

Analys av likvärdig betygssättning mellan elevgrupper och skolor

Jämförelser mellan betyg och nationella prov i årskurs 9



Publikationen finns att ladda ner som
kostnadsfri PDF från Skolverkets webbplats:
skolverket.se/publikationer

ISSN: 1103-2421

ISRN: SKOLV-R-475-SE

Omslagsbild: Elke Welzbacher, Skolverket
Skolverket, Stockholm, 2019

Förord.....	3
Sammanfattning	5
Endast små skillnader i betygssättning mellan olika elevgrupper.....	5
Det är stora skillnader i betygssättning mellan olika skolor	5
Betygssättningen är i stor utsträckning relativ	5
Betygssättningen är mer generös i fristående skolor.....	6
Analysen visar på grundläggande problem med betygssystemet.....	6
1. Inledning	7
Vad rapporten handlar om.....	7
Varför det är viktigt med likvärdig betygssättning	7
Betyg och nationella prov som mått på kunskaper	9
Nationella prov som indikator på likvärdig betygssättning	11
Analysmetod och data.....	13
2. Skillnader i betygssättning av olika elevgrupper	17
Skillnader i betygssättning mellan flickor och pojkar	17
Skillnader i betygssättning mellan utlandsfödda elever och elever födda i Sverige.....	19
Skillnader i betygssättning utifrån socioekonomisk bakgrund.....	20
3. Skillnader i betygssättning mellan skolor	22
Skillnader i betygssättning mellan skolor totalt.....	22
Skillnader i betygssättning mellan olika kategorier av skolor.....	26
Hur mycket av skolskillnaderna i betygssättning kan förklaras av skolfaktorer?.....	29
Hur har skolskillnaderna i betygssättning förändrats över tid?	31
4. Sammanfattat resultat	34
5. Diskussion	36
Betygssättningen av elevgrupper	36
Betygssättningen mellan skolor.....	36
Bilaga B1	43
B1.1 Datakällor och exkluderingar.....	43
B1.2 Variabler som ingår i analyserna.....	45
B1.3 Metoder.....	52
B1.4 Osäkerhet i skattningarna av systematiska avvikelser i betygssättning	56
Bilaga 2 Kompletterande analyser	59
B2.1 Skillnader i betygssättning uppdelat på provbetyg	59
B2.2 Skillnader i betygssättning utifrån betygssteg	60
B2.3 Skillnader i skattning av betygssättning beroende på val av statistisk metod	61
B2.4 Spridning i genomsnittlig avvikelse mellan kommunala skolor respektive mellan fristående skolor	62
B2.5 Relativ betygssättning mellan kommunala respektive mellan fristående skolor	63
B2.6 Andel av variationen i betygssättning mellan skolor som kan förklaras av huvudman ...	64
B2.7 Andelen behöriga lärare på skolan och betygssättning.....	64
B2.8 Relativ betygssättning över tid	66
Referenser	67

Förord

Sedan den första rapporten 2004 har Skolverket flera gånger gjort omfattande analyser som bygger på jämförelser mellan betyg och nationella prov. Skolverket ger dessutom ut årliga pm om relationen mellan betyg och nationella prov för grund- och gymnasieskolan. I den här rapporten har vi utvecklat metoderna och analyserna på flera sätt. Det har vi kunnat göra tack vare värdefulla synpunkter från flera medarbetare på Skolverket och från flera forskare som är verksamma på området. Vi vill därför särskilt tacka *Alli Klapp*, Göteborgs universitet, *Christina Wikström*, Umeå universitet, *Björn Tyrefors*, Institutet för näringslivsforskning och *Jonas Vlachos*, Stockholms universitet. Rapporten är skriven av *Jonas Sandqvist* och *Anders Auer* vid Skolverkets analysavdelning. Alla analyser, tolkningar och slutsatser som görs i rapporten är helt och hållet Skolverkets egna.

Peter Fredriksson
Generaldirektör

Jonas Sandqvist
Undervisningsråd

Sammanfattning

I denna rapport analyserar vi hur betygen i grundskolans årskurs 9 sätts i relation till resultaten på de nationella proven. En central utgångspunkt är att de nationella proven fungerar som en indikator för att bedöma om betygssättningen är likvärdig mellan skolor och elevgrupper. Stora skillnader i betygssättning i förhållande till de nationella proven mellan olika skolor och mellan olika elevgrupper indikerar att lärarna sätter betyg utifrån olika måttstockar och att betygssättningen inte är likvärdig.

Vi analyserar betygssättningen mellan flickor och pojkar, mellan utlandsfödda elever och elever födda i Sverige samt mellan elever med olika socioekonomiska bakgrund. Vi analyserar den totala skillnaden i betygssättning mellan skolor och hur utvecklingen ser ut över tid. Dessutom undersöker vi skillnader i betygssättning mellan kommunala och fristående skolor samt mellan skolor med olika prestationsnivå (olika hög nivå på provbetygen).

Vi kommer fram till följande huvudresultat:

Endast små skillnader i betygssättning mellan olika elevgrupper

Vi ser endast små skillnader i betygssättning i förhållande till de nationella proven för olika elevgrupper. Betygssättningen gynnar flickor och utlandsfödda elever, men dessa skillnader är små. När det gäller elever med olika socioekonomisk bakgrund så syns inga märkbara skillnader alls i betygssättningen.

Det här är ett i grunden positivt resultat. På det stora hela förmår lärarna att sätta betyg utan att gynna eller missgynna någon elevgrupp.

Det är stora skillnader i betygssättning mellan olika skolor

Skillnaden i betygssättning i relation till de nationella proven mellan olika skolor är stor. Det finns många skolor där en stor del av eleverna får högre betyg än provbetyg och det finns många skolor där de flesta elever får samma eller lägre betyg än provbetyg.

Många insatser har genomförts både på den statliga och den lokala nivån för att åstadkomma en mer likvärdig betygssättning men inget tyder på att skillnaderna mellan skolors betygssättning har minskat mellan åren 2005 och 2017.

Betygssättningen är i stor utsträckning relativ

Analysen visar att betygssättningen i hög grad är relativ, det vill säga att den relateras till den genomsnittliga prestationsnivån hos eleverna på skolorna. En elev har betydligt lägre sannolikhet att få ett högre betyg givet sitt resultat på det nationella provet om eleven går i en skola med hög genomsnittlig prestationsnivå.

Den relativa betygssättningen kan förklara mellan 25–40 procent av de totala skillnaderna mellan olika skolors betygssättning i relation till de nationella proven. De faktiska betygsskillnaderna mellan skolor skulle alltså ha varit betydligt större om det inte vore för den relativa betygssättningen.

Betygssättningen är mer generös i fristående skolor

Rapporten analyserar också betygssättningen mellan skolor med kommunal respektive fristående huvudman. Det är en stor variation i betygssättning i relation till de nationella proven mellan såväl olika kommunala skolor som mellan olika fristående skolor, men i genomsnitt har fristående skolor en mer generös betygssättning. Men skillnaden i betygssättning mellan kommunala och fristående skolor kan endast förklara en mycket liten del, endast omkring 1,5 procent, av de totala skillnaderna i betygssättning mellan skolor.

Analysen visar på grundläggande problem med betygssystemet

Att betygssättningen i relation till de nationella proven varierar så mycket mellan olika skolor generellt samt det tydliga inslaget av relativ betygssättning visar på en grundläggande problematik i betygssystemet att åstadkomma en nationellt likvärdig bedömning av eleverna. En förklaring till den bristande likvärdigheten är svårigheterna för lärare på olika skolor att finna en samsyn i bedömningen. Betygssättningen av elevernas prestationer sker utifrån en tolkning av kursplanernas kunskapskrav och de värdeord som finns för att särskilja de olika betygsstegen. Det är en övermäktig uppgift för lärarna att åstadkomma en gemensam tolkning som leder till en nationellt likvärdig betygssättning av eleverna. De grundläggande problemen är inbyggda i betygssystemet och kan inte tillskrivas lärarkollektivet. För att nå en markant ökad likvärdighet i betygssättningen krävs systemförändringar i betygssystemet, exempelvis att de nationella proven på något vis styr betygssättningen på skolnivå. Sådana systemförändringar behöver utredas noggrant och vägas mot tänkbara negativa effekter som förändringarna kan få.

Detta utesluter inte vikten av en fortsatt strävan att ge lärare bättre förutsättningar att utveckla en ökad samsyn om hur kunskapskraven ska förstås och på vilka grunder betyg sätts. Här har lärarutbildningen ett ansvar. Lärarstudenterna behöver få grundläggande kunskaper i betyg och bedömning. Här har också skolans huvudmän och rektorer ett ansvar. Lärarna behöver inom ramen för det kollegiala lärandet ges utrymme att diskutera betyg och bedömning. Även Skolverket har ett ansvar. Lärarna ska från Skolverket kunna ta del av ett användarvänligt och behovsanpassat bedömningsstöd som kan stödja den enskilda läraren vid betygssättning och vara ett underlag i det kollegiala lärandet.

1. Inledning

Vad rapporten handlar om

Skolverket har tidigare gjort analyser av betygssättningen och hur den förhåller sig till resultaten på nationella prov, både för grundskolan och gymnasieskolan.¹ Fokus i analyserna har varit på skolskillnader i betygssättning i relation till nationella prov. Detta undersöks även i denna rapport som handlar om betygssättningen i årskurs 9. Vi undersöker skillnader i betygssättning i relation till nationella prov mellan skolor över tid.² Dessutom undersöker vi skillnader i betygssättning i relation till nationella prov mellan kommunala och fristående skolor och mellan skolor med olika prestationsnivå (olika nivå på provresultaten). Vi undersöker även betygssättning i relation till nationella prov för tre olika elevgrupper: flickor och pojkar, utlandsfödda och födda i Sverige samt elever med olika socioekonomisk bakgrund. Vi undersöker betygssättningen i samtliga tio ämnen som har nationella prov.³

I förhållande till föregående rapporter om betygssättning har vi utvecklat analysmetoderna och vi har även tillgång till fler och bättre data om elevernas bakgrund. Detta kommer att beskrivas kortfattat men finns mer utförligt beskrivet i bilagan till rapporten.⁴

Varför det är viktigt med likvärdig betygssättning

Det finns flera anledningar till varför en likvärdig betygssättning är viktig.⁵ Betygen i årskurs 9 har flera funktioner. En central funktion är att betygen fungerar som urvalsinstrument till gymnasieskolan och dess olika program. Betygen är helt avgörande för om eleven kommer in på ett nationellt program eller inte och dessutom till vilket nationellt program.⁶ Om inte elevens betyg räcker till för att börja på ett nationellt program så hänvisas eleven till ett introduktionsprogram.⁷ Om det är fler sökande än det finns platser avgör betygen även vilken gymnasieskola en elev kommer in på. Ju högre betyg eleven får desto fler program och skolor har eleven att välja på och ju lägre betyg desto färre valmöjligheter.

¹ Skolverket (2004), (2007), (2009) och (2016a). Se även Skolverket (2015a) och (2015b).

² Vi använder begreppet ”skola” i rapporten även om det juridiska begreppet i skollagen sedan i juli 2011 är ”skolenhet”.

³ Matematik, engelska, svenska, fysik, biologi, kemi, geografi, historia, religionskunskap samt samhällskunskap.

⁴ En excelfil med rapportens figurer och kompletterande analyser kan hämtas på samma ställe där rapporten finns på Skolverkets webbplats. Även en teknisk rapport kommer att finnas tillgänglig där de statistiska analyserna är utförligt beskrivna.

⁵ Skolverket (2018b) s.6. Se Klapp (2018) för en grundläggande genomgång av betygens funktioner och vad de mäter.

⁶ <https://www.skolverket.se/regler-och-ansvar/ansvar-i-skolfragor/antagning-till-gymnasieskolan>

⁷ Betygen kan även avgöra vilka inriktningar på introduktionsprogram som en elev får gå.

Detta innebär att om betygssättningen inte är likvärdig kan det leda till att elever som konkurrerar om studieplatser behandlas olika. Elever som har kunskaper och förutsättningar att klara en utbildning kan få stå tillbaka för elever med sämre kunskaper men som har fått en mer generös betygssättning. Elever som fått en generös betygssättning och precis kommit in på ett nationellt program kanske inte har kunskaper nog att klara gymnasiestudierna medan elever som fått en restriktiv betygssättning men inte kommit in kanske skulle ha klarat studierna. Elever som får en relativt generös betygssättning kommer kanske in på den utbildning de önskar medan elever som får en relativt sträng betygssättning kan tvingas gå på ett program eller en skola de inte önskar.⁸ En icke likvärdig betygssättning innebär också att samhället inte tar tillvara elevernas kompetens på ett optimalt sätt.

Det finns också ett rättssäkerhetsperspektiv att beakta. Betygssättning är myndighetsutövning och betygen går inte att överklaga, till skillnad mot de flesta andra beslut som rör enskilda individer.⁹ Om betygen ska vara rättssäkra förutsätter det att elevernas kunskaper blir bedömda på ett likvärdigt sätt. Om elever och föräldrar uppfattar att betygssättningen inte är likvärdig riskerar det att skada deras tilltro till skolan och samhället.¹⁰

Slutligen är betygen också viktiga som informationsbärare på såväl skol- som systemnivå.¹¹ Betyg förutsätts vara viktig information om skolorna när föräldrar och elever ska välja skola.¹² Om betygen är mer generöst satta på vissa skolor kan dessa framstå som att de håller bättre kvalitet i undervisningen än vad de faktiskt gör.¹³ Betyg tillhör också den mest centrala information som finns att tillgå för skolhuvudmän och rektorer om de skolor de ansvarar för. I flera skolämnen är betygen det enda resultatmått som finns om elevernas kunskaper.¹⁴ Om betygen inte speglar elevernas faktiska kunskaper blir det svårt att veta hur elevernas kunskaper i olika skolor verkligen ser ut och om insatser som gjorts av stat och huvudmän för att höja elevernas kunskaper verkligen gett avsedda effekter. En

⁸ Diamond & Persson (2016), som använder en annan metod än i denna rapport för att avgöra diskriminerande betygssättning, visar på positiva effekter för en elev att bli generöst betygssatt, både när det gäller gymnasieresultat och arbetsmarknadsutfall, jämfört med elever som fått en mindre generös betygssättning.

⁹ SOU 2010:96 s.54ff.

¹⁰ Se Skolverket (2016a) s.59ff och SOU 2010:96 s.68ff för genomgångar av elevers syn på hur likvärdig betygssättning är.

¹¹ Betygen är grundläggande i den officiella statistiken för skolväsendet som Skolverket publicerar på totalnivå och om enskilda skolors resultat.

¹² Se t ex <http://www.valjaskola.se/> Heller Sahlgren & Jordahl (2016).

