till så kallad klass-obalans. I denna

rapport beräknas Brier Score-värden för samtliga prediktioner och jämförs mot en ”naiv” modell som skulle prediktera alla elever som majoritetsklassen (d.v.s. som behöriga) med 100 procents sannolikhet. Brier Score-värden beräknas även uppdelat på obehöriga respektive behöriga elever i testdata för att ta hänsyn till klass-obalansen.

3.2 Skolenhets- och huvudmannanivå

På skolenhets- och huvudmannanivå ligger fokus på modellens förmåga att producera skattade andelar obehöriga elever som ligger nära de faktiska andelarna i datamaterialet. Ett litet avstånd mellan skattade och faktiska andelar obehöriga elever visar på god prediktiv förmåga. Avstånd mäts som skillnaden mellan skattad och faktisk andel obehöriga elever (i procentenheter) för en skolenhet eller huvudman. I likhet med hur statistiken levereras kommer endast skolenheter och huvudmän där elevantalet överstiger 5 att beaktas.

Den skattade andelen obehöriga elever beräknas som summan av predikterade sannolikheter att bli obehörig till gymnasiet dividerat med antalet elever på skolenheten eller huvudmannen i datamaterialet. Faktiska andelar obehöriga elever beräknas som summan av antalet elever som registrerats som obehöriga till gymnasiet dividerat med det totala antalet elever på skolenheten eller huvudmannen. Summeringar görs per testdataset (exempelvis HT20) för att efterlikna hur statistiken tas fram i den verkliga statistikproduktionen. Med andra ord behandlas prediktioner för samma skolenhet/huvudman separat per gång som skolenheten/huvudmannen förekommer i datamaterialet.

För att utvärdera den prediktiva förmågan på skolenhets- och huvudmannanivå analyseras följande:

1. Fördelningen av skillnader mellan skattade och faktiska andelar obehöriga elever
2. Korrelation
3. Bias
4. Mean Squared Error

Korrelation

För att mäta styrkan på det linjära sambandet mellan skattade och faktiska andelar obehöriga elever beräknas Pearsons korrelationskoefficient. Korrelationskoefficienten antar värden mellan -1 (perfekt negativt samband) och 1 (perfekt positivt samband), där höga positiva värden är önskvärda i detta fall för att påvisa att andelar som bygger på prediktioner överensstämmer i hög grad med faktiska andelar utifrån datamaterialet.

Korrelationskoefficienter analyseras utifrån ett gränsvärde på 0,70 och redovisas tillsammans med den statistiska signifikansen av ett hypotestest av nollhypotesen att inget linjärt samband finns.

Bias
Bias avser om skillnader mellan skattade och faktiska andelar obehöriga elever tenderar att vara negativa (systematiska underskattningar av faktiska andelar) eller positiva (systematiska överskattningar). För att undersöka bias redovisas den genomsnittliga skillnaden bland skolenheter/huvudmän tillsammans med den statistiska signifikansen av ett hypotestest av nollhypotesen att den genomsnittliga skillnaden är lika med noll. Bias-värden som överstiger gränsvärdet på 0,10 (10 procentenheter) anses värda att belysa i denna rapport.

Mean Squared Error (MSE)
Mean Squared Error mäter det genomsnittliga kvadrerade avståndet mellan skattade och faktiska andelar obehöriga elever. Detta mått antar värdet 0 om skattade och faktiska andelar överensstämmer perfekt och ger större vikt åt stora skillnader. För att sätta MSE-värden i ett sammanhang redovisas även MSE-värden från en modell som predikterar obehörighet slumpmässigt, där resursfördelningsmodellen presterar bättre än slumpen om dess MSE-värden understiger slumpmodellens MSE-värden.

3.3 Indexutfall
För att utvärdera skillnaden i indexvärden mellan att använda predikterade sannolikheter respektive faktiska utfall analyseras:

1. Fördelningen av skillnader
2. Korrelation, indexvärden
3. Korrelation, ranking
4. Bias, indexvärden
5. Bias, ranking

Utvärderingen av indexutfall fokuserar på huvudmannanivån. Då index beräknas årsvis i den faktiska statistikproduktionen följer denna rapport samma princip och presenterar värden årsvis. För att få en överblick beräknas även genomsnittliga värden över testdata-åren.

Korrelation, indexvärden
Det linjära sambandet mellan skattade och faktiska indexvärden undersöks med Pearsons korrelationskoefficient. Se beskrivningen i Avsnitt 3.2 för mer information om detta utvärderingsmått.

Korrelation, ranking
Beroende på om predikterade sannolikheter eller faktiska utfall används kan rankingen på behovsindexet för en huvudman variera. För att mäta styrkan på sambandet mellan en huvudmans ranking utifrån skattade respektive faktiska indexvärden beräknas Spearmans rangkorrelationskoefficient. I detta sammanhang mäter Spearmans korrelationskoefficient sambandet mellan hur huvudmän rangordnas när skattade respektive faktiska indexvärden används (huvudmannen

med högst indexvärde får rank 1, huvudmannen med näst högst indexvärde får rank 2, etc.).

Spearmans korrelationskoefficient har ett högt positivt värde om huvudmän har en liknande rangordning utifrån skattade och faktiska indexvärden. I denna rapport undersöks korrelationer utifrån ett gränsvärde på 0,70.

Bias, indexvärden
För att undersöka om det finns en systematisk över- eller underskattning av indexvärden beräknas den genomsnittliga skillnaden mellan skattade och faktiska indexvärden samt ranker (bias). Se beskrivningen i Avsnitt 3.2 för mer information om detta utvärderingsmått. Ett gränsvärde på 50 indexenheter används i denna rapport.

Bias, ranking
Se beskrivningen i Avsnitt 3.2 för mer information om detta utvärderingsmått. Ett gränsvärde på 50 ranker används i denna rapport.

4. Resultat

4.1 Elevnivå

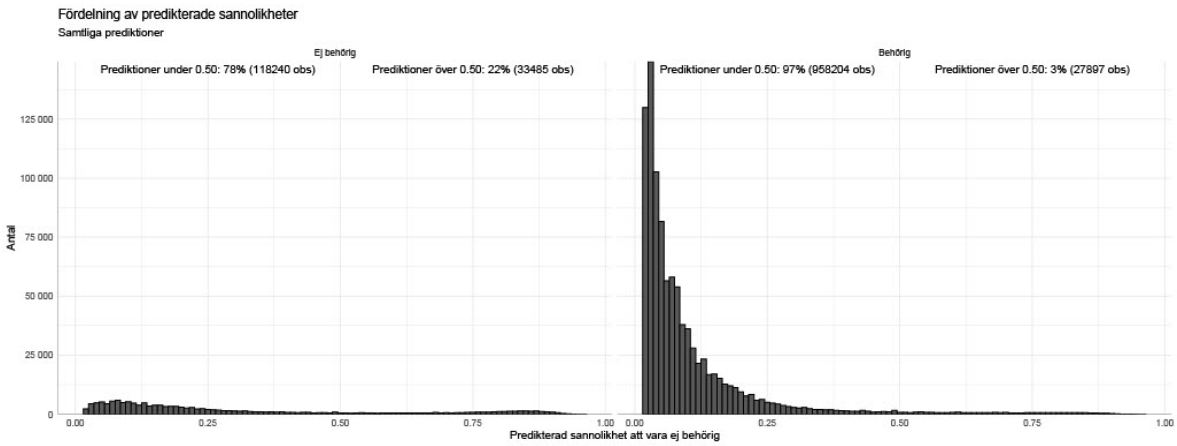
Modellens robusthet

Tabell 3 visar resultatet av modellträningen per träningsdataset. Genomsnittet av skattade *random effects* används som intercept, varför inget p-värde redovisas för denna koefficient. Tabell 3 visar stora likheter i skattade koefficientvärden mellan de olika träningsdataunderlagen. Samtliga skattade koefficienter är statistiskt signifikanta. Därmed visar Tabell 3 på en modellkonfiguration som är robust över tid och där ingående förklaringsvariabler har ett statistiskt signifikant samband med utfallet som studeras.

Predikterade sannolikheter – fördelning

Fördelningar av predikterade sannolikheter totalt sett för obehöriga respektive behöriga elever redovisas i Figur 1. I avsnittet ”Diagram och tabeller” bland rapportens bilagor redovisas motsvarande fördelningar uppdelat på bakgrundvariablers kategorier.

När det gäller predikterade sannolikheter för obehörighet för elever som senare fått behörighet förväntar vi oss att dessa ska vara låga i värde. Figur 1 överensstämmer i stort med dessa förväntningar, där 97 procent av prediktionerna har en predikterad sannolikhet på under 50 procent. När det gäller elever som inte fått behörighet visar Figur 1 att regressionsmodellen presterar sämre, där majoriteten av eleverna (78 procent) tilldelats predikterade sannolikheter för obehörighet under 50 procent. Fördelningarna utifrån förklaringsvariablernas kategorier visar i stort på samma mönster. Det är värt att notera att frånsteg från mönstret sker när antalet elever är jämnt fördelade mellan obehöriga och behöriga elever. I fall där det finns en obalans mellan obehöriga och behöriga elever tenderar fördelningarna att likna fördelningen för kategorin som har flest observationer.



Figur 1. Fördelning av predikterade sannolikheter, samtliga prediktioner.

*Tabell 3. Skattade regressionskoefficienter över tid. *** = statistiskt signifikant skattad koefficient på 5-procentsnivån.*

	Träningsdata			
Variabel	**2014–2018**	**2015–2019**	**2016–2020**	**2017–2021**
Intercept (genomsnitt)	-4,12 (-)	-4,08 (-)	-4,08 (-)	-4,03 (-)
Vårdnadshavares utbildningsnivå				
Förgymnasial	1,20***	1,19***	1,17***	1,18***
Gymnasial	0,79***	0,78***	0,77***	0,77***
Eftergymnasial	Referens	Referens	Referens	Referens
Okänd	1,15***	1,07***	1,03***	1,01***
Invandring				
Inom tre år från grundskoleavgång	2,42***	2,37***	2,33***	2,29***
4–6 år	0,93***	0,93***	0,96***	0,94***
Mer än 6 år eller född i Sverige	Referens	Referens	Referens	Referens
Vårdnadshavares inkomst				
Upp till 60 % av medianen	0,94***	0,97***	1,00***	1,01***
Över 60 och upp till 80 % av medianen	0,73***	0,76***	0,79***	0,80***
Över 80 % och upp till medianen	0,67***	0,69***	0,71***	0,70***
Över medianen och upp till 120 % av medianen	0,47***	0,47***	0,48***	0,47***
Över 120 % och upp till 150 % av medianen	0,24***	0,24***	0,26***	0,22***
Över 150 % av medianen	Referens	Referens	Referens	Referens
Kön				
Pojke	0,24***	0,23***	0,23***	0,21***
Flicka	Referens	Referens	Referens	Referens
Ekonomiskt bistånd				
Ja	0,51***	0,49***	0,48***	0,46***
Nej	Referens	Referens	Referens	Referens
Familj				
Folkbokförd på samma adress som båda vårdnadshavarna	Referens	Referens	Referens	Referens
Ej folkbokförd på samma adress som båda vårdnadshavarna	0,29***	0,29***	0,28***	0,28***
Antal syskon				
0	0,34***	0,35***	0,33***	0,29***
1	Referens	Referens	Referens	Referens
2	0,21***	0,23***	0,23***	0,24***
3	0,42***	0,43***	0,44***	0,45***
4	0,59***	0,60***	0,62***	0,61***
5 eller fler	0,82***	0,84***	0,85***	0,85***
Bostadsområdets socioekonomiska status				
Kategori 1 (svagt)	0,48***	0,47***	0,46***	0,46***
Kategori 2	0,37***	0,36***	0,35***	0,36***
Kategori 3	0,23***	0,24***	0,24***	0,26***
Kategori 4 (starkt)	Referens	Referens	Referens	Referens
Eleven har inte kunnat föras till något bostadsområde	1,47***	1,45***	1,48***	1,47***

C-index
I Tabell 4 redovisas C-indexvärden och Brier Score-värden för prediktioner på elevnivå. På totalnivå har C-indexet ett värde på 0,79, vilket visar på tillfredsställande prediktionsförmåga. För flera bakgrundskategorier överstiger C-indexet gränsvärdet även om det kan urskiljas att C-indexet precis understiger 0,70 för vissa fall. För elever som inte har kunnat föras till något bostadsområde är C-indexet lägst och understiger gränsvärdet.

Brier Score
Tabell 4 visar att regressionsmodellen presterar bättre än den ”naiva” modellen totalt sett samt för de flesta bakgrundsvariablerna. Det bör dock noteras att Brier Score-värdena i flera fall endast skiljer sig med någon decimal, bland annat totalt sett. Regressionsmodellen presterar på samma nivå som en ”naiv modell” när det gäller elever där föräldrarna har eftergymnasial utbildning samt elever i områden som är relativt socioekonomiskt starka. Dessa är kategorier där majoriteten av eleverna är att förväntas få behörighet till gymnasiestudier.

Samtidigt som modellen har en tillfredsställande prediktiv förmåga totalt sett utifrån C-indexvärdet har modellen svårigheter att korrekt tilldela höga sannolikheter för obehörighet till obehöriga elever. Detta framgår i Tabell 4 av Brier Score-värdet totalt sett för obehöriga elever, vilket överstiger gränsvärdet på 0,50. Samtliga Brier Score-värden för behöriga elever understiger gränsvärdet med flertalet Brier Score-värden för obehöriga elever visar på relativt låg prediktiv förmåga. Särskilt höga Brier Score-värden för obehöriga elever kan urskiljas för elevkategorier där majoriteten av elever kan förväntas få behörighet till gymnasiestudier, exempelvis elever från hushåll som inte mottar ekonomiskt bistånd och från hushåll i socioekonomiskt starka områden.

Tabell 4. Utvärderingsmått på elevnivå.

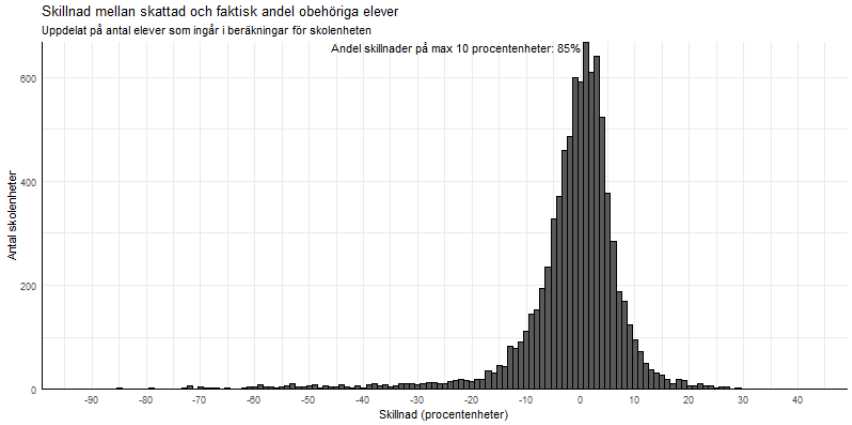
Variabel	C-index: Totalt sett	Brier Score	Brier Score: Naiv modell	Brier Score: Obehöriga elever	Brier Score: Behöriga elever
Samtliga observationer	0,79	0,10	0,13	**0,56**	0,03
Vårdnadshavares utbildningsnivå					
Förgymnasial	**0,67**	0,23	0,42	0,27	0,20
Gymnasial	**0,67**	0,14	0,19	**0,63**	0,03
Eftergymnasial	0,73	**0,06**	**0,06**	**0,77**	0,01
Okänd	**0,67**	0,25	0,57	0,07	0,48
Invandring					
Inom tre år från grundskoleavgång	**0,68**	0,26	0,54	0,07	0,47
4–6 år från grundskoleavgång	**0,69**	0,20	0,29	0,32	0,15
Mer än 6 år från grundskoleavgång eller född i Sverige	0,76	0,09	0,11	**0,72**	0,01
Vårdnadshavares inkomst					
Upp till 60 % av medianen	0,73	0,18	0,28	0,40	0,10
Över 60 och upp till 80 % av medianen	0,73	0,14	0,20	**0,52**	0,05
Över 80 % och upp till medianen	0,70	0,12	0,15	**0,65**	0,03
Över medianen och upp till 120 % av medianen	0,70	0,08	0,10	**0,77**	0,01
Över 120 % och upp till 150 % av medianen	0,72	0,05	0,10	**0,82**	0,004
Över 150 % av medianen	0,74	0,03	0,06	**0,87**	0,002
Kön					
Pojke	0,79	0,10	0,14	**0,54**	0,03
Flicka	0,79	0,09	0,12	**0,58**	0,02
Ekonomiskt bistånd					
Ja	**0,67**	0,24	0,42	0,25	0,22
Nej	0,77	0,09	0,11	**0,65**	0,02
Familj					
Folkbokförd på samma adress som båda vårdnadshavarna	0,79	0,08	0,10	**0,53**	0,05
Ej folkbokförd på samma adress som båda vårdnadshavarna	0,74	0,14	0,20	**0,60**	0,02
Antal syskon					
0	0,78	0,12	0,16	0,46	0,05
1	0,77	0,07	0,08	**0,70**	0,01
2	0,77	0,09	0,12	**0,63**	0,02
3	0,75	0,13	0,19	**0,53**	0,04
4	0,74	0,16	0,24	0,46	0,06
5 eller fler	0,70	0,20	0,32	0,39	0,11
Bostadsområdets socioekonomiska status					
Kategori 1 (svagt)	0,73	0,16	0,24	0,43	0,08
Kategori 2	0,75	0,12	0,16	**0,58**	0,03
Kategori 3	0,75	0,09	0,11	**0,68**	0,02
Kategori 4 (starkt)	0,76	**0,05**	**0,05**	**0,76**	0,01
Eleven har inte kunnat föras till något bostadsområde	**0,61**	0,27	0,36	0,18	0,32

4.2 Skolenhetsnivå

Skillnad i skattad och faktisk andel obehöriga elever – fördelning

Figur 2 visar fördelningen av skillnader (differenser) mellan skattade och faktiska andelar obehöriga elever per skolenhet. Figur 2 visar att skillnader är i stort symmetriskt fördelade runt nollvärdet, där 85 procent av skillnaderna är inom 10 procentenheter. Det förekommer fall i datamaterialet där den skattade andelen obehöriga elever noterbart skiljer sig mot det faktiska utfallet, särskilt där andelen obehöriga elever underskattas med över 40 procentenheter.