¹³ Det är ett särskilt stort problem i ett skolsystem som det svenska där skolval och konkurrens mellan skolor är en central del. Se till exempel Holmlund m fl (2014) s.20ff. Vlachos (2018).

¹⁴ Det finns 20 ämnen i grundskolan och 11 ämnen (inklusive svenska som andraspråk) som har nationella prov.

icke likvärdig betygssättning gör det också svårt för myndigheter, forskare och andra att studera skolans kunskapsutveckling.¹⁵

Betyg och nationella prov som mått på kunskaper

I läroplanen uttrycks kortfattat vad betygen ska grunda sig på:

Betyget uttrycker i vad mån den enskilda eleven har uppnått de nationella kunskapskrav som finns för respektive ämne. Som stöd för betygssättningen finns ämnesspecifika kunskapskrav för olika betygssteg.

Läraren ska... vid betygssättningen utnyttja all tillgänglig information om elevens kunskaper i förhållande till de nationella kunskapskraven och göra en allsidig bedömning av dessa kunskaper.¹⁶

Lärarens bedömningsunderlag i ett ämne i årskurs 9 innehåller elevprestationer från olika redovisningssätt och sammanhang, både sådant som är dokumenterat och som inte är dokumenterat.¹⁷ Läraren ska utifrån bedömningsunderlagen värdera kvaliteten på elevens kunnande utifrån kunskapskraven i ämnets kursplan. Kunskapskraven ska tolkas i relation till ämnets syfte och den undervisning som bedrivits. Skillnaden mellan olika betygssteg uttrycks genom så kallade värdeord i kunskapskraven.¹⁸ Det räcker emellertid inte med att varje lärare utgår från kunskapskraven och sina betygsunderlag om eleverna när de sätter betyg – för att nå en likvärdig betygssättning förutsätts också en samsyn mellan landets lärare.¹⁹

De analyser som bedrivits om lärares betygssättning visar på olika svårigheter när det gäller att få till stånd en likvärdig bedömning mellan lärare.²⁰ Mycket av analyserna har handlat om svårigheter för lärarna att hitta en gemensam tolkning av kunskapskraven.²¹ Ett annat problem som forskningen identifierat med betygssättningen är att även personliga egenskaper hos eleverna i viss mån vägs in i

¹⁵ Holmlund m fl (2014) s.293ff. Gustafsson m fl (2014) kap 3. SOU 2016:25 del 1 kap 5.

¹⁶ Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011. Se även Skolverket (2011) s.21ff och <https://www.skolverket.se/undervisning/grundskolan/betyg-i-grundskolan/satta-betyg-i-grundskolan>

¹⁷ Skolverket (2018b) s.32ff.

¹⁸ Tillämpades för första gången i årskurs 9 vårterminen 2013. Innan dess fanns betygsriterier. För en beskrivning av kunskapskraven och värdeorden se <https://www.skolverket.se/undervisning/grundskolan/betyg-i-grundskolan/satta-betyg-i-grundskolan#h-Kunskapskraven>

¹⁹ Skolverket (2018b) s.21f. Jmf Skolverket (1998) s.39 om jämförbarheten i betygssystemet.

²⁰ Se t ex Klapp (2018) och SOU 2010:96 bilaga 5 för en genomgång av forskningen. Se även Skolverket (2016a). För en internationell översikt av analyser av orsaker bakom diskriminerande betygssättning, se Malouff (2008) och Malouff & Thorsteinsson (2016).

²¹ Kunskapskraven tillämpades för första gången i årskurs 9 vårterminen 2013. Innan dess fanns betygsriterier.

bedömningen, exempelvis elevens vilja att anstränga sig och andra beteenden och attityder.²² Detta kan slå olika för olika elevgrupper.²³

De nationella proven har i jämförelse med betygen mindre problem med subjektiva inslag i bedömningen. Men inte heller de nationella proven kan ses som helt objektiva mått på elevers kunskaper, varken på individnivå eller mellan elevgrupper eller skolor.²⁴ Om en elevgrupp eller skola blir gynnad eller missgynnad när det gäller de nationella proven kan det bero på flera orsaker. Det kan ske vid själva bedömningen (rättningen) av proven. Detta har bland annat undersökts genom analyser av datamaterial bestående av nationella prov som bedömts av ordinarie lärare och sedan ombedömts av andra lärare. Analyserna har visat att det finns skillnader mellan skolor när det gäller bedömningen av proven och att elevgrupper i viss mån kan gynnas eller missgynnas i själva bedömningen av proven.²⁵ Den andra typen av gynnande eller missgynnande när det gäller nationella prov kan uppstå på grund av provens konstruktion och själva provsituationen. Det finns omfattande forskning som analyserar om olika elevgrupper (framförallt skillnader mellan kön) gynnas eller missgynnas av prov som bedömningsform jämfört med betyg som bedömningsform.²⁶ Valet av provfrågor och provfrågornas utformning kan också missgynna vissa elevgrupper.²⁷ Själva provsituationen kan också vara

²² Se exempelvis SOU 2010:96 bilaga 5. Klapp Lekholm & Cliffordson (2009). Lundahl m fl (2015). Thorsen (2014). Lundahl m fl (2010) s.150f. Fischer m fl (2013).

²³ Se Thorsen & Cliffordson (2012) som kommer fram till att flickor och elever i skolor med låg socioekonomisk sammansättning i högre grad blir bedömda utifrån personliga egenskaper. Jmf även Terrier (2015).

²⁴ För en generell diskussion om validitets- och reliabilitetsproblem gällande nationella prov se t ex Lind Pantzare (2018). Se även Skolverket (2004), SOU 2016:26 del 2, Lundahl (2009), Gustafsson m fl (2014), Tengberg & Skar (2017).

²⁵ Tyrefors Hinnerich m fl (2011), Tyrefors Hinnerich m fl (2015), Tyrefors Hinnerich & Vlachos (2012), Tyrefors Hinnerich & Vlachos (2013). Se även Gustafsson & Ericksson (2013) och Gustafsson m fl (2014) s.31 för en diskussion om ombedömningar. För en generell diskussion och redovisning av forskning om diskrepanser i bedömningen av nationella prov se Skolverket (2008) s.22ff. För ett exempel om samstämmighet i bedömningar av nationella prov i samhällskunskap, se Löfstedt (2018). För en analys på engelska data se Hinton & Higson (2017).

²⁶ En aktuell forskningsgenomgång görs i Wikström & Wikström (2017). Se även Voyer & Voyer (2014).

²⁷ Se t ex Reardon m fl (2018) som visar att testfrågornas utformning kan ha en stor betydelse för provresultaten för olika kön. Se Skolverket (2017) s.22 för en principiell diskussion inför konstruktionen av nationella prov. Se Skolverket (2009) s.45ff om huruvida utformningen av proven diskriminerar när det gäller elever med ett visst kön och elever med annat modersmål. Vad gäller de universitetsinstitutioner som utformar de nationella proven så sker deras analyser utifrån ett insamlat provmaterial där det enda som är känt om elevernas bakgrund är elevens kön och om eleven läst efter kursplanen i svenska eller svenska som andraspråk. Det görs också enkäter till lärare med frågor om hur proven fungerar för olika elevgrupper. Se Alm Fjellborg & Molin (2018) som med detta insamlade provmaterial analyserar hur olika provuppgifter i geografi fungerar för elever som följer kursplanen för svenska respektive svenska som andraspråk.

mer eller mindre svår att hantera för olika grupper av elever – vissa kan påverkas mer av negativ stress.²⁸

Ett särskilt problem är att de nationella proven av tidsskäl och andra skäl inte kan utvärdera elevernas kunskaper i förhållande till hela ämnets syfte, centrala innehåll och kunskapskrav. Man skulle därmed kunna hävda att elever på vissa skolor haft mer otur och har undervisats relativt lite om de kunskaper som mäts i just de provuppgifter som råkade komma på det nationella provet – och att eleverna har bättre kunskaper i områden som inte förekommer i provet. Mot att detta skulle generera större skillnader i provresultat mellan skolor talar att de nationella proven prövar centrala, representativa kunskaper i ämnena. Mycket talar för att undervisning och läromedel i hög grad styr in mot de kunskaper som prövas i de nationella proven.²⁹

Det finns ytterligare problem med de nationella provens likvärdighet. Lärare och skolor kan i olika utsträckning ha förberett sig för de nationella proven genom att träna på gamla prov. Det finns också ett växande problem med läckor och fusk när det gäller nationella prov.³⁰

Trots ovanstående problem gör vi bedömningen att jämförelser av hur betygssättningen sker i relation till de nationella proven för olika elevgrupper och skolor är en relevant indikator på hur generös betygssättningen är.

Nationella prov som indikator på likvärdig betygssättning

De nationella proven har alltid varit ett viktigt bedömningsunderlag i de ämnen som haft sådana.³¹ De ska dessutom numera, enligt den nya bestämmelse som trädde i kraft 29 juni 2018, ”särskilt beaktas” vid betygssättningen.³² De nationella proven ska fungera som ett stöd för en likvärdig betygssättning men inte vara det enda som avgör en elevs betyg.³³ Det finns alltså ingen självklar överensstämmelse mellan betyg och provbetyg för en enskild elev. Det är fullt rimligt att läraren bedömer att betygsunderlagen vid en samlad bedömning ger skäl för ett annat betyg än vad eleven fått på det nationella provet. Det innebär att det inte är möjligt att använda de nationella proven som en indikator på likvärdig betygssättning för enskilda elever.

På aggregerad nivå, till exempel för en skola så kan man förvänta sig högre samstämmighet mellan betygen och de nationella provresultaten. Om de nationella

²⁸ Eklöf & Nyroos (2013) påvisar en högre grad av provstress hos flickor men det är oklart om detta påverkar prestationen. Se även Heissel m fl (2018) om hur provstress kan påverka prestationsförmågan hos olika elevgrupper.

²⁹ Skolverket (2007) s.68.

³⁰ Riksrevisionen (2018).

³¹ Se t ex Skolverket (2004) s.34ff. Skolverket (2015c) s.251.

³² <https://www.skolverket.se/undervisning/grundskolan/nationella-prov-i-grundskolan/genomfora-och-bedoma-prov-i-grundskolan#Provresultatetsbetydelseforbetyget>

³³ Skolverket (2018b) s.37. För en internationell jämförelse av system för betygssättning se Lundahl m fl (2017) s.161f. Skolverket (2016a) s.101f.

proven korrekt och tillförlitligt sätter gränserna för de olika betygsstegen så skulle man förvänta sig att en skolas betygsgenomsnitt i ett ämne ligger nära det nationella provets genomsnitt. Dessvärre förhåller det sig inte riktigt så. Det är svårt att göra nationella prov som mäter betygsgränserna på exakt samma vis över tid, vilket också visar sig i statistiken där provbetygen varierar ganska mycket från år till år.³⁴

Det finns inte heller några bestämmelser för hur mycket de nationella proven ska styra eller påverka en enskild skolas genomsnittliga betyg i respektive ämne.³⁵ Det är därför inte heller möjligt att uttala sig om huruvida en enskild skola ligger för högt eller för lågt i förhållande till de nationella proven – ett nationellt prov kan vara ovanligt svårt ett visst år och det kan då vara naturligt att skolorna avviker mer i genomsnitt det året jämfört med ett år då provet var relativt lätt.

Det är först när man jämför flera skolor (eller olika elevgrupper) som de nationella proven kan användas som en indikator på likvärdig betygssättning. Våra analyser har som utgångspunkt att skolornas (eller elevgruppernas) avvikelser mellan betyg och nationella prov bör vara ungefär likartad för att man ska kunna tala om att betygssättningen är likvärdig. Om det däremot finns betydande skillnader i hur mycket skolorna avviker indikerar detta ett likvärdighetsproblem med betygssättningen. Det går inte att säga vilken skola eller vilken elevgrupp som betygssätts ”korrekt” eller ”rättvist” i relation till de nationella proven, det vill säga vilken absolut nivå på avvikelserna mellan betyg och provbetyg som indikerar en rättvis betygssättning. Utgångspunkten i denna rapport är istället att avvikelserna mellan betyg och nationella prov bör vara *likartade mellan skolor och elevgrupper* för att betygssättningen kan antas vara likvärdig.³⁶

Bakom detta ligger ett antagande om att det är svårt att hitta legitima skäl, skäl som är förenliga med betygssystemets intentioner, för att en skola eller en elevgrupp uppvisar väsentligt högre (eller lägre) betyg i relation till de nationella proven jämfört med en annan skola eller elevgrupp. Exempelvis: givet att analyserna görs på ett rättvisande sätt samt att inte de nationella proven missgynnar flickor på något avgörande vis är det svårt att finna legitima skäl till att flickor ska få väsentligt högre eller lägre betyg än pojkar i relation till resultaten på de nationella proven. Det är på samma vis svårt att finna några legitima skäl till varför en skola skulle ha betydligt högre betyg i förhållande till de nationella proven än en annan skola. Det går visserligen att argumentera för att en skola skulle vara bättre än en annan på att sätta in stödinsatser efter att de nationella proven har genomförts och bedömts – och av det skälet har större avvikelser mellan betyg och nationella prov.

³⁴ SOU 2016:25 kap 13.1.

³⁵ SOU 2016:25 kap 12.2.

³⁶ Om alla skolor har en hög avvikelse mellan betyg och nationella prov, en hög avvikelse som ser likadan ut mellan alla skolor, så indikerar detta inget stort likvärdighetsproblem i betygssättningen. Däremot kan en sådan situation ge en bild av att det nationella provets betygsgränser inte på ett bra sätt speglar lärarnas uppfattning av betygsgränserna.

Men å andra sidan återstår inte många veckor av stödinsatser innan betygen ska vara satta.³⁷ Denna eventuella orsak bör ha en mindre betydelse när vi gör analyser på nationell nivå mellan alla skolor. Skillnader mellan skolors betygssättning i relation till de nationella proven bör därmed vara en bra indikator på hur likvärdigt betygen sätts.

Det finns också empiriska belägg för att samma betyg står för skilda kunskaper på olika skolor, beroende på skolans relation mellan betygsresultat och nationella prov. Analyser visar att elever som gått på skolor med en relativt generös betygssättning (höga genomsnittliga avvikelser mellan betyg och provbetyg i förhållande till övriga skolor) tenderar att klara sig sämre i gymnasiet jämfört med elever som fått samma betyg men som gått i skolor med en mer restriktiv betygssättning.³⁸ Detta visar att de analyser av skillnader mellan olika elevgrupper och skolor vad gäller avvikelser mellan betyg och provbetyg som vi gör i rapporten också säger något om hur likvärdig betygssättningen är.

Analysmetod och data

När vi jämför betygssättningen i relation till de nationella proven för olika elevgrupper och olika skolkategorier är det viktigt att så långt som möjligt se till att inte andra faktorer än de vi verkligen vill undersöka påverkar resultaten. När vi exempelvis undersöker betygssättningen av utlandsfödda elever vill vi försäkra oss om att det resultat vi får inte beror på att utlandsfödda elever har lägre socioekonomisk bakgrund. Därför innehåller analyserna en kontroll av flera variabler.³⁹ De variabler som ingår i våra analyser är:

- Elevens kön (flicka eller pojke).
- Elevens migrationsbakgrund (utlandsfödd eller född i Sverige).
- Elevens socioekonomiska bakgrund (ett socioekonomiskt index som baseras på föräldrarnas *utbildningsnivå*, *inkomst* och *grad av bidragstagande*).
- Huvudman (om skolan är kommunal eller fristående).
- Skolans genomsnittliga prestationsnivå (skolans genomsnittliga provbetyg i det aktuella ämnet).

När vi till exempel analyserar betygssättningen i relation till de nationella proven för utlandsfödda elever i jämförelse med elever födda i Sverige så är det utifrån principen *allt annat lika*. Jämförelsen görs mellan elever som har exakt samma värden på övriga fyra bakgrundsvariabler. Vi jämför till exempel två elever som har samma kön, har samma nivå på den socioekonomiska bakgrunden och som båda går på en kommunal skola (eller båda på fristående skola) och där skolans genom-

³⁷ Provdatum finns på Skolverkets webbplats <https://www.skolverket.se/undervisning/grundskolan/nationella-prov-i-grundskolan/provdatum-i-grundskolan>. För en diskussion om stödinsatser som en orsak till skillnader i betygssättning mellan skolor, se Skolverket (2004) s.71 och Skolverket (2007) s.72.

³⁸ Skolverket (2007) kap 4.3.

³⁹ I de fall det inte görs en kontroll så framgår det av framställningen.

snittliga provbetyg i det aktuella ämnet dessutom är det samma. Det enda som skiljer de två eleverna är att den ena är utlandsfödd och den andra är född i Sverige.

En annan central faktor vi måste ta hänsyn till är om de elevgrupper vi jämför har olika prestationsnivå (vilket i olika grad gäller för alla våra jämförelser). Gör man inte det kan jämförelsen av betygssättningen mellan elevgrupperna och skolkategorierna bli missvisande.

Det finns här två förhållanden som kan snedvrider analyserna åt olika håll.

1. *Provbetygseffekt*

Sannolikheten för en elev att få ett högre betyg än provbetyg är till viss del beroende av *vilket* provbetyg en elev har fått. Den största skillnaden är att elever som fått betyget F på det nationella provet (det vill säga underkänt) i genomsnitt har en betydligt högre sannolikhet att få ett högre betyg (det vill säga betyg E som är godkänt betyg), jämfört med elever som fått ett godkänt provbetyg (det vill säga E-A).⁴⁰ Det betyder att om den ena av de två elevgrupperna man jämför har en högre andel elever med F i provbetyg så måste man justera för detta, annars riskerar man att dra slutsatsen att elevgruppen med högre andel F i provbetyg betygssätts mer generöst när den egentliga orsaken är att elever med F i provbetyg generellt betygssätts mer generöst, oavsett vilken elevgrupp eleven tillhör.

2. *Fördelningseffekt*

Det andra förhållandet innebär att en elevgrupp med högre prestationsnivå i genomsnitt har kunskaper som ligger närmare ett högre betyg, jämfört med en elevgrupp med lägre prestationsnivå som har samma provbetyg. Om man inte tar hänsyn till detta kan det framstå som att elevgruppen med den högre prestationsnivån blir mer generöst betygsatt när den egentliga orsaken är att de i genomsnitt ligger på en högre kunskapsnivå inom varje enskilt provbetygsintervall.

Ovanstående två förhållanden snedvrider analyserna åt motsatta håll, vilket innebär att de till viss del tar ut varandra. Men man kan inte utgå från att dessa effekter alltid är av samma storlek och därför måste man kontrollera för båda effekterna för att inte riskera att resultaten blir missvisande. I bilagan finns en beskrivning av hur vi kontrollerar för prestationsnivån mellan elevgrupper och skolkategorier.⁴¹

⁴⁰ Det finns också en så kallad tak-effekt, att elever som fått A i provbetyg inte kan få högre betyg. En elevgrupp som har relativt många elever med A i provbetyg får av detta skäl en lägre avvikelse mellan betyg och nationella prov. Detta måste analysen också ta hänsyn till även om det är ett mindre problem än andelen elever som fått F. För en diskussion om möjliga orsaker till att elever med F i provbetyg i högre grad får ett högre betyg se Skolverket (2007) s.68ff. Skolverket (2009) s.65ff. Skolverket (2004) s.70ff.

⁴¹ Bilaga B1.3. Se också den separata tekniska rapporten (kommande) för detaljerad beskrivning av hela metodansatsen.

Analyserna fokuserar huvudsakligen på åren 2013–2015. 2013 var det första året med den nya betygsskalan och 2015 är det senaste året för vilket vi har tillgång till uppgifter om det socioekonomiska indexet. Genom att starta 2013 får vi tillgång till alla tio ämnen som idag har nationella prov. Dessutom undviker vi eventuella problem som kan uppstå på grund av att betygsskalan ändrades mellan 2012 och 2013. Vi har analyserat varje år för sig men sedan slagit ihop resultaten för dessa tre år för att få ett stabilare mått. Orsaken till detta är att de nationella proven tenderar att variera i svårighetsgrad (de kravgränser som sätts i proven för olika betyg).⁴² Detta skulle kunna påverka olika grupper olika. Vi nyttjar även en längre tidserie, 2005–2017, när vi studerar de samlade skolskillnaderna över tid. I bilagan finns en utförlig beskrivning av våra data.

⁴² Skolverket (2017) s.33ff. SOU 2016:25 bilaga 5.

Förklaring av några begrepp som används i rapporten:

Systematisk skillnad i betygssättning

Om en elevgrupp A uppvisar en större genomsnittlig avvikelse mellan betyg och provbetyg jämfört med en annan elevgrupp B, tolkas detta som att det finns en systematisk skillnad i betygssättning mellan elevgrupperna, till fördel för elevgrupp A. Den systematiska skillnaden i betygssättning avser de skillnader som finns *efter* kontroll för övriga bakgrundsvariabler.

Generös/restriktiv betygssättning

När det finns en systematisk skillnad i betygssättning, dvs. när det finns skillnader i den genomsnittliga avvikelsen mellan betyg och provbetyg mellan två elevgrupper, säger vi att den elevgrupp som uppvisar störst avvikelse betygssätts mer generöst eller mindre restriktivt (jämfört med den andra elevgruppen). Den elevgrupp som uppvisar minst avvikelse sägs istället betygssättas mer restriktivt eller mindre generöst (jämfört med den andra elevgruppen).

Faktisk skillnad i betyg

Med faktisk skillnad i betyg avses hur mycket betygen i genomsnitt skiljer sig mellan till exempel två elevgrupper, utan hänsyn till provbetygen. Detta innebär att om en elevgrupp i genomsnitt har högre betyg jämfört med en annan elevgrupp (oavsett hur de har presterat på det nationella provet), så finns en faktisk skillnad i betyg mellan elevgrupperna. Även den faktiska skillnaden i betyg avser skillnaden *efter* kontroll för bakgrundsvariabler. Den faktiska skillnaden i betyg är med som en referenspunkt som hjälp för att tolka storleken på skillnaden i betygssättning mellan elevgrupper och skolor.

2. Skillnader i betygssättning av olika elevgrupper

I detta kapitel undersöker vi om det förekommer en mer generös i betygssättning i relation till de nationella proven av olika elevgrupper. Utgångspunkten är, utifrån resonemanget i kapitel 1, att relationen mellan betyg och provbetyg bör vara ungefär lika mellan olika elevgrupper, givet att analysen på ett korrekt sätt tar hänsyn till olika sammansättning vad gäller övriga bakgrundsvariabler samt elevgruppernas olika prestationsnivå.

Vi jämför hur betygssättningen skiljer sig åt mellan tre olika elevgrupperingar: flickor jämfört med pojkar, utlandsfödda elever jämfört med elever födda i Sverige samt elever med hög socioekonomisk bakgrund jämfört med elever med låg socioekonomisk bakgrund.

I figuren nedan redovisas två staplar per ämne. De mörkblå staplarna anger skillnaden i betygssättning i relation till de nationella proven mellan elevgrupperna. Om värdet på stapeln är skilt ifrån noll innebär det att den ena eller den andra elevgruppen får en relativt generös betygssättning i förhållande till de nationella proven i respektive ämne. Värdet anges i betygspoäng. Betygspoängen kan variera mellan 0 och 20 i varje ämne (F=0, E=10, D=12,5, C=15, B=17,5, A=20).

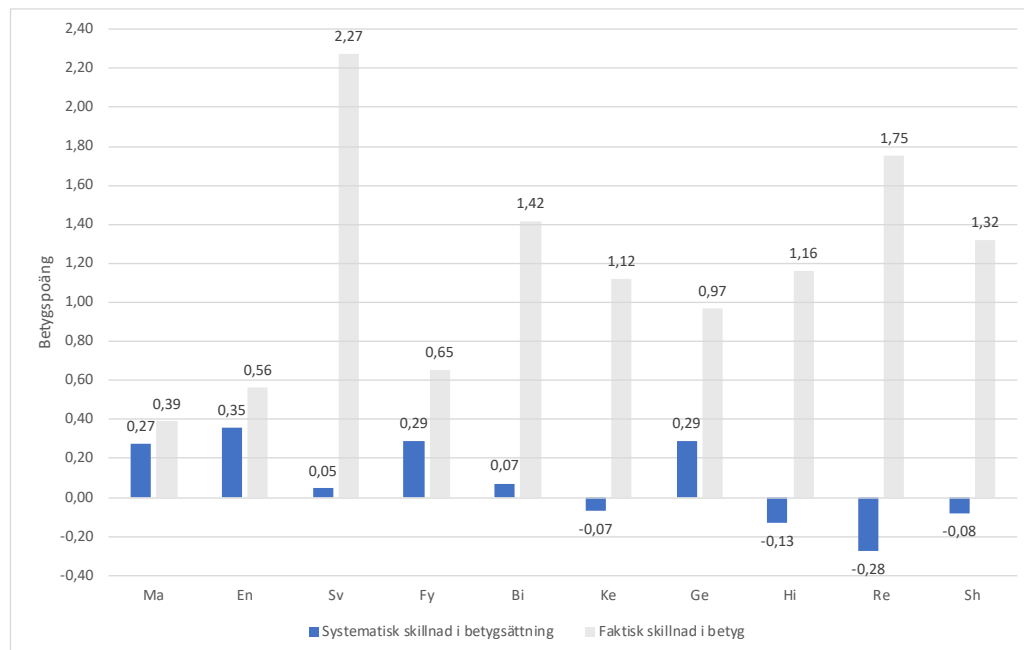
De ljusgrå staplarna visar de faktiska skillnaderna i betyg i respektive ämne mellan de elevgrupper som jämförs. Även de faktiska betygsskillnaderna som redovisas har kontrollerats för övriga bakgrundsvariabler.⁴³ Även de faktiska betygsskillnaderna är uttryckta i betygspoäng. Genom att redovisa de faktiska betygsskillnaderna ges en referenspunkt för att kunna tolka storleken på betygssättningsskillnaderna. Med det sagt är det ändå de absoluta skillnaderna som ligger till grund för bedömningen om betygsskillnaderna är stora eller små.

Skillnader i betygssättning mellan flickor och pojkar

De mörkblå staplarna i figur 1 visar skillnaderna i betygssättning mellan flickor och pojkar. Positiva värden betyder att flickor får en mer generös betygssättning i förhållande till de nationella proven än pojkar. De ljusgrå staplarna visar de faktiska betygsskillnaderna mellan könen. Positiva värden betyder att flickor har högre betyg i ämnet. Alla värden är uttryckta i betygspoäng som kan variera mellan 0 och 20.

⁴³ De faktiska skillnader i betyg som redovisas är så kallade nettoskillnader, vilket innebär att det är skillnaden i betyg mellan två elevgrupper efter att hänsyn tagits till övriga bakgrundsvariabler, enligt samma princip som när vi skattar skillnaden mellan motsvarande elevgrupper i betygssättningen i förhållande till proven.

Figur 1. Skillnader i betygssättning i förhållande till de nationella proven mellan flickor och pojkar, samt faktiska skillnader i betyg mellan samma elever. Uttryckt som genomsnittlig betygspoäng, genomsnitt för åren 2013–2015.



Positiva värden innebär att flickor får högre betyg i förhållande till de nationella proven (mörkblå staplar) respektive erhåller högre faktiska betyg (ljusgrå staplar), givet att alla övriga bakgrundsfaktorer är lika. Betygs-poängen kan variera mellan 0 och 20 i varje ämne. Justerad för provbetygs- och fördelningseffekter. Samtliga skillnader i betygssättning är statistiskt signifikanta (5 procent) utifrån approximativa standardfel (se bilaga B1.4).

I figur 1 visar de mörkblå staplarna hur mycket betygssättningen i förhållande till de nationella proven skiljer sig mellan flickor och pojkar. Den samlade bilden är splittrad. I vissa ämnen ser det ut som flickor gynnas i betygssättningen i förhållande till de nationella proven, i andra ämnen ser det ut att vara tvärtom, i några ämnen syns inte några skillnader alls. Vi återkommer till en samlad bild av skillnaden i betygssättning mellan kön i slutet av rapporten (figur 10).