I avsnittet ”Diagram och tabeller” bland rapportens bilagor redovisas fördelningen av predikterade sannolikheter på skolenhetsnivå utifrån antalet elever som ingår i beräkningarna för skolenheterna samt hur stor andel av eleverna på skolenheten som är obehöriga i datamaterialet. Av dessa figurer framgår att en större spridning och förekomst av noterbara underskattningar främst förekommer bland mindre skolenheter (0–50 elever) och skolenheter med en relativt hög andel obehöriga elever (21+ procent). Exempelvis för små skolenheter är 68,6 procent av skillnaderna inom 10 procentenheter, att jämföra mot 97,1 procent för skolenheter med 201+ elever.



Figur 2. Skillnad mellan skattad och faktisk andel obehöriga elever, skolenhetsnivå.

Korrelation

I tabell 5 redovisas utvärderingsmått på skolenhetsnivå. Bland annat framgår att samtliga korrelationer är signifikant skilda från noll, samtidigt som flertalet av korrelationerna understiger gränsvärdet på 0,70, bland annat totalt sett. För samtliga skolenheter understiger korrelationen gränsvärdet på 0,70. Starka samband mellan skattade och faktiska andelar obehöriga elever kan utläsas för större skolenheter (101–200 elever och 200+ elever). Det bör noteras att korrelationen för skolenheter med 51–100 elever är precis under gränsvärdet.

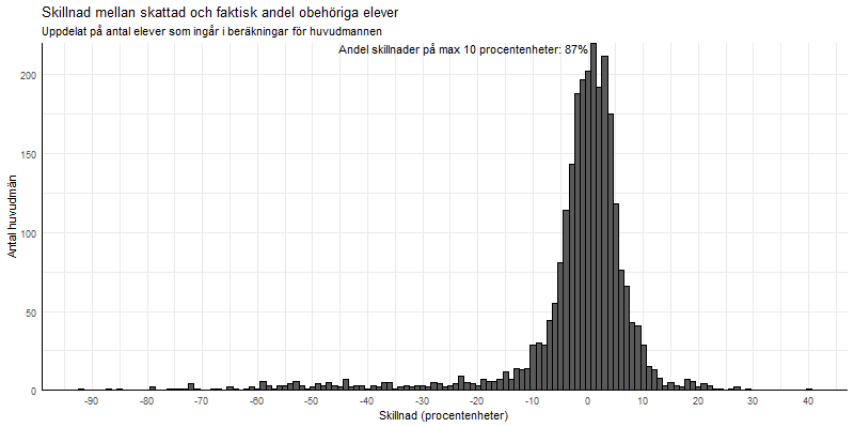
Bias
Tabell 5 visar att bias-värden är försumbara i nästintill samtliga fall. För skolenheter med en relativt hög andel obehöriga elever urskiljs en negativ bias, där faktiska andelar i genomsnitt underskattas med cirka 12 procentenheter.

*Tabell 5. Utvärderingsmått på skolenhetsnivå. *** = statistiskt signifikant korrelationsvärde (5 procents signifikansnivå).*

Variabel	Korrelation	Bias	MSE	MSE: slumpmodell
Samtliga observationer	**0,54*** **	-0,02***	0,01	0,03
Skolstorlek (elever)				
0–50	**0,37*** **	-0,05***	**0,03**	**0,03**
51–100	**0,67*** **	-0,006***	0,006	0,02
101–200	0,80***	-0,005***	0,003	0,02
201+	0,84***	-0,0004***	0,002	0,02
Andel obehöriga elever (%)				
0–10	**0,32*** **	0,04***	0,003	0,02
11–20	**0,31*** **	-0,004***	0,003	0,02
21+	**0,06*** **	**-0,12*** **	0,04	0,05

Mean Squared Error
Tabell 5 visar att regressionsmodellen presterar bättre än slumpen i nästintill samtliga fall. För små skolenheter (0–50 elever) har modellen samma prediktiva förmåga som en slumpmodell. Det bör noteras att MSE-värden skiljer sig med enbart någon decimal i vissa fall.

4.3 Huvudmannanivå
På huvudmannanivå kan ett liknande mönster som på skolenhetsnivå urskiljas, med undantag att regressionsmodellen presterar bättre än slumpmodellen även för mindre huvudmän. Fördelningar av predikterade sannolikheter på huvudmannanivå redovisas i Figur 3 samt i avsnittet ”Diagram och tabeller” bland rapportens bilagor. Utvärderingsmått på huvudmannanivå redovisas i Tabell 6.



Figur 3. Skillnader mellan skattad och faktisk andel obehöriga elever, huvudmannanivå.

Tabell 6. Utvärderingsmått på huvudmannanivå.

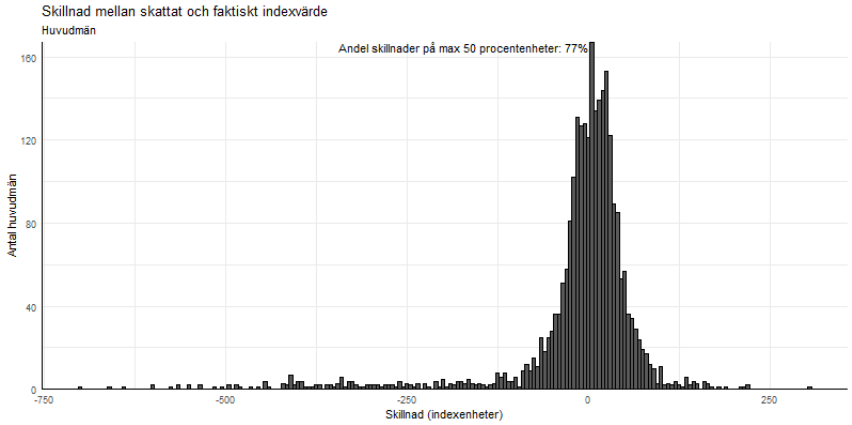
Variabel	Korrelation	Bias	MSE	MSE: slumpmodell
Samtliga observationer	**0,47*** **	-0,02***	0,02	0,08
Skolstorlek (elever)				
0–50	**0,39*** **	-0,07***	0,05	0,32
51–200	**0,57*** **	-0,002	0,01	0,01
201–500	0,81***	-0,001	0,002	0,02
501+	0,84	-0,006***	0,0008	0,02
Andel obehöriga elever (%)				
0–10	**0,35*** **	0,04***	0,003	0,03
11–20	**0,34*** **	-0,002	0,002	0,02
21+	**-0,03**	**-0,17*** **	**0,07**	0,05

4.4 Indexutfall

Skillnad mellan skattade och faktiska indexvärden – fördelning

Figur 4 visar fördelningen av skillnader (differenser) mellan skattade och faktiska indexvärden per huvudman. Av Figur 4 framgår att majoriteten av skillnaderna bildar en symmetrisk fördelning kring nollvärdet. Bland skillnaderna överstiger 23 procent av dem 50 indexenheter, där flera noterbara negativa skillnader förekommer.

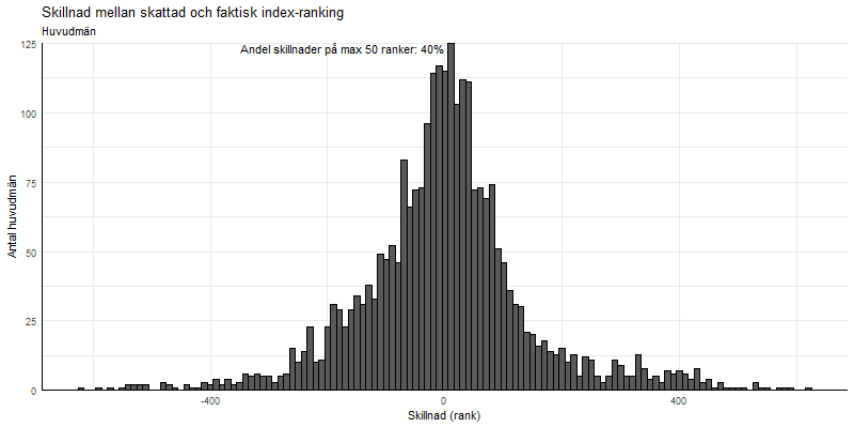
I avsnittet ”Diagram och tabeller” bland rapportens bilagor redovisas skillnader i indexvärden uppdelat på elevantal samt andel obehöriga elever. Dessa figurer visar att modellens träffsäkerhet ökar i samband med att elevantalet stiger samt när andelen obehöriga elever är relativt låg. Det framgår även att större negativa underskattningar förekommer främst bland huvudmän med relativt låga elevantal (0–50 elever) samt med höga andelar obehöriga elever (21+procent).



Figur 4. Skillnad mellan skattat och faktiskt indexvärde.

Skillnad i ranking – fördelning

I Figur 5 illustreras fördelningen av skillnader (differenser) mellan den rank (rank 1 = högsta indexvärde) som huvudmannen blivit tilldelad utifrån skattade och faktiska indexvärden. Figur 5 visar på att skillnaderna i rank är centrerade kring nollvärdet men även att majoriteten av skillnaderna (60 procent) överstiger 50 ranker. Tillsammans med rank-diagrammen i avsnittet ”Diagram och tabeller” bland rapportens bilagor syns att variationen i rank-skillnader är särskilt stor bland relativt små huvudmän (0–50 elever). För huvudmän med relativt låga (höga) andelar obehöriga elever tenderar huvudmannen att tilldelas ett högre (lägre) resursbehov än om faktiska utfall hade använts.