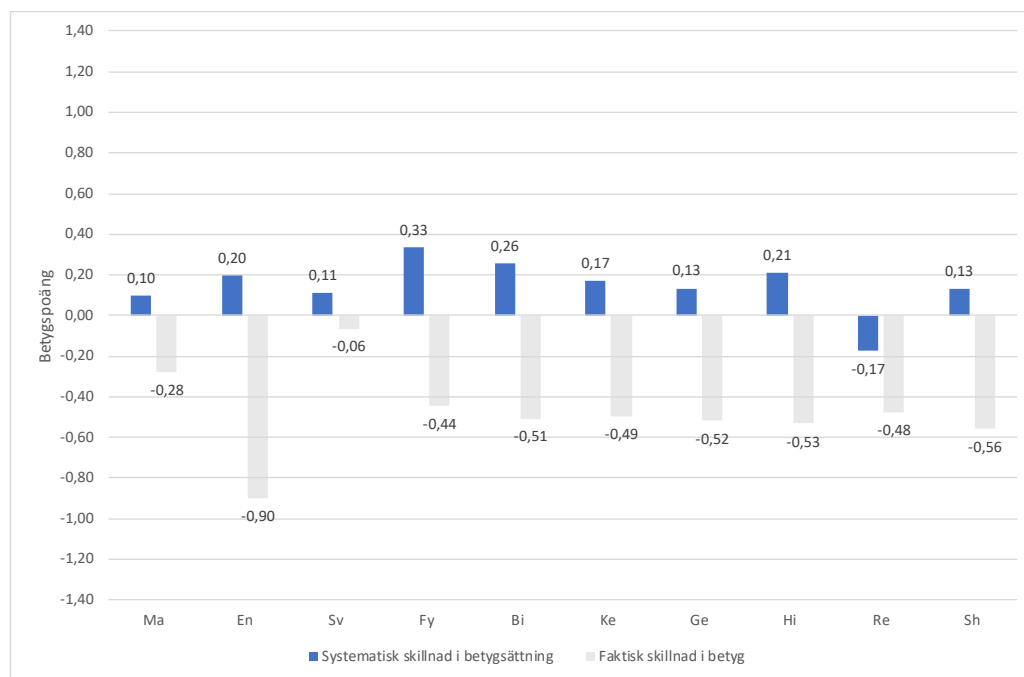
Flickor har högre betyg än pojkar i de flesta ämnen och det har varit så under lång tid.⁴⁴ Detta syns även tydligt i de ljusgrå staplarna i figur 1. En jämförelse mellan skillnaden i betygssättning (mörkblå staplar) och de faktiska betygsskillnaderna (ljusgrå staplar) visar att en mer gynnsam betygssättning ser ut att kunna förklara en relativt stor del av de högre betyg som flickor har i matematik, engelska och fysik. Just i dessa tre ämnen är också de faktiska betygsskillnaderna mellan kön som lägst. I övriga ämnen har betygssättningen en liten betydelse i förhållande till de faktiska betygsskillnaderna.

⁴⁴ Se t ex Sjögren (2017), Voyer & Voyer (2016) och Skolverket (2006).

Skillnader i betygssättning mellan utlandsfödda elever och elever födda i Sverige

I figur 2 visar skillnaderna i betygssättning mellan utlandsfödda elever och elever födda i Sverige (mörkblå staplar). Positiva värden betyder att utlandsfödda elever får en mer generös betygssättning i förhållande till de nationella proven än elever födda i Sverige. De grå staplarna visar de faktiska betygsskillnaderna mellan samma grupper. Positiva värden betyder att utlandsfödda elever har högre betyg i ämnet. Alla värden är uttryckta i betygspoäng som kan variera mellan 0 och 20.

Figur 2. Skillnader i betygssättning i förhållande till de nationella proven mellan utlandsfödda elever och elever födda i Sverige, samt faktiska skillnader i betyg mellan samma elever. Uttryckt som genomsnittlig betygspoäng, genomsnitt för åren 2013–2015.



Positiva värden innebär att utlandsfödda elever gynnas i betygssättningen (blå staplar) respektive erhåller högre faktiska betyg (ljusgrå staplar), givet att alla övriga bakgrundsfaktorer är lika. Betygspoängen kan variera mellan 0 och 20 i varje ämne. Justerad för provbetygs- och fördelningseffekter. Samtliga skillnader i betygssättning är statistiskt signifikanta (5 procent) utifrån approximativa standardfel (se bilaga B1.4).

Enligt de mörkblå staplarna i figur 2 syns vissa systematiska skillnader i betygssättning i relation till de nationella proven till utlandsfödda elevers fördel i alla ämnen, förutom i religionskunskap. Vi återkommer till en sammanställning i figur 10.

De faktiska skillnaderna i betyg (ljusgrå staplar) är negativa, det vill säga utlandsfödda elever har i genomsnitt lägre betyg än elever födda i Sverige, efter att vi tagit

hänsyn till övriga bakgrundsfaktorer.⁴⁵ En slutsats är alltså att de faktiska betygsskillnaderna mellan utlandsfödda elever och elever med svensk bakgrund skulle alltså vara ännu något större om hänsyn togs till skillnaderna i betygssättning i relation till de nationella proven.

Skillnader i betygssättning utifrån socioekonomisk bakgrund

Elevernas socioekonomiska bakgrund mäts med ett index som bygger på föräldrarnas utbildningsnivå, inkomster och grad av bidragstagande.⁴⁶ Indexvariabeln är kontinuerlig (den antar många olika värden på en skala) varför det inte är självklart hur man ska dela in eleverna i två grupper. Vi undersöker därför betygssättningen i förhållande till nationella provet för elever som skiljer sig åt med en standardavvikelse i det socioekonomiska indexet.⁴⁷

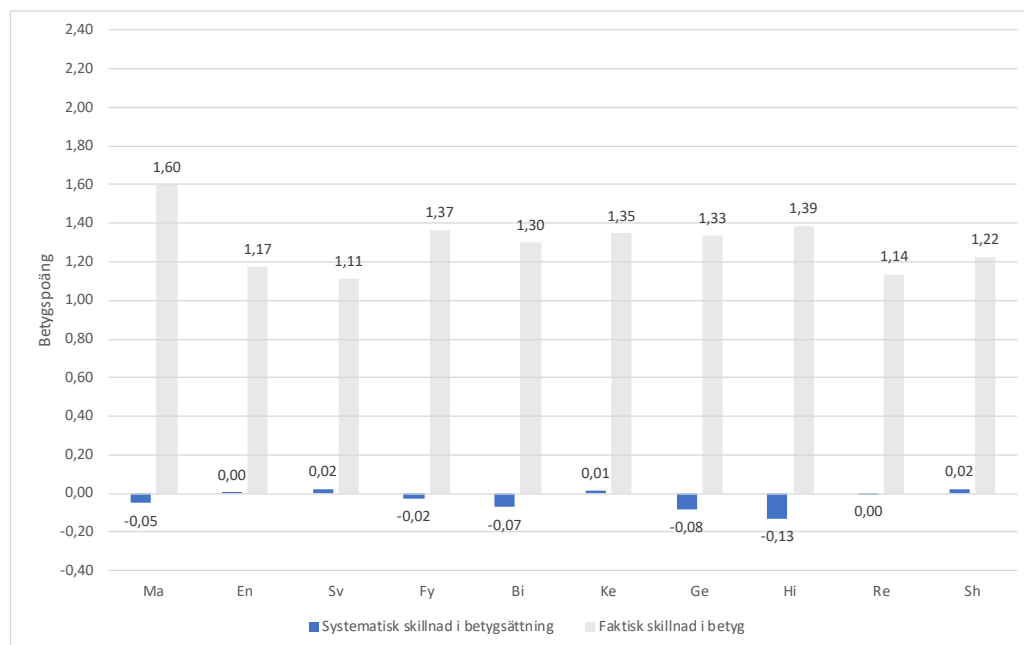
De mörkblå staplarna i figur 3 visar skillnaderna i betygssättning mellan elever med låg och hög socioekonomisk bakgrund. Positiva värden betyder att elever med hög socioekonomisk bakgrund får en mer generös betygssättning i förhållande till de nationella proven. De ljusgrå staplarna visar de faktiska betygsskillnaderna mellan samma grupper. Positiva värden betyder att elever med hög socioekonomisk bakgrund har högre betyg i ämnet. Alla värden är uttryckta i betygspoäng som kan variera mellan 0 och 20.

⁴⁵ Det viktiga här är inte den exakta nivån på de faktiska betygsskillnaderna utan att det viktiga är att få en referenspunkt för betygssättningsskillnaderna som är framtagna på samma sätt – med samma kontroll för bakgrundsvariablerna. Det är särskilt komplicerat att avgöra hur och med vilka variabler man ska kontrollera för att få fram rättvisande faktiska betygsskillnaderna mellan utlandsfödda elever och elever födda i Sverige. Se Grönqvist & Niknami (2017) s.46f för en diskussion.

⁴⁶ Bilaga B1.2.

⁴⁷ Jämförelsen görs över hela skalan av det socioekonomiska indexet, inte bara i mitten av skalan. För att få en mycket ungefärlig idé om vad en standardavvikelse i det socioekonomiska indexet innebär kan två elever jämföras där den enes föräldrar båda har gymnasial utbildning (3 år) medan den andres föräldrar båda har en 2-årig eftergymnasial utbildning. Dessutom har den ene elevens föräldrar en inkomst som ligger precis i mitten av inkomstfördelningen (topp 50 procent) medan den andra elevens föräldrar har en inkomst som ligger lite högre (topp 65 procent). Detta motsvarar en skillnad på ungefär en standardavvikelse i det socioekonomiska indexet.

Figur 3. Skillnader i betygssättning i förhållande till de nationella proven för elever som har en högre (en standardavvikelse) socioekonomisk bakgrund, samt faktiska skillnader i betyg mellan samma elever. Uttryckt som genomsnittlig betygspoäng, genomsnitt för åren 2013–2015.



Ett positivt värde innebär att elever med en standardavvikelse högre värde på det socioekonomiska indexet får högre betyg i förhållande till provbetygen (blå staplar) samt högre faktiska betyg (ljusgrå staplar), givet att alla övriga bakgrunds faktorer är lika. Betygs-poängen kan variera mellan 0 och 20 i varje ämne. Justerad för provbetygs- och fördelningseffekter. För engelska, fysik, kemi och religionskunskap är skillnaderna i betygssättning inte statistiskt signifikanta (5 procent) utifrån approximativa standardfel (se bilaga B1.4).

Figur 3 visar att det i praktiken inte finns några skillnader i betygssättning i förhållande till de nationella proven för elever med lägre och högre socioekonomisk bakgrund (mörkblå staplar). De faktiska skillnaderna i betyg (ljusgrå staplar) som finns mellan elever med lägre och högre socioekonomisk bakgrund är däremot betydande, och de kan inte förklaras med en generösare betygssättning.

3. Skillnader i betygssättning mellan skolor

I detta kapitel flyttar vi fokus från elevgruppsnivå till att undersöka skillnader i betygssättning mellan olika skolor. Det vi analyserar är hur lärare på olika skolor sätter betyg i relation till de nationella proven. Utgångspunkten, som vi beskrivit i kapitel 1, är att på skolnivå så bör betygssättningen i relation till de nationella proven vara likartad för att betygssättningen ska anses vara likvärdig.

Konkret så analyserar vi den genomsnittliga avvikelsen (nettoavvikelsen) mellan betygen och provbetygen per skola och undersöker sedan hur mycket skolorna skiljer sig åt i dessa nettoavvikelser.⁴⁸ I föregående kapitel tolkades resultaten utifrån skillnader i betygspoäng. När skolors nettoavvikelser analyseras i de inledande figurerna i detta kapitel (figur 4a–c och 5) tolkas resultaten istället utifrån andelen elever som får ett högre betyg än provbetyget. Motiveringen är att underlätta tolkningen av figurerna.

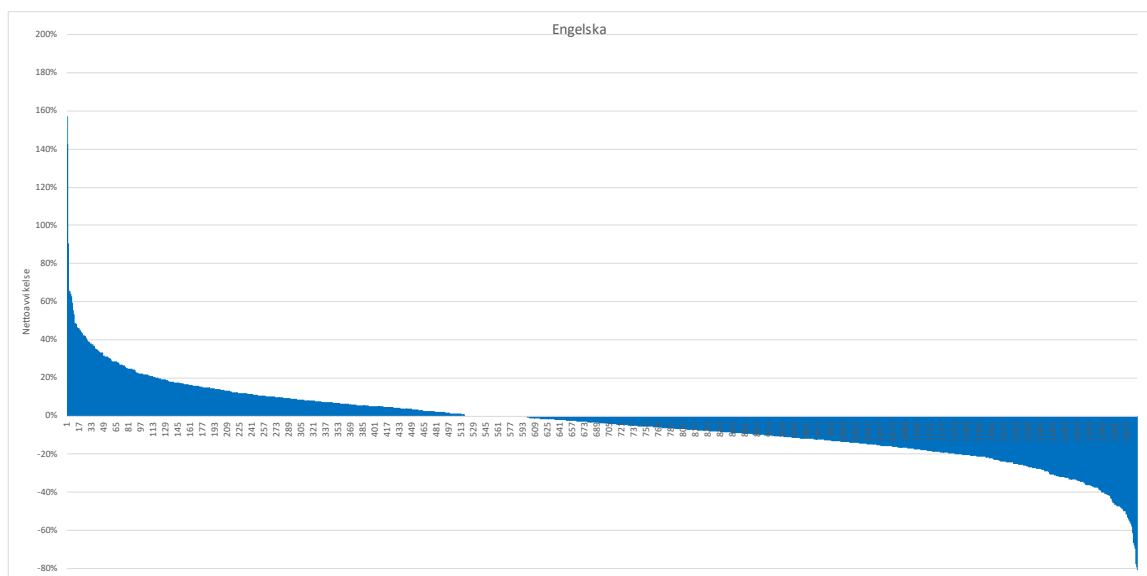
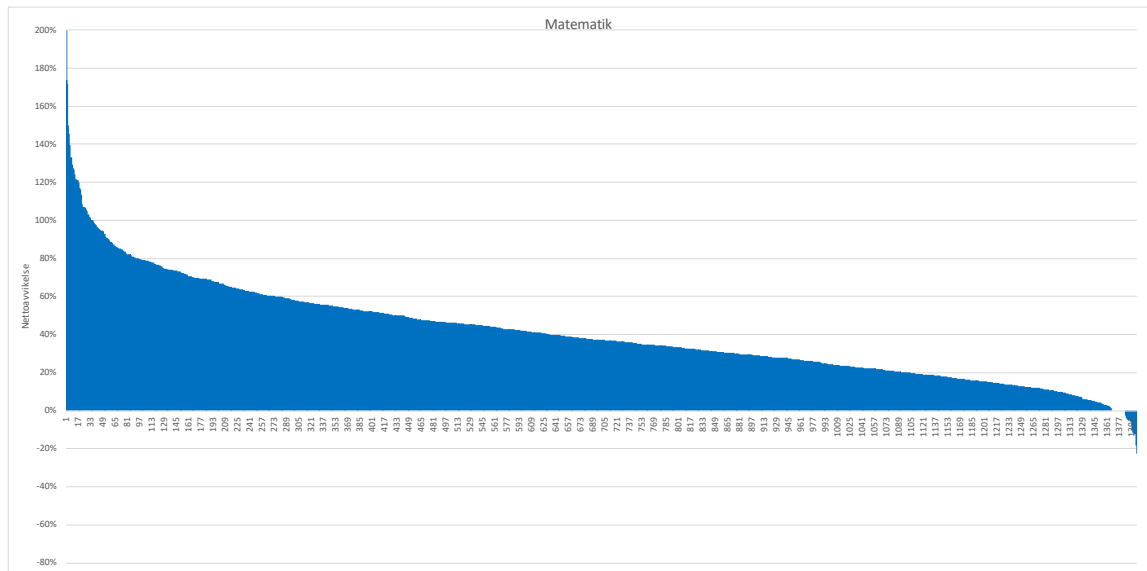
I kapitlets första avsnitt redovisas skillnader i skolors betygssättning i relation till de nationella proven totalt, det vill säga för samtliga skolor. I det andra avsnittet undersöks hur betygssättningen i relation till de nationella proven skiljer sig åt mellan kommunala och fristående skolor respektive för skolor med olika prestationsnivå. I det tredje avsnittet undersöks hur mycket skillnaderna i betygssättning mellan kommunala och fristående skolor samt mellan skolor med olika prestationsnivå kan förklara av de totala skillnaderna i betygssättning mellan skolor. Kapitlet avslutas med en analys av hur skolskillnaderna i betygssättning utvecklats åren 2005–2017.

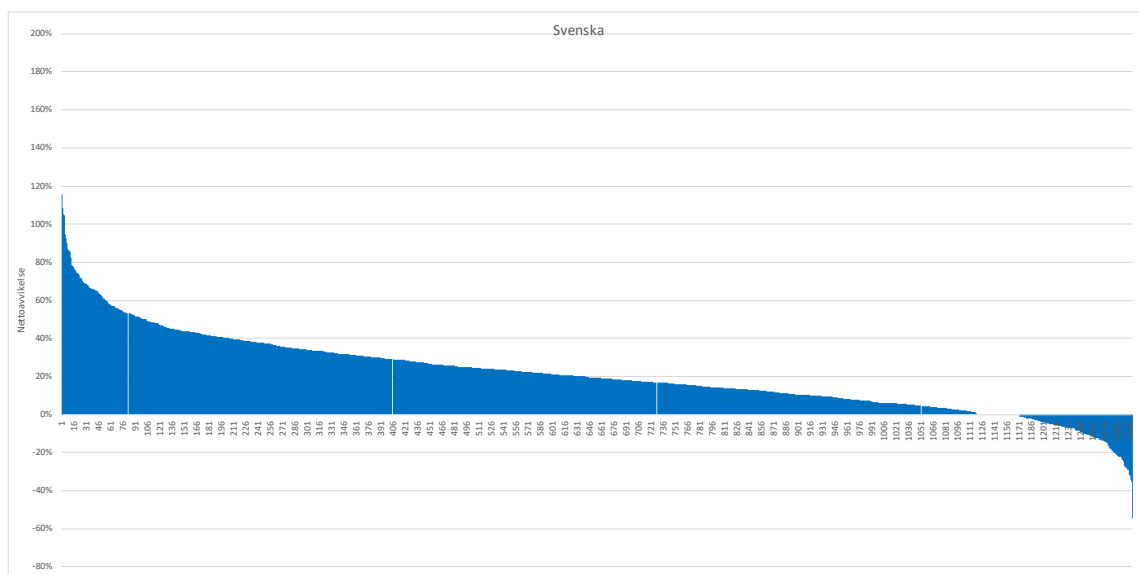
Skillnader i betygssättning mellan skolor totalt

Vi börjar med att på ett grafiskt åskådligt sätt illustrera skillnaderna i skolors betygssättning i förhållande till de nationella proven. I figur 4a–c redovisas skolornas nettoavvikelser i matematik, engelska och svenska för vårterminen 2015. Varje tunn stapel i figuren är en skolas nettoavvikelse (staplarna är så tunna att de knappt går att urskilja). Positiva värden innebär att betygen sätts högre än de nationella proven. Till skillnad från analyserna i föregående kapitel tar vi i denna analys inte hänsyn till elevernas bakgrundsfaktorer och därmed inte heller till elevsammansättningen på skolorna.

⁴⁸ En konkretisering av begreppet nettoavvikelse görs efter figur 4a–c.

Figur 4a–c. Skolskillnader i betygssättning i matematik, engelska och svenska 2015, uttryckt som nettoavvikelsen (den genomsnittliga avvikelsen mellan betyg och provbetyg) för varje skola.





Figuren visar varje enskild skolas genomsnittliga avvikelse (nettoavvikelse) mellan betyg och provbetyg (utifrån betygssteg) som en stapel på den vertikala axeln. Eftersom det finns mellan 1300 och 1400 skolor representerade i varje enskild figur så är det svårt att urskilja enskilda staplar. Skolorna är rangordnade så att skolor med högst nettoavvikelse är placerade längst till vänster och skolor med lägst nettoavvikelse längst till höger i figuren. Endast skolor med minst 15 elever i respektive ämne ingår i figuren och analyserna. Skolornas nettoavvikelser är ej justerade för provbetygs- eller fördelningseffekter.

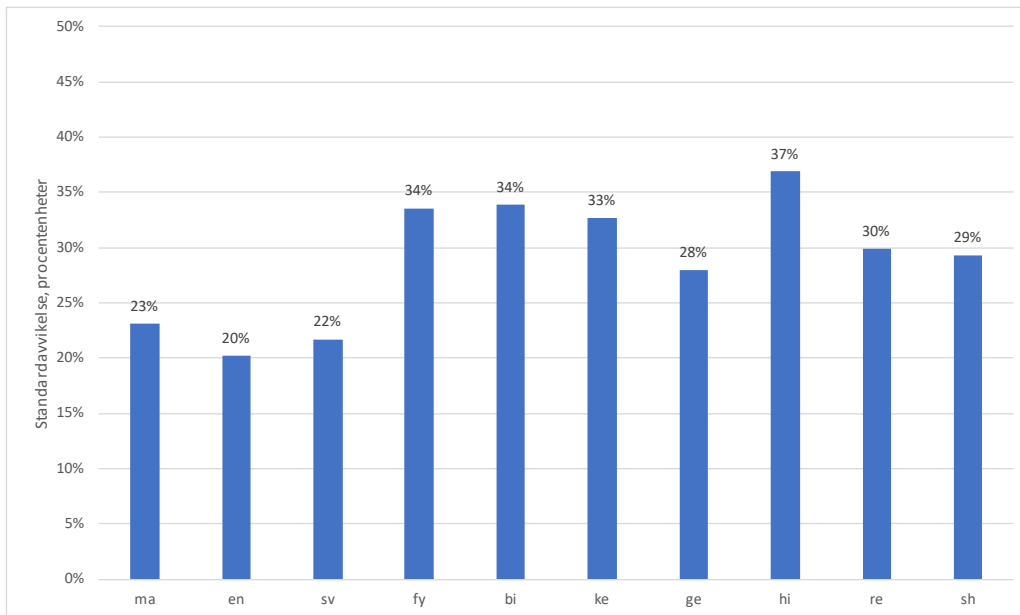
Figur 4a–c visar att det finns en stor variation mellan skolor när det gäller betygsättning i relation till de nationella proven. I matematik finns en skola med en nettoavvikelse på hela 200 procent. Det kan tolkas som att eleverna på den skolan i genomsnitt fått ett betyg som är hela två betygssteg högre än deras betyg på det nationella provet. I andra änden finns en skola som har en negativ nettoavvikelse på 23 procent, vilket kan tolkas som att på den skolan fick 23 procent av eleverna ett lägre betyg än provbetyg, medan övriga elever fick samma betyg som provbetyg.⁴⁹

Figuren för ämnet engelska visar att det finns minst lika många skolor med negativa nettoavvikelser som skolor med positiva nettoavvikelser. Men det finns fortfarande en stor variation i nettoavvikelser mellan olika skolor. Högst nettoavvikelse är en skola med 157 procent och lägst en skola med -83 procent. Även i svenska syns en stor variation i skolors nettoavvikelser och den varierar från en skola med 116 procent i nettoavvikelse till en skola med -55 procent.

⁴⁹ Detta är en hypotetisk tolkning. I verkligheten kan samma nettoavvikelse erhållas med andra kombinationer av elevers individuella avvikelser. Till exempel kan en skolas nettoavvikelse på 10 procent erhållas genom att 10 procent av eleverna får ett steg högre betyg än provbetyg och övriga 90 procent får samma betyg, eller genom att 20 procent av eleverna får ett steg högre betyg medan 10 procent får ett steg lägre och övriga 70 procent får samma betyg som provbetyg. Varje skolas avvikelser mellan betyg och nationella prov finns redovisad på Skolverkets webbplats: <https://www.skolverket.se/skolutveckling/statistik>

Figur 4a–c visar tydligt att det finns stora skillnader i skolors betygssättning i förhållande till resultaten på de nationella proven. För att få ett mer samlat mått på dessa skolskillnader redovisas i figur 5 spridningsmättet standardavvikelsen i skolors nettoavvikelser, genomsnittligt för åren 2013–2015 för respektive ämne.

Figur 5. Variation i skolors nettoavvikelser, genomsnitt 2013–2015.



Spridningsmättet är standardavvikelsen i skolornas nettoavvikelse mellan betyg och provbetyg (utifrån betygsssteg), uttryckt i procent. Standardavvikelsen kan tolkas som hur mycket skolorna i genomsnitt avviker från det nationella medelvärde av nettoavvikelsen. Endast skolor med minst 15 elever i respektive ämne ingår i analyserna. Data är viktad efter skolorlek och justerad för provbetygs- och fördelningseffekter.

Figur 5 visar att skillnaderna i skolors betygssättning, utifrån skolornas nettoavvikelser mellan betyg och nationella prov, är minst i matematik, svenska och engelska och störst i no- och so-ämnena. Senare i kapitlet redovisas utvecklingen av nettoavvikelserna över tid.

Skillnader i betygssättning mellan olika kategorier av skolor

Förra avsnittet visade att det finns betydande skillnader mellan skolors betygssättning i förhållande till de nationella proven för samtliga ämnen. I detta avsnitt undersöker vi skillnaden i betygssättning i relation till de nationella proven för två olika kategoriseringar av skolor: skolor med kommunal respektive fristående huvudman samt mellan skolor som skiljer sig åt i prestationsnivå.

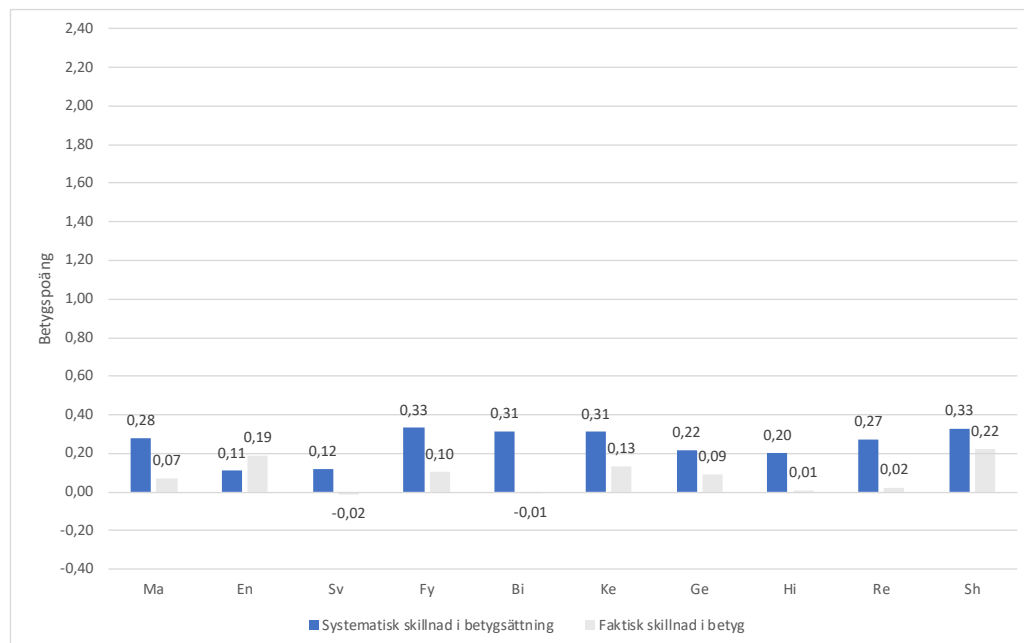
På samma sätt som när det gäller skillnader mellan elevgrupperna i kapitel 2 har vi i analyserna tagit hänsyn till skolornas olika sammansättning när det gäller elevernas socioekonomiska bakgrund, migrationsbakgrund etc.

Skillnader i betygssättning mellan fristående och kommunala skolor

De mörkblå staplarna i figur 6 beskriver liksom tidigare skillnaderna i betygssättning i relation till de nationella proven. Ett positivt värde innebär högre betyg i relation till de nationella proven för elever i fristående skolor. Skillnaden är uttryckt i betygspoäng (0–20).

De ljusgrå staplarna visar faktiska skillnader i betyg (uttryckt i betygspoäng) mellan elever i fristående skolor och kommunala skolor efter kontroll för övriga variabler.

Figur 6. Skillnader i betygssättning i förhållande till de nationella proven mellan elever i fristående och kommunala skolor samt faktiska skillnader i betyg mellan samma elever. Uttryckt som genomsnittligt betygspoäng, genomsnitt för åren 2013–2015.



Positiva värden innebär att elever i fristående skolor får högre betyg i förhållande till de nationella proven (mörkblå staplar) respektive erhåller högre faktiska betyg (ljusgrå staplar), givet att alla övriga bakgrundsfaktorer är lika. Betygspoängen kan variera mellan 0 och 20 i varje ämne. Justerad för provbetygs- och fördelningseffekter. Samtliga skillnader i betygssättning är statistiskt signifikanta (5 procent) utifrån approximativa standardfel (se bilaga B1.4).

Figur 6 visar (mörkblå staplar) att fristående skolor i genomsnitt sätter högre betyg i förhållande till provbetygen jämfört med kommunala skolor, efter kontroll för övriga variabler. Denna skillnad i betygssättning syns i samtliga ämnen. Skillnaden är något lägre i svenska och engelska jämfört med övriga ämnen. En närmare analys uppdelad på olika provbetyg visar att elever med ett underkänt provbetyg (F) i särskilt hög utsträckning får ett högre betyg (E) på fristående skolor.⁵⁰

Elever i fristående skolor får i genomsnitt något högre betyg än elever i kommunala skolor, efter kontroll för övriga bakgrundsvariabler (ljusgrå staplar). Skillnaderna i betygssättning (mörkblå staplar) är större än skillnaderna i faktiska betyg i alla ämnen utom i engelska. Det betyder att om hänsyn togs till skillnader i betygssättning skulle istället elever i kommunala skolor få något högre betyg än elever i fristående skolor, efter kontroll för övriga bakgrundsvariabler. Vi återkommer till en samlad bild i figur 10.

Ovanstående resultat bygger på genomsnitt och bakom dessa genomsnitt finns en stor variation. Det innebär att det finns många fristående och kommunala skolor som uppvisar såväl låga som höga avvikelser i betygssättningen i relation till de nationella proven. Bland kommunala skolor är spridningen (mätt som standardavvikelsen) i skolors nettoavvikelse i exempelvis matematik 21,3 procentenheter medan den är 24,6 procentenheter bland fristående skolor. Spridningen i skolors nettoavvikelse är större mellan fristående skolor för samtliga ämnen.⁵¹

Skillnader i betygssättning mellan låg- och högpresterande skolor

Den andra indelningen av skolor vi gör är utifrån skolornas olika prestationsnivå. Det vi vill undersöka med denna uppdelning är om det förekommer *relativ betygssättning*, det vill säga om lärarna sätter betyg i relation till prestationsnivån hos de elever de själva har. Vi har inte uppgifter om vilka elever lärarna har så vi får nöja oss med att jämföra prestationsnivån på skolan. En fråga är om och i så fall i vilken utsträckning lärare sätter betyg i förhållande till prestationsnivån på skolan – är det svårare att, givet att eleven fått ett visst betyg på nationella provet, få ett högre betyg på en skola med högre genomsnittlig prestationsnivå?