Figur 5. Skillnad mellan skattad och faktisk index-ranking.

Korrelation, indexvärden
I Tabell 7 redovisas utvärderingsmått för indexutfall. Tabell 7 visar att starka samband mellan skattade och faktiska indexvärden finns för stora huvudmän (201–500 samt 501+ elever). I flertalet fall uppvisar korrelationerna värden som understiger gränsvärdet på 0,70. Särskilt för huvudmän med höga andelar obehöriga elever (21+ procent) uppvisas svaga samband, där korrelationsvärden inte är statistiskt signifikant skilda från noll.

Korrelation, ranking
Tabell 7 visar att rangordningen av huvudmän överensstämmer i tillfredsställande grad totalt sett, och i synnerhet när elevantalet är högt. I likhet med resultaten för indexvärden är samband särskilt svaga för huvudmän med höga andelar obehöriga elever.

Bias, indexvärden
Tabell 7 visar att bias-värden är försumbara totalt sett. Regressionsmodellen presterar särskilt bra för huvudmän med höga elevantal. Noterbara underskattningar av faktiska indexvärden förekommer dock bland mindre huvudmän samt bland huvudmän med relativt höga andelar obehöriga elever. Även om bias-värden inte överstiger gränsvärdet för huvudmän med 0–10 procent obehöriga elever kan det noteras att faktiska indexvärden tenderar att överskattas för denna grupp samtidigt som indexvärden i genomsnitt underskattas för huvudmän med 21+ procent obehöriga elever.

Bias, ranking
När det gäller ranking-bias visar Tabell 7 att bias-värden generellt sett är försumbara. I likhet med resultaten för indexvärden visar bias-värden för ranker att resursbehovet för huvudmän med höga andelar obehöriga elever tenderar att underskattas medan behovet för huvudmän med 0–10 procent obehöriga elever i genomsnitt överskattas. För huvudmän med 0–10 procent obehöriga elever överstiger bias-värdet inte gränsvärdet.

Tabell 7. Utvärderingsmått, indexvärden och ranker.

Variabel		År	Korrelation: index	Korrelation: rank	Bias: index	Bias: rank
Samtliga observationer		*2018*	**0,46*** **	0,70***	-15,23***	0
		2019	**0,45*** **	0,71***	-16,16***	0
		2020	**0,48*** **	0,70***	-13,18***	0
		2021	**0,49*** **	**0,69*** **	-5,96	0
		Genomsnitt	**0,47**	0,70	-12,63	0
Skolstorlek (elever)	***0–50***	*2018*	**0,41*** **	**0,58*** **	**-65,11*** **	38,48***
		2019	**0,29*** **	**0,36*** **	**-106,01*** **	49,21***
		2020	**0,38*** **	**0,52*** **	**-52,79*** **	28,74
		2021	**0,41*** **	**0,61*** **	-12,64	0,37
		Genomsnitt	**0,37**	**0,52**	**-59,14**	29,20
	51–200	*2018*	**0,46*** **	0,71***	-0,02	2,13
		2019	**0,48*** **	0,75***	3,03	-5,23
		2020	**0,65*** **	0,80***	6,77	-17,10
		2021	**0,68*** **	0,75***	-0,37	-8,54
		Genomsnitt	**0,57**	0,75	2,35	-7,19
	201–500	*2018*	0,85***	0,81***	9,81***	-44,31***
		2019	0,83***	0,80***	2,33	-27,90***
		2020	0,80***	0,74***	-0,47	-18,84***
		2021	0,82***	0,80***	-3,03	10,19
		Genomsnitt	0,83	0,79	2,16	-20,22
	501+	*2018*	0,84***	0,82***	-2,84	-6,83
		2019	0,87***	0,87***	-2,99	1,65
		2020	0,85***	0,86***	-2,47	7,76
		2021	0,78***	0,81***	-1,41	15,67
		Genomsnitt	0,84	0,84	2,43	4,56
Andel obehöriga elever (%)	***0–10***	*2018*	**0,35*** **	**0,52*** **	28,23***	-49,77***
		2019	**0,51*** **	**0,61*** **	26,73***	-39,28***
		2020	**0,34*** **	**0,50*** **	29,67***	-36,97***
		2021	**0,26*** **	**0,42*** **	37,84***	-43,87***
		Genomsnitt	**0,37**	**0,51**	30,62	-42,47
	11–20	*2018*	**0,34*** **	**0,41*** **	-0,64	-15,23
		2019	**0,38*** **	**0,44*** **	1,39	-21,72***
		2020	**0,33*** **	**0,40*** **	3,39	-24,55***
		2021	**0,33*** **	**0,38*** **	5,07***	-12,87
		Genomsnitt	**0,35**	**0,41**	2,30	-18,59
	21+	*2018*	**-0,03**	**-0,09**	**-130,45*** **	**128,88*** **
		2019	**-0,06**	**-0,14**	**-135,02*** **	**121,34*** **
		2020	**-0,01**	**-0,03**	**-120,82*** **	**112,39*** **
		2021	**-0,04**	**-0,03**	**-90,45*** **	**87,31*** **
		Genomsnitt	**-0,04**	**-0,07**	**-119,19**	**112,48**

5. Sammanfattande diskussion

5.1 Modellens styrkor och svagheter

Modellens styrkor

Utifrån de analyser som genomförts dras slutsatsen att den nuvarande regressionsmodellen bygger på betydelsefulla förklaringsvariabler och producerar resultat som är robusta över tid. Regressionsmodellen bedöms ha en god prediktiv förmåga på elevnivå totalt sett. För skolenheter/huvudmän med höga elevantal eller en relativt låg andel elever som senare blir obehöriga bedöms modellen fungera väl.

Modellen producerar skattade indexvärden som totalt sett inte leder till noterbara över- eller underskattningar av faktiska indexvärden eller hur det relativa resursbehovet ser ut bland huvudmän. Däremot finns utvecklingsmöjligheter för modellen.

Svagheter

Resultaten i denna rapport visar att träffsäkerheten hos modellen kan problematiseras för flera elevkategorier, särskilt dem där majoriteten av eleverna kan förväntas få behörighet till yrkesprogrammet. Vidare visar analyserna på flera fall där samband mellan skattade och faktiska andelar samt skattade och faktiska indexvärden är svaga för relativt små skolenheter/huvudmän och för skolenheter/huvudmän med relativt höga andelar elever som senare blir obehöriga till yrkesprogrammet. För dessa skolenheter/huvudmän visar resultaten på en risk att resursbehov underskattas. Eftersom modellen har relativt låg träffsäkerhet för obehöriga elever görs bedömningen att modellen får betydande problem att prediktera elevresultat för skolor som har få elever eller hög andel obehöriga elever.

Över lag visar denna rapport på utvecklingsmöjligheter hos den nuvarande modellen. Det bör finnas möjligheter att se över både modellansats och förklarande variabler till modellen. Framför allt bör det undersökas om det finns möjligheter att förbättra förmågan att identifiera obehöriga elever.

Slutligen understryks vikten av att tolka den statistiska modellen utifrån ämneskännedom om vilka samband som den tar respektive inte tar hänsyn till. Avvikelser från modellens mönster är att förvänta. Det finns antagligen ett flertal faktorer som förklarar elevprestationer, vilka inte är lämpliga att ha med i en resursfördelningsmodell. Det är därför av stor vikt att utveckla en dynamisk modell, regelbundet utvärdera modellprestationer och tillämpa ett kritiskt förhållningssätt.

5.2 Utvecklingsområden i en modell-översyn

Val av inputs till modellen

Utifrån resultaten i denna rapport är rekommendationen att se över om fler förklarande variabler kan stärka modellens förklaringsförmåga. I den aktuella tillämpningen bör dessa variabler inte bara reflektera resursbehov men även ej oavsiktligt skapa incitament som kan leda till beteendeändringar hos skolor som är kontraproduktiva för utbildningsmål eller jämlikhet. Presumtiva förklarande variabler bör analyseras utifrån både ämneskunskap och statistiska egenskaper (så som statistisk signifikans). Genom att noggrant välja variabler som både bär förklaringskraft, har en meningsfull tolkning i sammanhanget och är rättvisa säkerställs att modellen tjänar sitt avsedda syfte med rättvis resursfördelning.

Det är även värt att se över hur inputs definieras och vilka kategorier som ska förekomma. Exempelvis skulle inkomstvariabler kunna vara kontinuerliga, i stället för indelade i inkomstkategorier, och därmed kunna bidra med mer information till modellen.

Hänsyn till skillnad i antalet behöriga och obehöriga elever

I dataunderlaget till denna rapport är det en majoritet av eleverna som blir behöriga till gymnasiet. Detta innebär att modellen tränas och optimeras utifrån en större mängd behöriga än obehöriga elever, vilket bedöms vara en orsak till att den nuvarande regressionsmodellen är sämre på att identifiera obehöriga elever för vissa bakgrundskategorier. En central utvecklingsmöjlighet i modelleringsförfarandet är därmed att ta hänsyn till fördelningen mellan behöriga och obehöriga elever i modellträningsfasen. Det finns flera sätt att göra detta på, exempelvis genom att tilldela olika vikter till elever i testdata från de olika kategorierna eller genom att använda enbart ett urval av eleverna från den större kategorin så att det finns en jämn fördelning av behöriga och obehöriga elever i det slutliga datamaterialet.