Skolans prestationsnivå mäts med det genomsnittliga provbetygsresultatet i respektive ämne. Liksom för socioekonomisk bakgrund, bygger måttet vi använder för att undersöka relativ betygssättning på en kontinuerlig variabel, skolans genomsnittliga provbetyg i det aktuella ämnet. Detta innebär att det inte finns två givna grupper att jämföra. Vi använder återigen standardavvikelsen – vilket innebär att vi jämför betygssättningen mellan två hypotetiska elever med samma

⁵⁰ Bilaga B2.1.

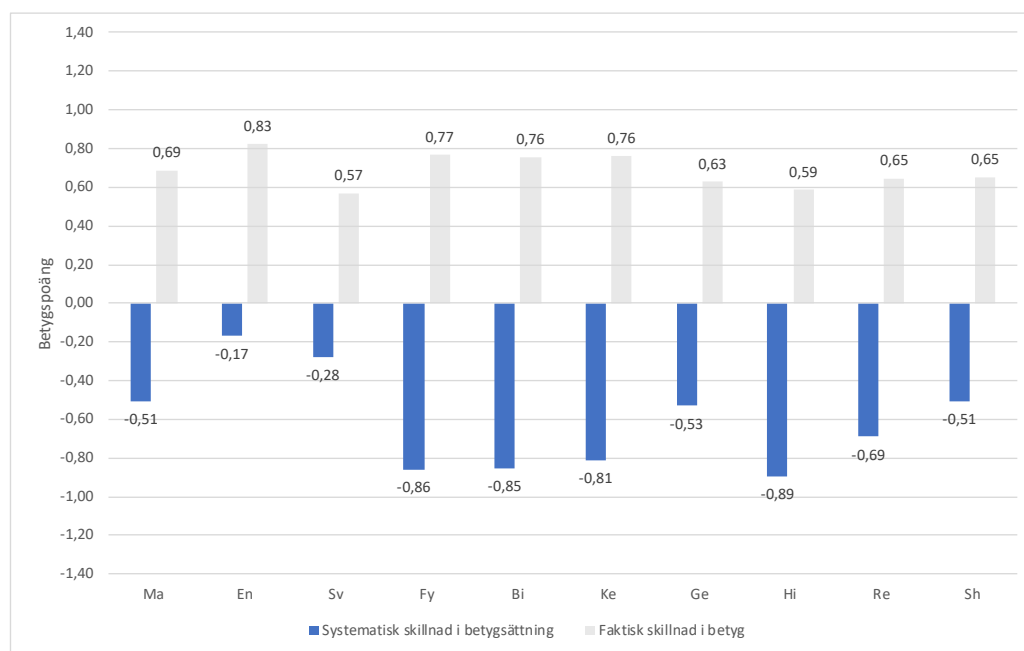
⁵¹ Även inom de olika fristående skolkoncernerna finns stora skillnader i betygssättning i relation till de nationella proven mellan skolor tillhörande samma koncern. Bilaga B2.4.

bakgrundsegenskaper men som går på olika skolor och där det skiljer en standardavvikelse i genomsnittligt provbetyg mellan skolorna.⁵²

De mörkblå staplarna i figur 7 beskriver liksom tidigare skillnaderna i betygssättning i relation till de nationella proven. Ett positivt värde innebär högre betyg i relation till de nationella proven för elever som går i skolor med en standardavvikelse högre provbetyg. De ljusgrå staplarna är de faktiska skillnaderna i betyg mellan skolor som har en standardavvikelse högre provbetyg.

Även här är skillnaderna uttryckta i betygspoäng (0–20). För båda staplarna är skillnaderna som tidigare efter kontroll för övriga variabler (kön, migrationsbakgrund, socioekonomisk bakgrund och huvudman).

Figur 7. Skillnader i betygssättning i förhållande till de nationella proven för elever i skolor som skiljer sig med en standardavvikelse i genomsnittligt provbetyg samt faktiska skillnader i betyg mellan samma elever. Uttryckt som genomsnittlig betygspoäng, genomsnitt för åren 2013–2015.



Negativa värden på de mörkblå staplarna innebär att elever som går i skolor med en standardavvikelse högre provbetyg får lägre betyg i förhållande till de nationella proven. Positiva värden på de ljusgrå staplarna innebär att elever som går i skolor med en standardavvikelse högre provbetyg får högre faktiska betyg, givet att alla övriga bakgrundsfaktorer är lika. Betygs-poängen kan variera mellan 0 och 20 i varje ämne. Justerad för provbetygs- och fördelningseffekter. Samtliga skillnader i betygssättning är statistiskt signifikanta (5 procent) utifrån approximativa standardfel (se bilaga B1.4).

⁵² Jämförelsen av standardavvikelsen görs över hela skalan för nivån på provbetygen, inte bara i mitten av skalan. I matematik motsvarar en standardavvikelse i skolornas genomsnittliga provbetyg ungefär 2 betygspoäng medan det i engelska motsvarar ungefär 1,5 betygspoäng.

Figur 7 visar att det finns påtagliga effekter av relativ betygssättning i samtliga ämnen. På skolor som ligger på en högre genomsnittsnivå i de nationella proven är betygssättningen lägre i förhållande till de nationella proven (mörkblå staplar). I fysik är till exempel skillnaden i betyg för en elev som går i en skola som har en standardavvikelse högre provresultat -0,86 betygspoäng, jämfört med en elev som går på en skola med en standardavvikelse lägre provresultat.⁵³ Effekterna av relativ betygssättning är störst i no-ämnena, följt av so-ämnena och minst i svenska och engelska, men även i svenska och engelska är inslaget av relativ betygssättning inte obetydligt med skillnader på -0,28 respektive -0,17 betygspoäng till nackdel för elever i skolor med högre genomsnittligt provresultat.⁵⁴

I vissa ämnen är den relativa betygssättningen lite högre och i andra lite lägre i förhållande till de faktiska betygsskillnaderna (ljusgrå staplar). Vi återkommer till en samlad bild i figur 10.

En närmare analys uppdelad på provbetyg visar att den relativa betygssättningen finns för samtliga provbetyg men är särskilt tydlig för elever som fått F i provbetyg.⁵⁵

Hur mycket av skolskillnaderna i betygssättning kan förklaras av skolfaktorer?

I detta kapitel har visats att det finns betydande skillnader i hur skolorna sätter betyg i förhållande till resultaten på de nationella proven. Det visades också att betygssättningen i förhållande till de nationella proven i genomsnitt är mer generös på fristående skolor, allt annat lika. Dessutom visades att betygssättningen skiljer sig systematiskt mellan skolor med genomsnittligt höga och låga provresultat, det vi kallar relativ betygssättning.

I detta avsnitt undersöks *hur mycket* av dessa totala skillnader i hur skolorna sätter betyg i förhållande till resultaten på de nationella proven, som kan förklaras av relativ betygssättning respektive skillnader i betygssättning mellan kommunala och fristående skolor. Dessutom undersöks om och i vilken utsträckning lärarbehörigheten på skolorna kan förklara något av dessa skillnader i skolors betygssättning.

Relativ betygssättning

I figur 8 visas hur stor *andel* av de totala skillnaderna i betygssättning mellan skolor (som visades i figur 4a–c och 5) som kan förklaras av relativ betygssättning. Det är tyvärr inte möjligt att räkna ut en exakt andel så därför redovisas här ett undre golv

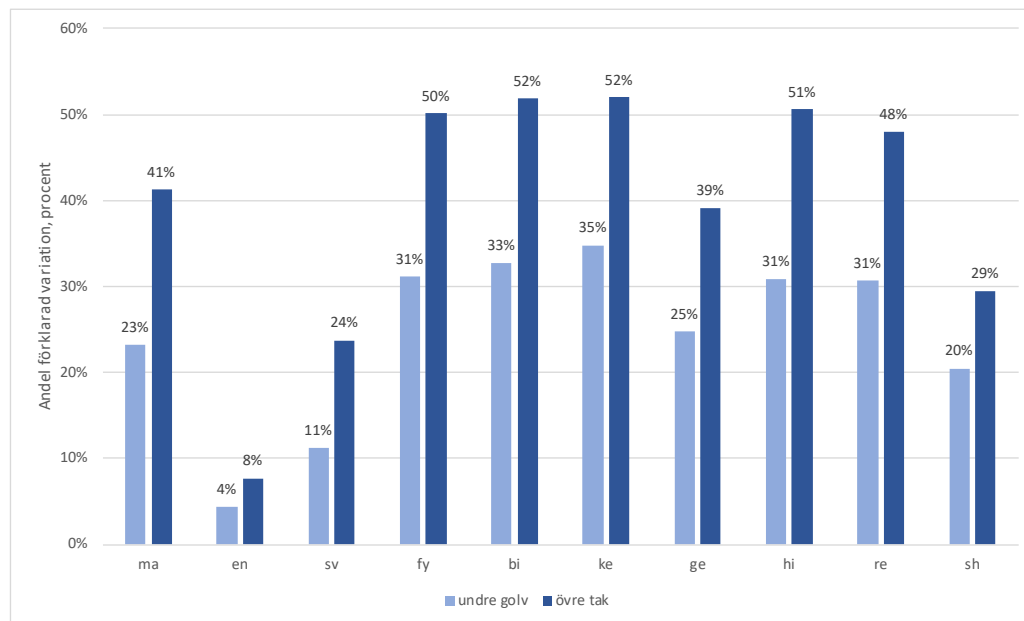
⁵³ I fysik motsvarar en standardavvikelse en skillnad på drygt 2 betygspoäng mellan skolornas genomsnittliga betygspoäng.

⁵⁴ Nivån på betygspoängen ska inte jämföras med nivån för huvudman eftersom skillnaden mellan skolor här är beräknad på en standardavvikelse – inte mellan två distinkta kategorier.

⁵⁵ Bilaga B2.3.

(ljusblå staplar) respektive ett övre tak (mörkblå staplar) för andelen av de totala skillnaderna i betygssättning som kan förklaras av relativ betygssättning.⁵⁶

Figur 8. Beräknade andelar (undre golv respektive övre tak) av de totala skillnaderna i skolors betygssättning som kan förklaras av relativ betygssättning. Genomsnitt 2013–2015.



Se bilaga B1.3 för mer information om hur analysen gått till.

Figur 8 visar att för no-ämnena och för historia och religionskunskap kan ungefär mellan 30–50 procent av skillnaderna i skolornas betygssättning förklaras med relativ betygssättning. För matematik, geografi och samhällskunskap ligger förklaringsgraden ungefär mellan 20–40 procent. För svenska ligger förklaringsgraden av relativ betygssättning mellan ungefär 10–25 procent. I engelska kan endast 4–8 procent av de totala skolskillnaderna i betygssättning förklaras av relativ betygssättning. Att nivån i engelska och svenska ligger lägre hänger ihop med att det var dessa ämnen som i figur 7 visade på lägst betydelse av relativ betygssättning. Sett över samtliga tio ämnen innebär detta att någonstans mellan 25–40 procent av de totala skolskillnaderna kan förklaras med prestationsnivån på skolorna, det vi kallar relativ betygssättning.

⁵⁶ Analysen görs med variansdekomposition där vi utnyttjar två olika modeller, en golvmodell respektive en takmodell, som ger en underskattning- respektive en överskattning av andelen som kan förklaras av faktorerna på skolnivå. Se bilaga B1.3 för en mer detaljerad redovisning av metoden och resultaten.

Kommunala och fristående skolor

Till skillnad från relativ betygssättning, kan skillnader i betygssättning mellan fristående och kommunala skolor endast förklara omkring 1,5 procent av de totala skillnaderna i betygssättning mellan skolor.⁵⁷ Då andelen elever som i vår analys går i fristående skola är relativt låg, cirka 16 procent, i kombination med att den systematiska avvikelsen i betygssättning mellan kommunala och fristående skolor är förhållandevis begränsad, är det inte konstigt att förklaringsgraden för de totala skolskillnaderna är så pass liten. Men även om andelen elever i fristående skolor skulle varit betydligt större, till exempel 50 procent, givet samma systematiska avvikelse i betygssättning, skulle förklaringsgraden sannolikt inte vara i närheten av den som syns för relativ betygssättning.

Andelen behöriga lärare på skolan

Vi har undersökt om lärarbehörighet har någon betydelse för betygssättningen utifrån hypotesen att lärarbehörighet ger en ökad säkerhet i bedömningen. Tyvärr har vi inte tillgång till uppgifter om vilka elever som lärarna undervisar varför vi istället har undersökt den genomsnittliga lärarbehörigheten på varje skola. Hypotesen bekräftas – för flertalet ämnen så är betygssättningen i förhållande till de nationella proven mer generös för skolor med en lägre lärarbehörighet, men betydelsen är mycket liten.⁵⁸ Det gör att skillnaden i lärarbehörighet mellan skolor endast kan förklara en bråkdel, cirka 0,5 procent, av de totala skillnaderna i hur skolorna sätter betyg i förhållande till resultaten på de nationella proven.⁵⁹

Hur har skolskillnaderna i betygssättning förändrats över tid?

Vi har visat att skillnaderna mellan skolors betygssättning i förhållande till de nationella proven är stora. Skillnaderna är stora utifrån perspektivet att vi förväntar oss att betygssättningen i relation till de nationella proven bör vara ungefär likartade mellan skolor för att man ska tro att betygen sätts på ett likvärdigt sätt. Men något objektivt mått på vad som är stora och små avvikelser mellan skolors betygssättning i förhållande till de nationella proven finns inte. Ett annat sätt att undersöka likvärdigheten i betygssättning mellan skolor är därför att undersöka utvecklingen över tid. Om skillnaderna mellan skolors betygssättning i relation till de nationella proven minskar så kan man tolka det som att lärarna med tiden når en mer likvärdig tolkningsgrund för betygssättningen.

Skolverket har publicerat analyser av skolors betygssättning i relation till de nationella proven över tid vid flera tillfällen.⁶⁰ I figur 9a–c visas hur skolskillnaderna i betygssättning i relation till proven utvecklats åren 2005–2017.

⁵⁷ Bilaga B2.5.