Val av modell

Resultaten i denna rapport visar på att modellen är relativt ”grov”. Då modellens syfte är inom prediktion skulle utveckling kunna ske genom att anpassa modellen utifrån ett maskininlärnings-ramverk.

Inom maskininlärning finns en stor mängd modeller var syfte är att klassificera observationer utifrån träning på kända utfall (*supervised learning*), där logistisk regression är ett exempel. En relativ fördel med logistisk regression är att regressionskoefficienter har en direkt tolkning. En relativ nackdel är att samband mellan inputs och utfallsvariabel måste definieras explicit samt att icke-linjära samband

inte fångas upp av modellen[1]. En utvecklingsmöjlighet skulle vara att testa flera prediktiva modeller i det aktuella sammanhanget.

Träning av modell
Prediktiva modeller inom maskininlärning brukar tränas genom att dela upp data i träningsdata, testdata (utvärderingsdata) samt prediktionsdata. I det nuvarande modelleringsförfarandet tränas och utvärderas resursfördelningsmodellen på samma dataunderlag. Detta kan medföra att utvärderingsmått visar överdrivet optimistiska värden samt innebär att underlag saknas för att rutinmässigt utvärdera modellens prestationsförmåga på nya observationer som modellen inte ”sett” i modellträningsfasen. Förslagsvis skulle det nuvarande modelleringsförfarandet utökas med så kallad korsvalidering, där träningsdata skulle kunna delas upp i exempelvis fem delar och där modellen i tur och ordning tränas på fyra av delarna och sedan testas (gör prediktioner för) den sista delen. Detta tillvägagångssätt brukar i litteraturen kallas för *K-fold cross-validation*.

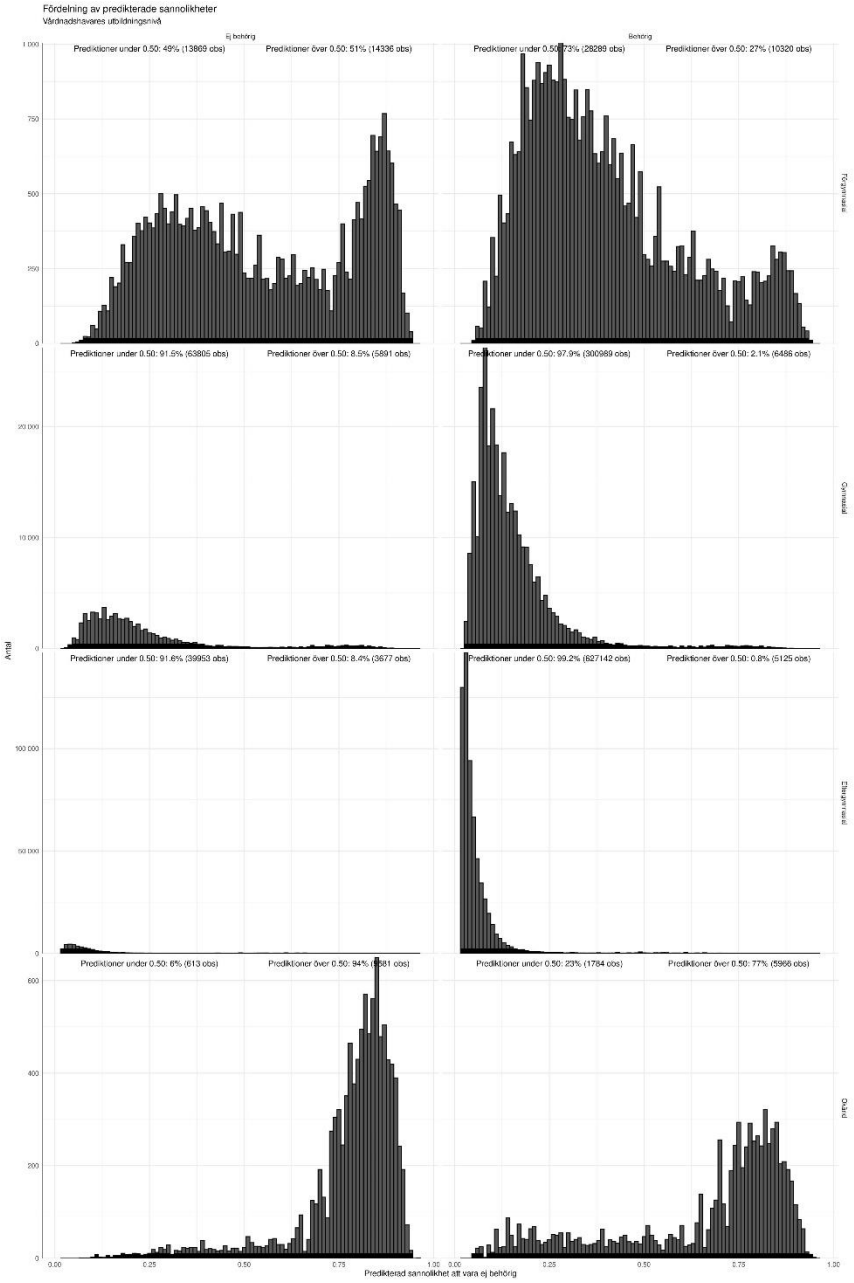
Osäkerhetsmått för indexvärden
Resultaten i denna utvärdering visar att prediktioners träffsäkerhet varierar mellan olika grupper av elever, skolenheter och huvudmän. Resultaten indikerar även att volatiliteten i skattade indexvärden bland annat varierar utifrån elevantalet. I dagsläget levereras skattade indexvärden som punktskattningar, där ingen redovisning görs av skattningarnas osäkerhet. Genom att ta fram osäkerhetsmarginaler (utifrån att skattade indexvärden bygger på modellbaserade elevprediktioner) för skattade indexvärden skulle tolkningen av enskilda indexvärdens tillförlitlighet underlättas.

Behandling av elever, skolenheter och huvudmän som kan förväntas vara olika resten av dataunderlaget
Det är även värt att överväga om flera grupper av elever, skolenheter eller huvudmän ska exkluderas från modellen och få modellberäknade värden på andra vis. Exempelvis förekommer resursskolor i träningsdata, vilka kan förväntas ha prestationer och behov som skiljer sig från majoriteten av skolenheter i datamaterialet. Det finns eventuellt en risk att dessa enheter även stör prediktioner och indexberäkningar för övriga enheter. I en eventuell översyn bör man titta närmre på hur dessa enheter kan hanteras.

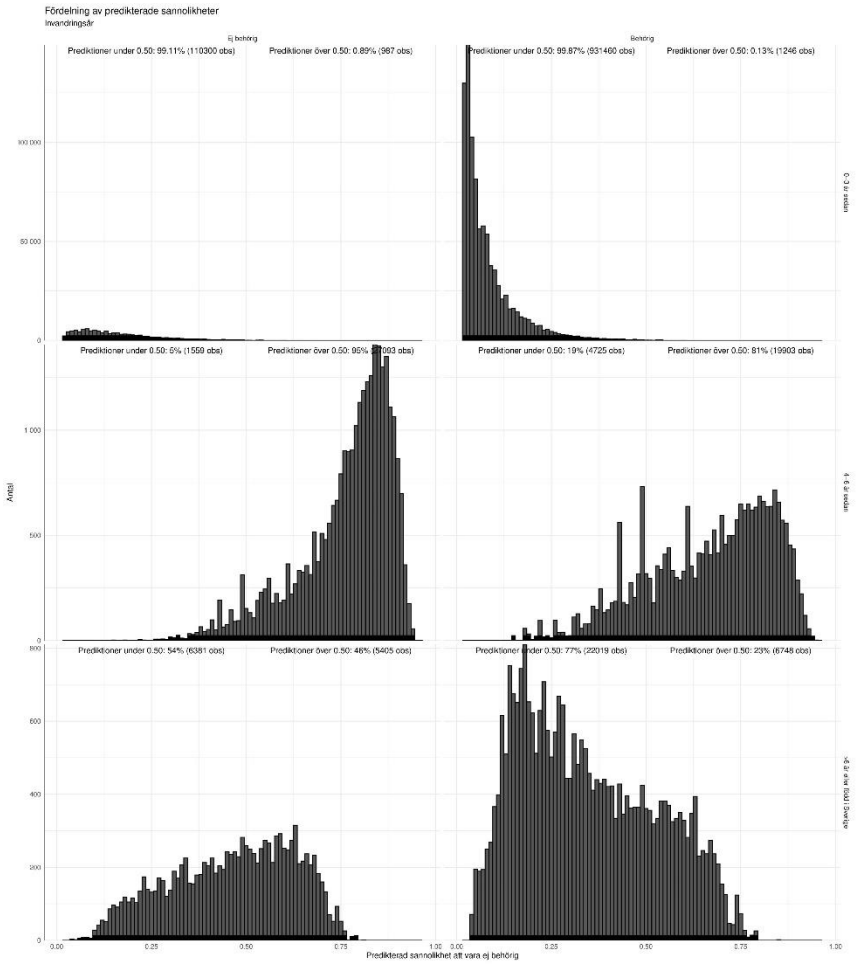
[1] Exempelvis har en enklare träning av en så kallad *Random Forest* testats inom ramen för denna utvärdering, med lovande resultat.

Bilagor

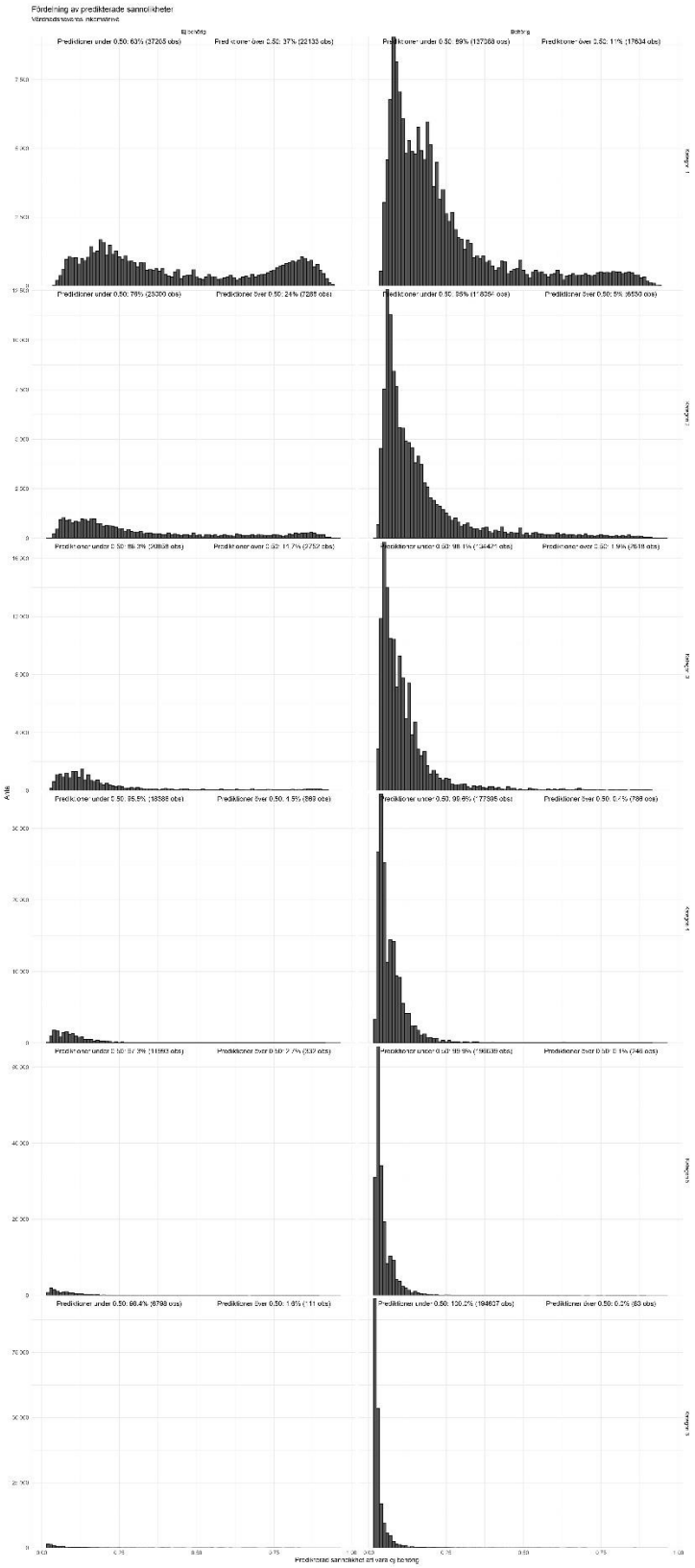
Diagram och tabeller: elevnivå



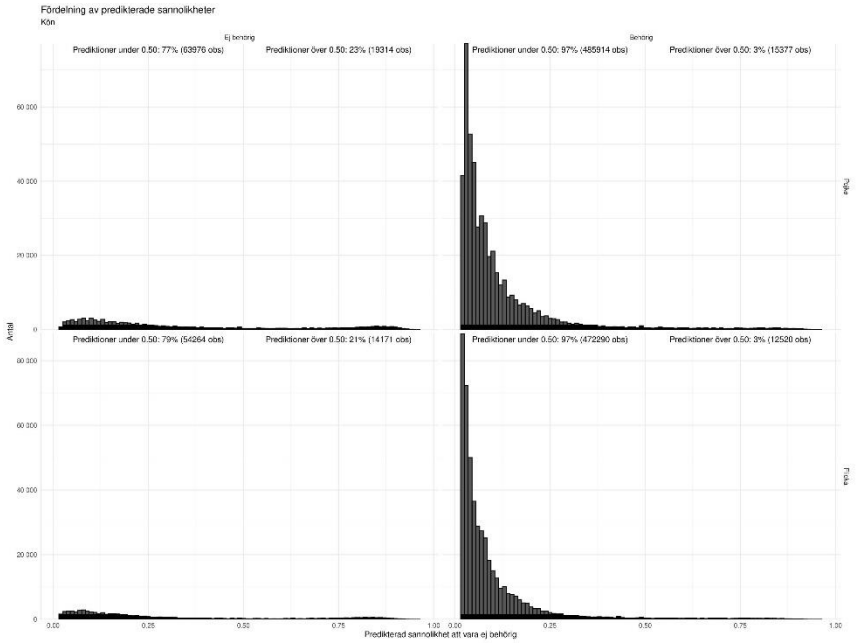
Figur 6. Fördelning av predikterade sannolikheter, vårdnadshavares utbildningsnivå.



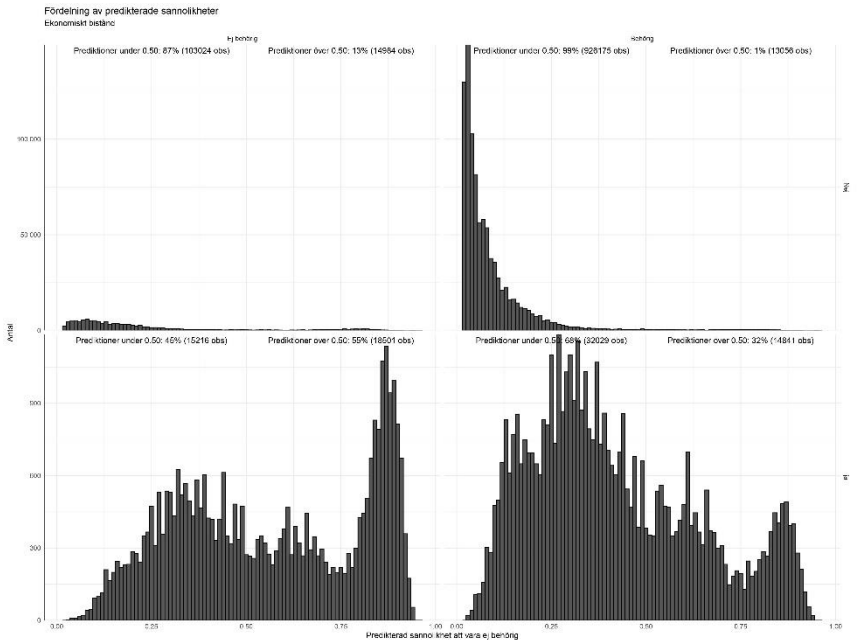
Figur 7. Fördelning av predikterade sannolikheter, invandringsår.



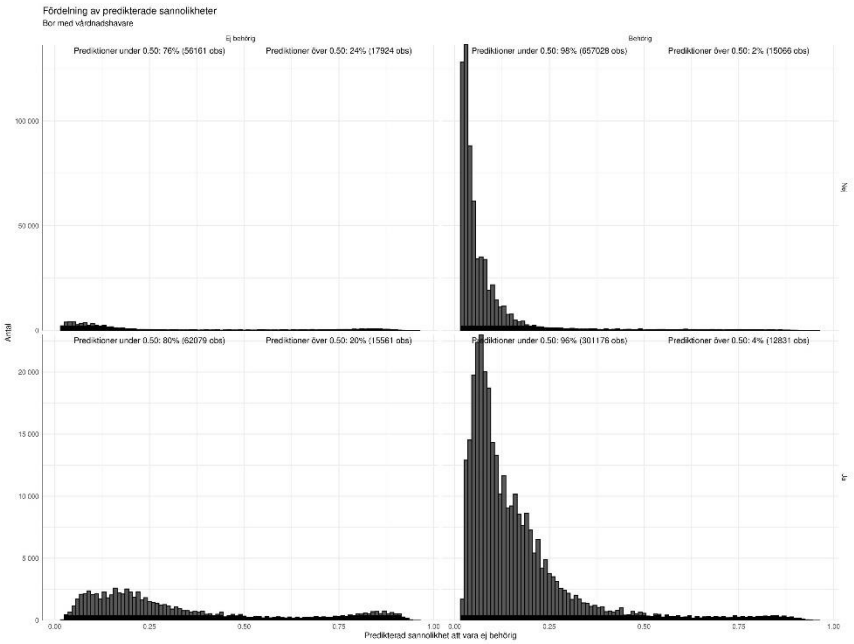
Figur 8. Fördelning av predikterade sannolikheter, vårdnadshavares inkomstnivå.