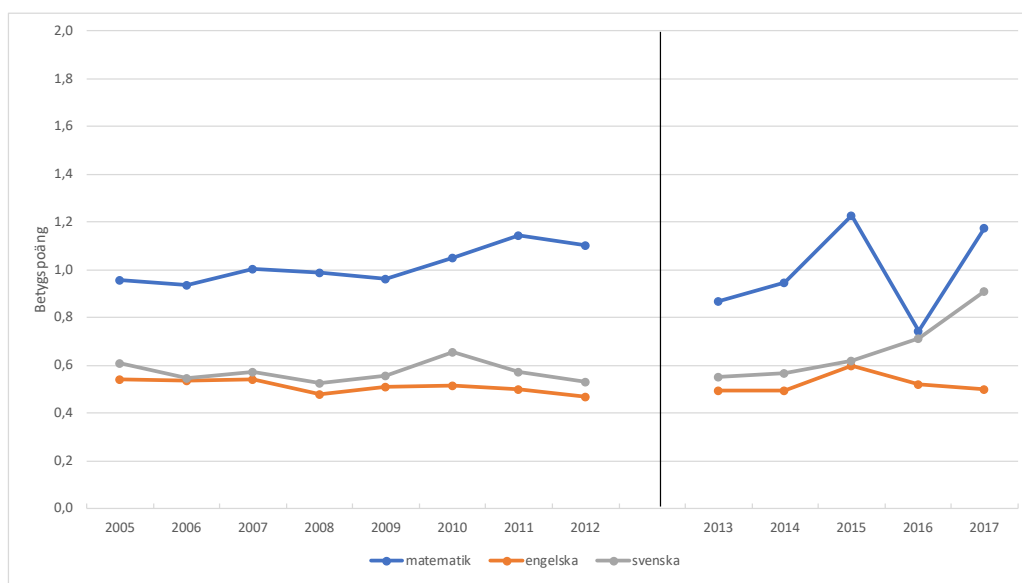
⁵⁸ Bilaga B2.7

⁵⁹ Fristående skolor har i genomsnitt en högre andel obehöriga lärare. Det är en delförklaring bakom att fristående skolor har en mer generös betygssättning men det är inte hela förklaringen.

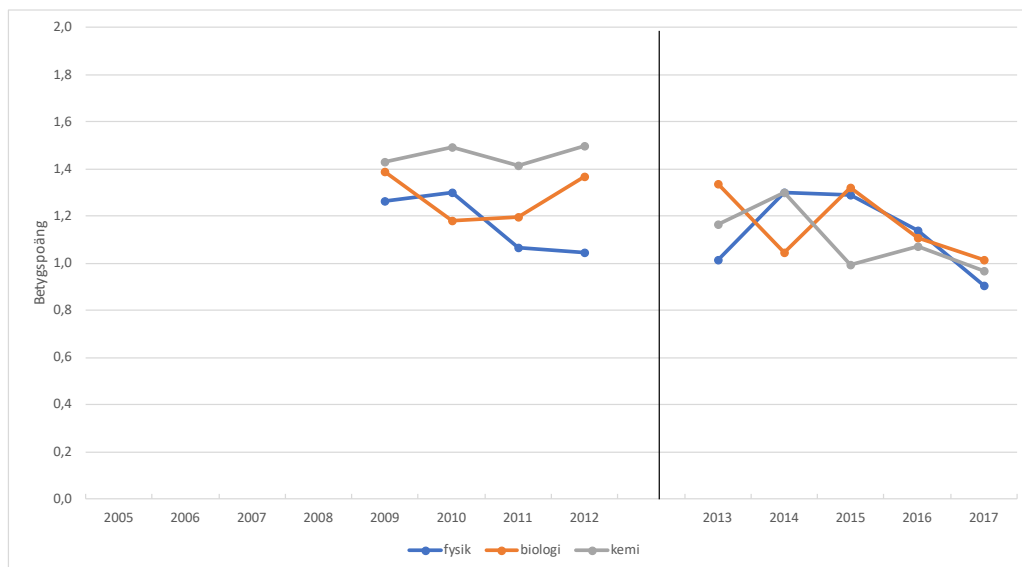
⁶⁰ Se t ex Skolverket (2007) kap 4, Skolverket (2016a) s.80ff. Det görs dessutom årliga beskrivningar av utvecklingen över tid, för den senaste beskrivningen, se Skolverket (2019).

Skillnaden mellan skolornas betygssättning är, precis som i figur 5, beskrivet med spridningsmåttet standardavvikelse men skillnaden är här uttryckt i betygspoäng som kan variera mellan 0 och 20 per ämne.

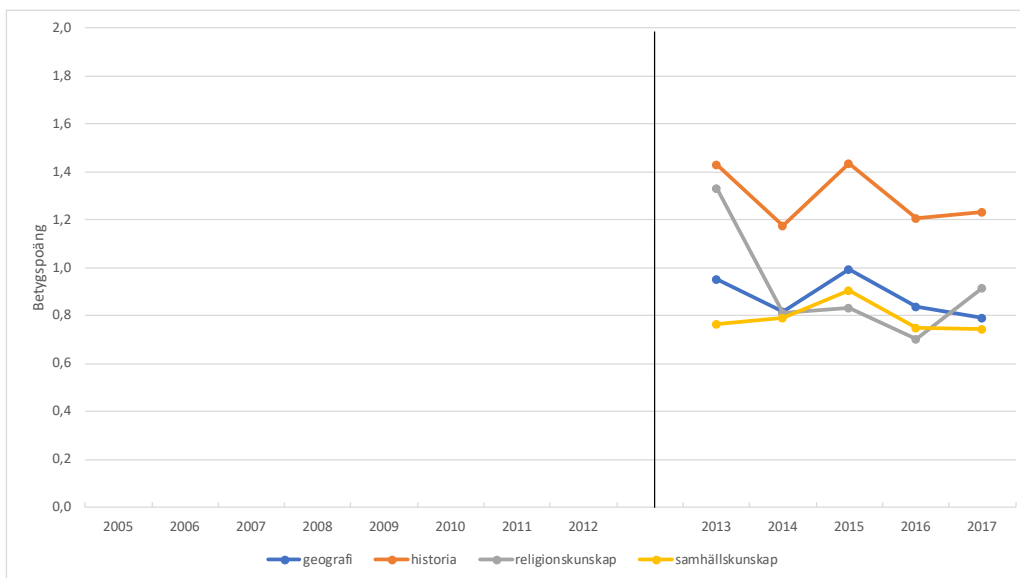
Figur 9a. Variation i skolors nettoavvikelse mellan betyg och provbetyg, uttryckt som standardavvikelse i betygspoäng, för matematik, engelska och svenska, åren 2005–2017.



Figur 9b. Variation i skolors nettoavvikelser mellan betyg och provbetyg, uttryckt som standardavvikelse i betygspoäng för fysik, biologi och kemi, åren 2009–2017.



Figur 9c. Variation i skolors nettoavvikelser, uttryckt som standardavvikelse i betygspoäng för geografi, historia, religionskunskap och samhällskunskap, åren 2013–2017.



Här används standardavvikelsen i betygspoäng istället för i betygsssteg i syfte att minimera effekter av bytet av betygsskala 2013. Betygspoängen kan variera mellan 0 och 20 i varje ämne.

Eftersom vi inte kan garantera att bytet av betygsskala inte påverkar våra tidserier så bör varje tidsperiod (2005–2012 respektive 2013–2017) analyseras separat. Mellan 2005 och 2012 syns ingen trend över huvud taget i matematik, engelska och svenska. Från 2013 till 2017 kan en viss ökning i skolskillnader skönjas i svenska.

I no-ämnena varierar skillnaderna mer men en viss nedåtgående trend kan skönjas, om man bortser från att bytet av betygsskala 2013 kan ha påverkat måttet.

I so-ämnena syns ingen trend överhuvudtaget.

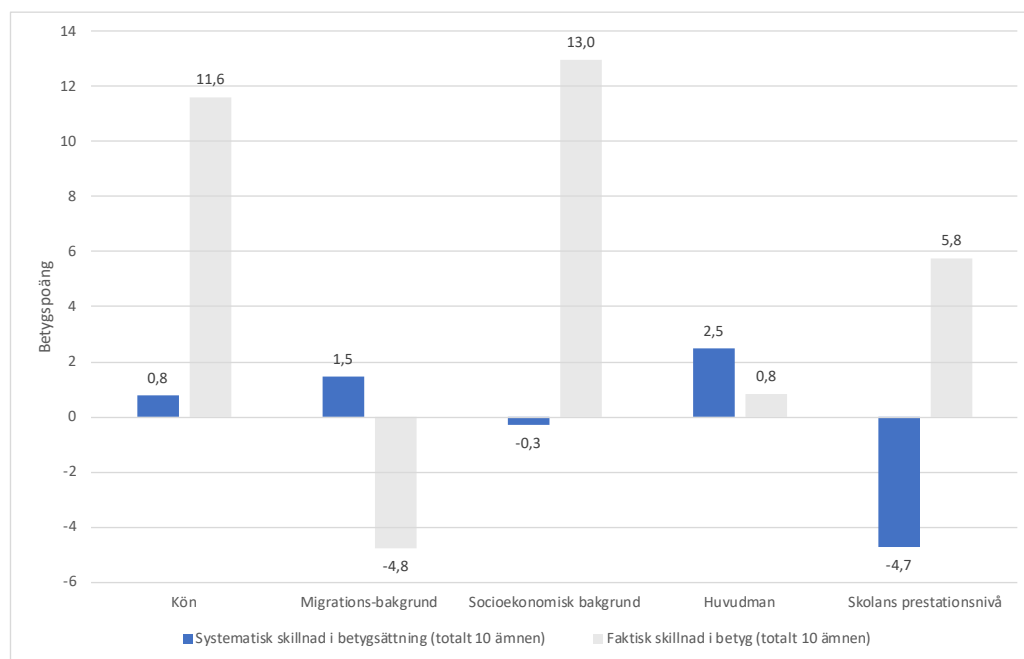
Sammantaget är bilden att inget tyder på att samstämmigheten i betygssättning mellan lärarna på landets skolor är på väg att förbättras på något avgörande sätt.

4. Sammanfattat resultat

I detta kapitel sammanfattas resultaten genom att vi presenterar en översikt av betygssättningen i relation till de nationella proven för elevgrupperna och skolkategorierna.

De mörkblå staplarna i figur 10 representerar den totala skillnaden i betygssättning i relation till de nationella proven summerat för samtliga tio ämnen som har nationella prov.⁶¹ De ljusgrå staplarna representerar de faktiska skillnaderna i betyg för de tio ämnen som har nationella prov. Resultaten är uttryckta i betygspoäng som här kan variera mellan 0 och 200 poäng.⁶²

Figur 10. Sammanfattande jämförelse av skillnader i betygssättning i relation till de nationella proven samt de faktiska skillnaderna i betyg för samma elever. Uttryckt som betygspoäng för samtliga tio ämnen. Genomsnitt åren 2013–2015.



Positiva värden för mörkblå och ljusgrå staplar betyder större *skillnad i betygssättning* respektive *faktisk skillnad i betyg* för: flickor, utlandsfödda, elever med högre socioekonomisk bakgrund, elever i fristående skolor och elever som går i skolor med en högre prestationsnivå.

Skillnad i betygssättning = summering av skillnaderna i betygssättningen i relation till de nationella proven för respektive ämne. Justerad för provbetygs- och fördelningseffekter.

Faktisk skillnad i betyg = den totala skillnaden i betygspoäng mellan kategorierna baserat på de tio ämnen som har nationella prov (max 200 poäng), efter kontroll för bakgrundsvariabler.

⁶¹ Summerat för de mörkblå staplarna i figur 1–3 och 6–7. För att få en rättvisande sammanvägning för relativ betygssättning har det dock gjorts en särskild justering. Se bilaga.

⁶² Tio ämnen med max 20 betygspoäng per ämne ger max 200 poäng.

När det gäller *kön* så ser vi i figur 10 att den faktiska skillnaden i betyg (ljusgrå stapel) är betydande, med 11,6 betygspoäng till flickornas fördel.⁶³ Betygssättning i relation till de nationella proven är, som vi sett i figur 1, till flickornas fördel i vissa ämnen och till pojkarnas fördel i vissa ämnen. Summerat blir skillnaden 0,8 betygspoäng till flickornas fördel. Även om betygssättningen har en viss betydelse så förklarar den bara en mycket liten del av de faktiska betygsskillnaderna som finns till flickornas fördel.

När det gäller *migrationsbakgrund* så har utlandsfödda elever i genomsnitt -4,8 betygspoäng lägre betyg än elever födda i Sverige – efter att vi kontrollerat för socioekonomisk bakgrund etc. Utlandsfödda gynnas något i betygssättningen i förhållande till de nationella proven. Vore det inte för detta skulle skillnaden i betyg gentemot elever födda i Sverige vara 1,5 betygspoäng större. Sammantaget bör betygssättningen till fördel för utlandsfödda elever betraktas som liten.

När det gäller *socioekonomisk bakgrund* undersöks skillnaden i betygssättning mellan elever som skiljer sig med en standardavvikelse i socioekonomisk bakgrund. Den faktiska skillnaden i betygspoäng är 13,0 poäng. Betygssättningen i relation till de nationella proven är mycket marginell, totalt 0,3 betygspoäng till fördel för elever med lägre socioekonomisk bakgrund. En rimlig slutsats utifrån denna analys är att det knappast går att tala om några skillnader i betygssättning utifrån elevers olika socioekonomiska bakgrund.

När det gäller skillnaden mellan *kommunala och fristående skolor* finns däremot en tydlig effekt av betygssättningen i relation till de nationella proven. När vi tagit hänsyn till bakgrundsfaktorer (socioekonomisk-, migrationsbakgrund etc.) så finns en faktisk skillnad i betyg på 0,8 betygspoäng till elever i fristående skolors fördel. Skillnaden i betygssättning i relation till de nationella proven visar att elever i fristående skolor gynnas med 2,5 betygspoäng. Det innebär att om det inte funnits några skillnader i betygssättning så skulle elever i kommunala skolor istället prestera lite bättre än elever i fristående skolor, givet en kontroll för övriga bakgrundsvariabler.

Slutligen har vi resultaten för *relativ betygssättning*. Vi finner här att det finns en betydande betygssättningseffekt som ”håller tillbaka” betygen för elever på skolor med högre prestationsnivå på de nationella proven. Om det inte vore för den relativa betygssättningen så skulle de faktiska betygsskillnaderna vara 4,7 betygspoäng större (ungefär 80 procent större), det vill säga öka från nuvarande 5,8 till 10,5 betygspoängs skillnad för skolor som har en standardavvikelse högre provbetyg.⁶⁴

⁶³ Kontroll för bakgrundsvariabler har en liten betydelse för kön då de av naturliga skäl inte skiljer sig åt mer än marginellt mellan flickor och pojkar.

⁶⁴(5,8) - (-4,7) = 10,5.