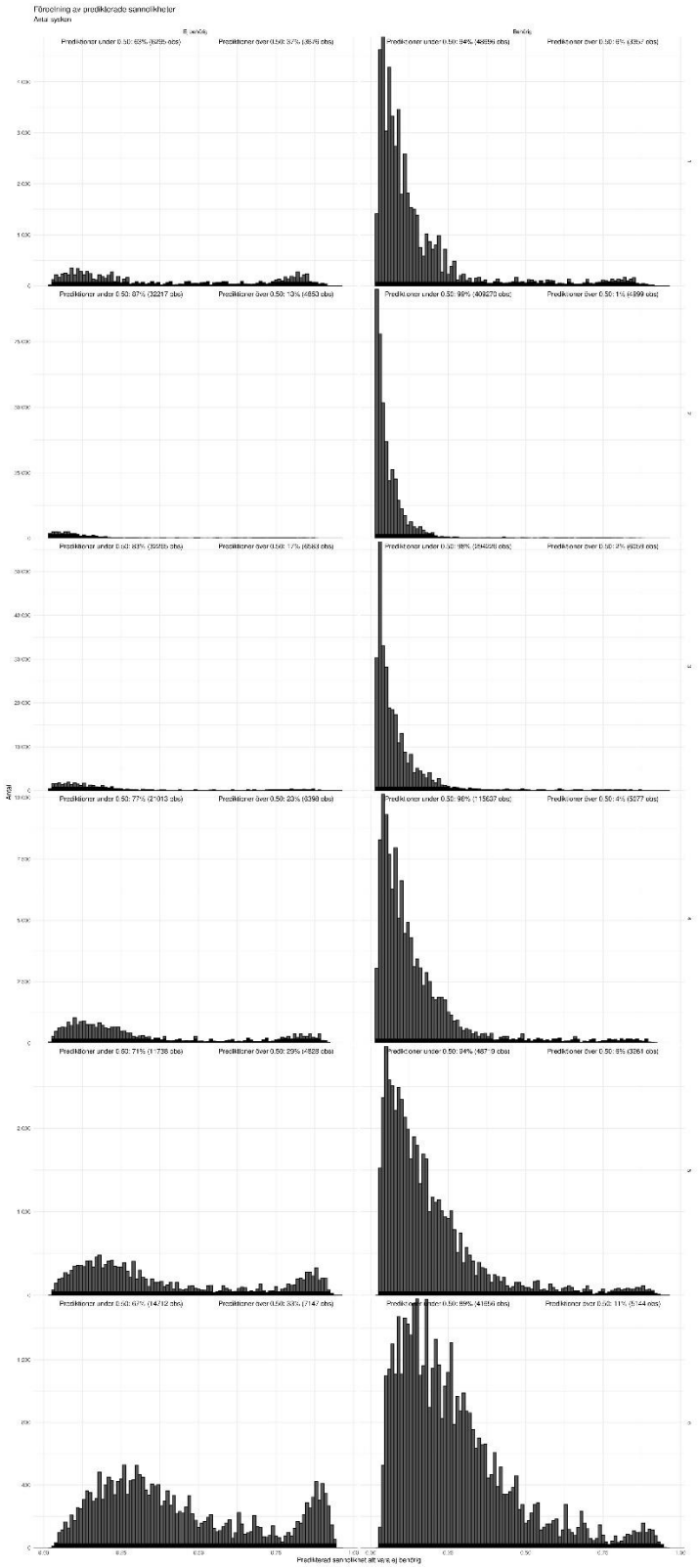
Figur 9. Fördelning av predikterade sannolikheter, kön.



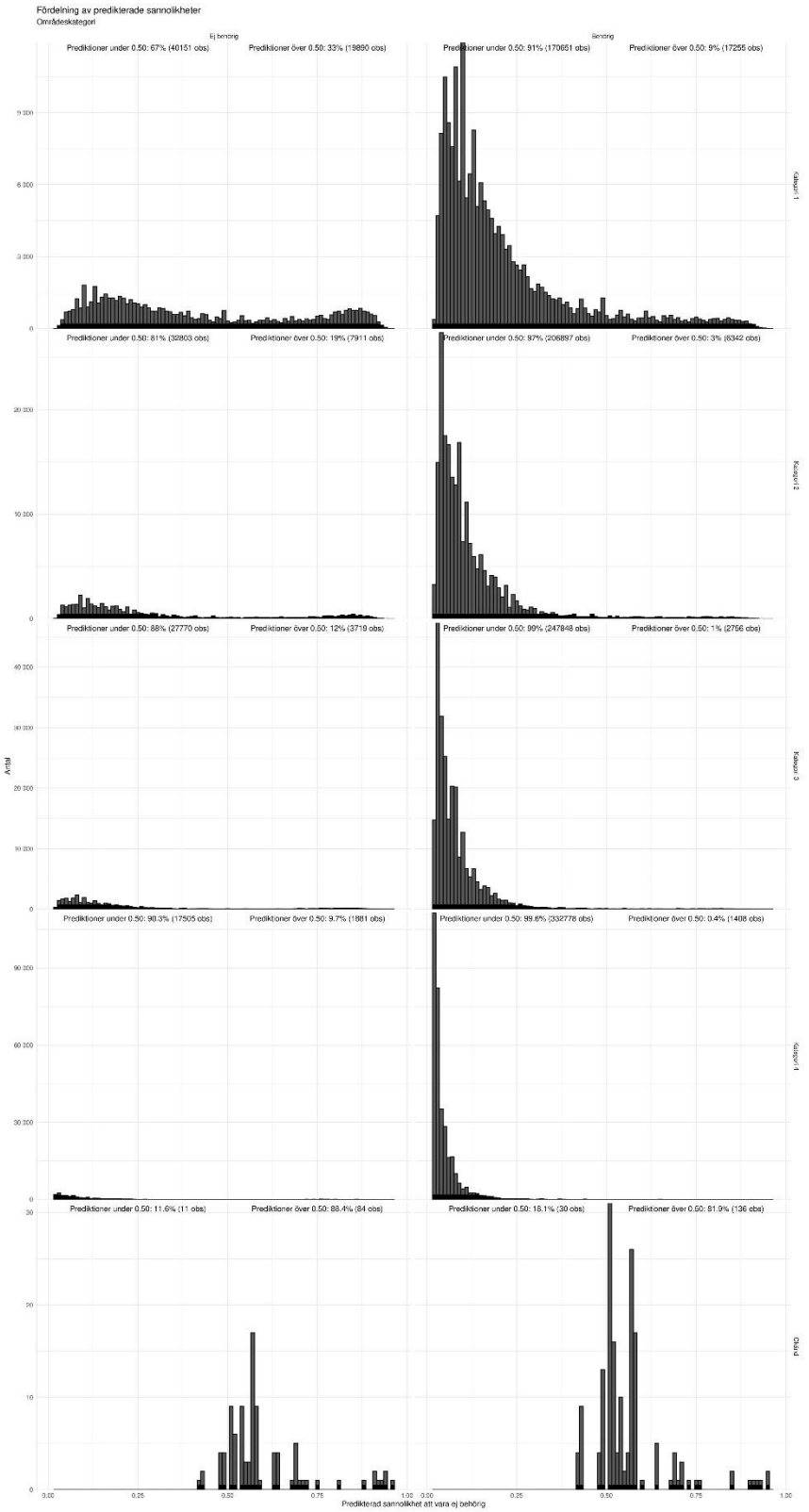
Figur 10. Fördelning av predikterade sannolikheter, ekonomiskt bistånd.



Figur 11. Fördelning av predikterade sannolikheter, bor med vårdnadshavare.

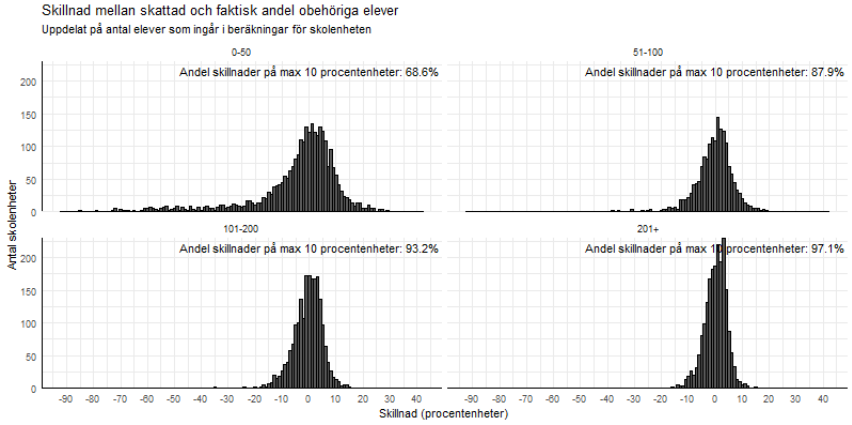


Figur 12. Fördelning av predikterade sannolikheter, Antal syskon. Här är kategori 1 = noll syskon och kategori 6 = 5 syskon eller fler.

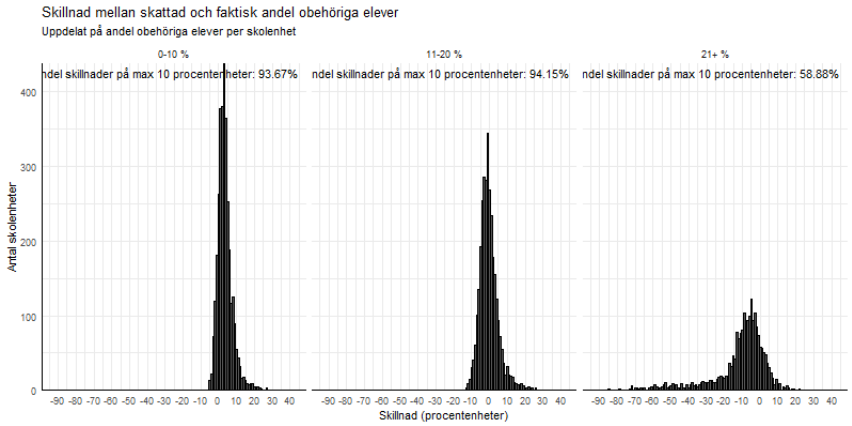


Figur 13. Fördelning av predikterade sannolikheter, områdeskategori.

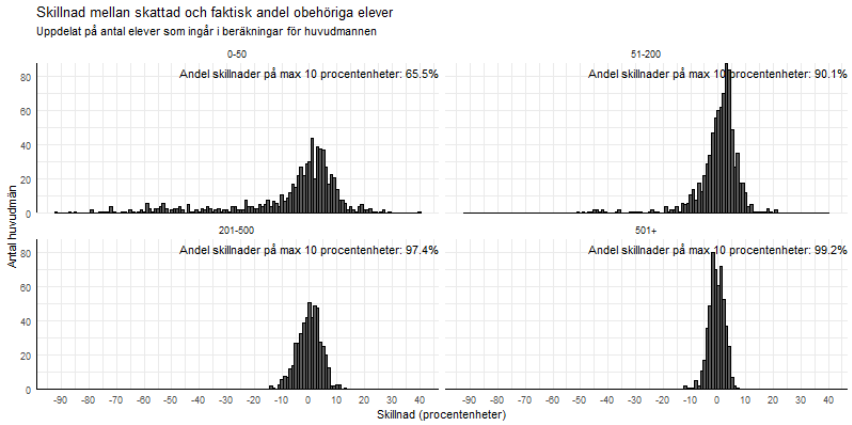
Diagram och tabeller: Skolenhets- och huvudmannanivå



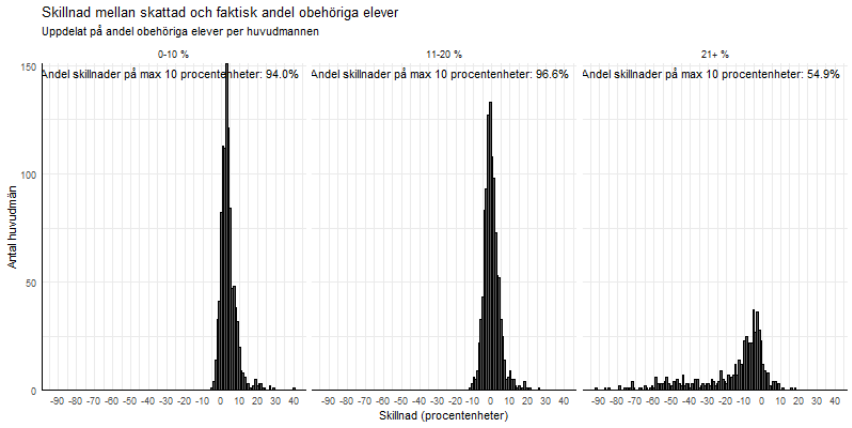
Figur 14. Skillnad i skattad och faktisk andel obehöriga elever på skolenhetsnivå, per skolstorlek.



Figur 15. Skillnad i skattad och faktisk andel obehöriga elever på skolenhetsnivå, utifrån andelen obehöriga elever på skolenheten.

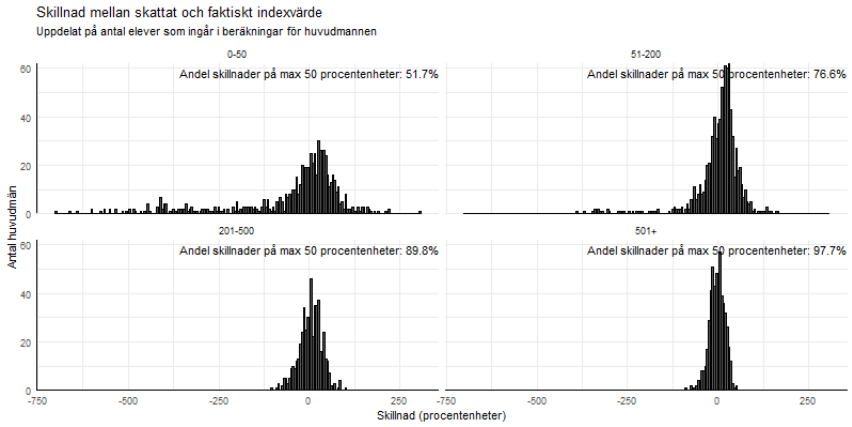


Figur 16. Skillnad i skattad och faktisk andel obehöriga elever på huvudmannanivå, per skolstorlek.

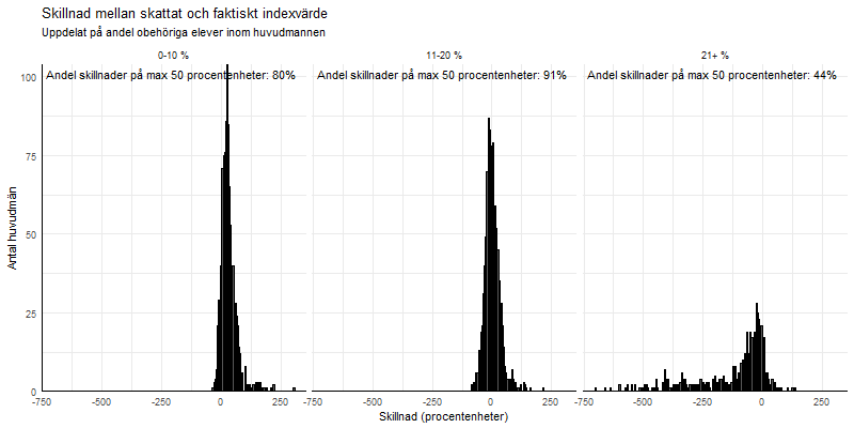


Figur 17. Skillnad i skattad och faktisk andel obehöriga elever på huvudmannanivå, utifrån andelen obehöriga elever inom huvudmannen.

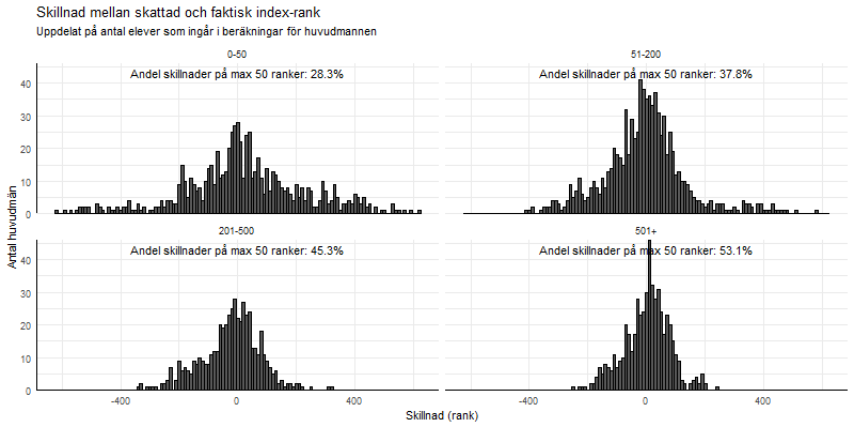
Diagram och tabeller: Index



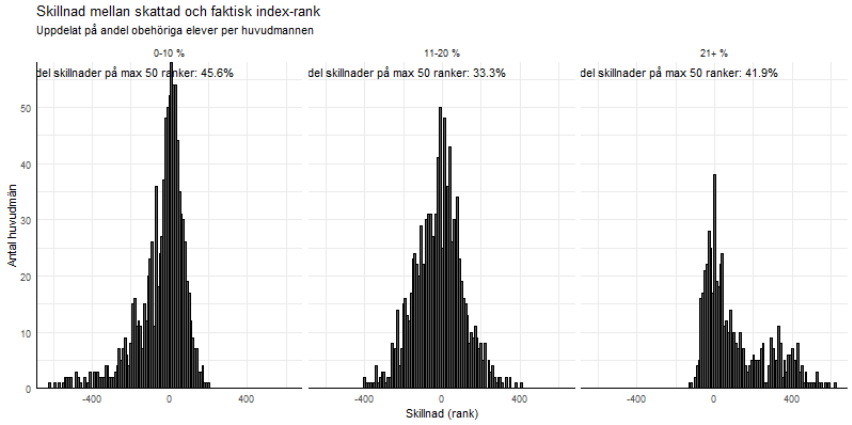
Figur 18. Skillnad mellan skattat och faktiskt indexvärde, per skolstorlek.



Figur 19. Skillnad mellan skattat och faktiskt indexvärde, per andel obehöriga elever.



Figur 20. Skillnad mellan skattad och faktisk index-rank, per skolstorlek.



Figur 21. Skillnad mellan skattad och faktisk index-rank, per andel obehöriga elever.