5. Diskussion

Betygssättningen av elevgrupper

När det gäller betygssättningen av olika elevgrupper är det övergripande resultatet att vi inte finner några större tecken på att de betygssätts på olika sätt i relation till de nationella proven. När det gäller elever med olika socioekonomisk bakgrund finner vi inga skillnader alls medan vi ser att flickor och utlandsfödda elever får en något mer gynnsam betygssättning i förhållande till provbetygen.⁶⁵

Vi har inte möjlighet att med våra data analysera vad som kan vara möjliga orsaker till dessa små skillnader i betygssättning till flickors och utlandsföddas fördel.⁶⁶ Det viktigaste resultatet är att vi inte ser några större tecken på att någon av de elevgrupper vi undersökt gynnas eller missgynnas i betygssättningen i förhållande till de nationella proven.

Betygssättningen mellan skolor

Om våra analyser visade på marginella skillnader i betygssättning mellan olika elevgrupper så visar analyserna av betygssättningen mellan skolor på ett betydligt större likvärdighetsproblem. Skillnaden mellan skolors betygssättning i relation till de nationella proven är betydande. Det finns gott om skolor där en stor andel av eleverna får högre betyg än provbetyg i ett ämne och det finns samtidigt skolor där en stor del av eleverna får lägre betyg än provbetyg i samma ämne. Rapportens utgångspunkt var att betygssättningen i relation till de nationella proven bör vara ungefär likartad mellan skolor och då kan inte de skolskillnader vi ser anses uppfylla kravet på en likvärdig betygssättning. Tidigare analyser har också visat att elever som gått i skolor med generös betygssättning i förhållande till de nationella proven i genomsnitt också klarar sig sämre på gymnasiet jämfört med elever som fått samma betyg men gått på skolor med en mindre generös betygssättning.⁶⁷ De skolskillnader vi ser visar därför på ett verkligt likvärdighetsproblem i betygssättningen.

Det är samtidigt svårt att definiera en objektiv nivå för hur stora betygsskillnaderna mellan skolor ska vara för att man ska anse att det finns stora likvärdighetsproblem i betygssättningen. Därför undersöktes även skolskillnader i betygssättning över tid för att se om likvärdigheten blivit bättre eller sämre. Analysen visar att skolskillna-

⁶⁵ Jmf Lindahl (2007) som med en liknande metod kommer fram till ett liknande resultat. Diamond & Persson (2016) använder en annan metod och finner inga skillnader i betygssättning mellan kön. Lindahl (2016) analyserar om lärarens kön har någon betydelse för betygssättningen. För resultat från andra länder se Ângelo & Balcão Reis (2018) och Kiss (2011).

⁶⁶ Se det inledande avsnittet *Betyg och nationella prov som mått på kunskaper* i kapitel 1 för en diskussion av vad forskning och analyser säger om vad som kan ligga bakom skillnader i betygssättning och provresultat mellan elevgrupper.

⁶⁷ Skolverket (2007) kap 4.3.

derna i betygssättning i relation till de nationella proven inte förändrats på något avgörande sätt under tidsperioden 2005–2017. Det finns en del som talar för att man skulle kunna förvänta sig minskade skolskillnader över tid. Man kan tänka sig att lärarna på olika skolor skulle vänja sig vid det nya betygssystemet som började tillämpas 1998 och med tiden nå en mer likartad bedömning. Vidare så har den bristande likvärdigheten i betygssättning uppmärksammats många gånger före och under tidsperioden och det har genomförts en rad insatser. Det finns många ambitiösa lokala insatser av olika huvudmän och ute på olika skolor för att förbättra likvärdigheten i betygssättningen. På den nationella nivån har det gjorts flera stora insatser som helt eller delvis syftat till en ökad likvärdighet i betygssättningen, exempelvis en ny betygsskala, nya kunskapskrav, enbart behöriga lärare ska få sätta betyg, mer fortbildning av lärare när det gäller betygssättning, ökade informationsinsatser om betygssättning, offentlig publicering av relationer mellan betyg och nationella prov för varje skola samt ökad inspektion av betygssättningen.⁶⁸ Om dessa insatser haft effekt kan vi inte svara på – utan dessa insatser kanske vi fått ökande skillnader mellan skolors betygssättning.

När det gäller skillnaden mellan huvudmän finns tydliga tecken på att betygen sätts mer generöst i förhållande till de nationella proven i fristående skolor.⁶⁹ Den mer generösa betygssättningen i relation till de nationella proven överstiger den fördel som elever i fristående skolor har i faktiska betyg över elever i kommunala skolor, efter att vi tagit hänsyn till skolornas elevsammansättning. Elever i kommunala skolor skulle alltså få något högre betyg än elever i fristående skolor i de tio ämnen som har nationella prov om det inte vore för den mer generösa betygssättningen för elever i fristående skolor. Den mer generösa betygssättning, till elever i fristående skolors fördel, är större i no- och so-ämnena, som inte haft nationella prov i så många år, än i matematik, svenska och engelska. Det finns även forskning som tyder på att betygen i icke-provämnen sätts ytterligare något mer generöst i fristående skolor.⁷⁰ Det finns dessutom analyser som tyder på att bedömningen av de nationella proven är mer generös i fristående skolor.⁷¹ Därmed finns en risk att de

⁶⁸ Se t ex Skolverket (2016a) s.22ff. Riksrevisionen (2011). Skolinspektionen (2018).

<https://www.skolverket.se/undervisning/grundskolan/bedomning-i-grundskolan>
<https://www.skolverket.se/undervisning/grundskolan/betyg-i-grundskolan/satta-betyg-i-grundskolan>

Se t ex Skolverket (2016a) s.111ff, SOU 2016:25, Skolverket (2016b), Gustafsson & Ericksson (2018) för en diskussion och förslag om nya åtgärder för en mer likvärdig betygssättning och de nationella provens roll.

⁶⁹ Vlachos (2018) finner också en mer generös betygssättning i fristående skolor. Även Lekholm (2011) finner med en annan metodansats en mer generös betygssättning för fristående skolor i vissa ämnen. Jmf Skolverket (2018a) s.30ff som med en annan metod finner liknade resultat på gymnasienivå.

⁷⁰ Vlachos (2018).

⁷¹ Tyrefors Hinnerich & Vlachos (2013). För en motsvarande analys för gymnasieskolan se Tyrefors Hinnerich & Vlachos (2017).

effekter som påvisas i denna rapports analyser kan vara en underskattning av den samlade skillnaden i betygssättning mellan kommunala och fristående skolor.

Samtidigt ska poängteras att det finns en mycket stor variation i betygssättning mellan skolor både inom gruppen fristående och inom gruppen kommunala skolor, vilket innebär att det finns många fristående skolor som är restriktiva i betygssättningen samtidigt som det finns många kommunala skolor som är generösa i betygssättningen. Våra analyser visar också att skillnaderna i betygssättning mellan huvudmän förklarar en mycket liten del (omkring 1,5 procent) av de totala skillnaderna i betygssättning mellan skolor, delvis beroende på att elever i fristående skolor i våra analyser endast utgör cirka 16 procent av samtliga elever. Det generella mönstret är ändå tydligt, elever som går i fristående skolor får i genomsnitt en mer generös betygssättning i förhållande till sina prestationer på de nationella proven. Men samtidigt är betygssättningen mellan huvudmän långt ifrån det största problemet när det gäller likvärdig betygssättning.

Det tydligaste resultatet rör relativ betygssättning. Det finns en klar tendens att betygen sätts i relation till prestationsnivån på skolorna.⁷² Vi såg i figur 8 att någonstans mellan 25–40 procent av de totala skillnaderna i betygssättning mellan skolor kunde förklaras med att betygen sätts i relation till skolornas prestationsnivå, det vi kallar relativ betygssättning. Vi såg i figur 10 att den faktiska skillnaden i betyg mellan skolor med en standardavvikelse högre provbetyg skulle öka från 5,8 till 10,5 betygspoäng om det inte vore för förekomsten av relativ betygssättning. Det innebär att betygsskillnaderna mellan hög- och lågpresterande skolor skulle vara nästan 80 procent större om det inte vore för den relativa betygssättningen. Vi har också sett att den relativa betygssättningen är minst lika stor om man enbart analyserar fristående skolor och till och med mellan skolor inom samma skolkoncerner. Detta är sammantaget starka indikationer på att det avgörande systemproblemet vad gäller likvärdig betygssättning sannolikt har att göra med att lärarna på olika skolor har problem att hitta gemensamma tolkningsgrunder för betygen. I inte obetydlig utsträckning anpassar lärarna betygssättningen till prestationsnivån hos eleverna på den egna skolan.

⁷² Forskning och analyser av relativ betygssättning är begränsad. Våra resultat är emellertid i linje med en analys som, med en annan metod, fann tydliga tecken på relativ betygssättning under tidsperioden 1990–2008, Fredriksson & Vlachos (2011) s.30f. Denna tidsperiod omfattar alltså även de sista åren (1990–97) med det tidigare normrelativa betygssystemet och även för den perioden fann forskarna en ungefär lika stor relativ betygssättning. Det bör dock påpekas att det sedan Läroplanen 1980 (Lgr 80) inte längre fanns en lika strikt koppling mellan standardprov och betyg som tidigare, se SOU 2016:25 s.135, Kilpatrick & Johansson (1994) s.24f. Se även Calsamiglia & Loviglio (2018) för en analys av relativ betygssättning i Katalonien. Det finns dessutom anledning att misstänka att den relativa betygssättningen är något starkare än vad vi visat här. Vi har utgått från prestationsnivån (provbetygen) på skolan och inte i klasserna. Det är rimligt att anta att lärarna sätter betyg mer i relation till eleverna i sina klasser jämfört med i relation till andra elever på skolan som läraren inte undervisar. Om det dessutom förekommer ”relativ” bedömning (rättning) av proven utifrån en skolans prestationsnivå vet vi inte då detta inte undersökts.

De resterande skillnaderna i betygssättning mellan skolorna, cirka 60–75 procent, kan alltså inte förklaras med relativ betygssättning. Vad kan ligga bakom dessa resterande skolskillnader i betygssättning? Det kan vi inte säga med någon säkerhet men det finns inget som utesluter att även dessa skillnader kan orsakas av svårigheter för lärarna att hitta gemensamma tolkningsgrunder i betygssättningen, även om det också kan finnas andra tänkbara förklaringar.⁷³

Att vi ser sådana stora skolskillnader i betygssättning och så tydliga effekter av relativ betygssättning har konsekvenser för skolsystemet. En konsekvens är att det uppstår en omotiverad skillnad så att betygen inte ger en rättvisande bild av elevernas kunskaper i ämnet. Det uppstår en osäkerhet om vad betygen egentligen säger om skolors resultat och elevers verkliga kunskaper. Är en skolas resultat så goda som betygen visar eller är det betygssättningen som är generös?

Ett särskilt problem som uppstår som en följd av resultatet om en relativ betygssättning gäller analyser av resultatskillnader mellan skolor. När vi analyserar resultatskillnader mellan skolor så leder förekomsten av relativ betygssättning till en underskattning av resultatskillnaderna i de ”faktiska” kunskaperna mellan skolor. Om detta påverkar trendanalyserna, där vi kunnat se ökande resultatskillnader (mellanskolvariation) mellan skolor över tid är mer osäkert.⁷⁴

En annan konsekvens är att den relativa betygssättningen skapar en osäkerhet om hur elevers betygsresultat påverkas av att gå i skolor med olika elevsammansättning. Får eleverna lägre betyg, givet provresultaten, om de går i en skola med hög prestationsnivå, det vill säga har skolor med en högre prestationsnivå en mindre generös betygssättning? Rapportens resultat tyder på det finns ett sådant samband i genomsnitt, även om det inte betyder att alla skolor med hög prestationsnivå har en restriktiv betygssättning. Samtidigt har Skolverket i en tidigare rapport konstaterat att elevers betygsnivå påverkas positivt om de går i en skola med en mer gynnsam socioekonomisk sammansättning jämfört med en skola med en mindre gynnsam sammansättning (det vi kallar skolnivåeffekter).⁷⁵ Även detta resultat gäller för ett genomsnitt och inte för alla skolor med en mer gynnsam socioekonomisk sammansättning. Hur ska man se på dessa två olika resultat?

De två olika resultaten är inte baserade på samma indelning av skolorna. Den relativa betygssättningen analyseras utifrån skolors olika provbetygsnivåer och skolnivåeffekterna analyseras utifrån skolors olika socioekonomiska samman-

⁷³ För en diskussion om tänkbara förklaringar bakom skolskillnader i betygssättning i förhållande till de nationella proven, se Skolverket (2007) s.68ff. Skolverket (2009) s.65ff. Skolverket (2004) s.70ff. Det har även under senare år alltmer uppmärksammats att lärare kan sättas under tryck från elever, föräldrar och rektorer att sätta mer generösa betyg. Om lärare på olika skolor i olika grad ger efter för påtryckningar kan detta också vara en faktor som påverkar skolskillnaderna i betygssättning i förhållande till de nationella proven. Se t ex Lärarnas riksförbund (2011).

⁷⁴ Skolverket (2018c) kap 4. Det ser inte som att den relativa betygssättningen på något avgörande sätt förändrats över tid. Se Bilaga B2.8.

⁷⁵ Skolverket (2018c) s.28ff och 89ff.

sättning. Det är inte samma sak. Men det finns förstås ett starkt samband mellan skolors provbetygsnivå och skolors socioekonomiska sammansättning.⁷⁶ En slutsats är därför att elever i genomsnitt lär sig lite mer om de går i en skola med en mer gynnsam socioekonomisk sammansättning men det är inte lika självklart att denna tillägnade kunskap kommer att återspeglas fullt ut i betygen. Ett annat sätt att uttrycka det är att den genomsnittligt positiva betydelsen för en elevs betyg av att gå i en skola med en mer gynnsam socioekonomisk sammansättning skulle vara större om det inte vore för den motverkande relativa betygssättningen.

Vad är då orsaken bakom de stora skolskillnaderna i betygssättning och de tydliga resultaten om relativ betygssättning? En central slutsats från Skolverkets tidigare granskningar av betygssättningen mellan skolor var att det är svårigheterna för lärare att göra nationellt gemensamma tolkningar av hur olika elevprestationer ska värderas och sedan betygssättas utifrån kursplanernas kunskapskrav (tidigare betygskriterier).⁷⁷ I en rapport från 2004, som baseras på intervjuer och enkäter med lärare, formuleras problemet som att många lärare kan uppleva en relativt stor trygghet i betygssättningen av sina egna elever men en osäkerhet över om andra lärare gör liknande bedömningar:

Likvärdigheten är dock problematisk. Skolverkets kvalitetsgranskning om betyg, Stockholm stads skolinspektörens rapporter och andra lokala rapporter tyder på att det finns betydande brister i likvärdigheten. Skolverkets panelstudier visar emellertid att många lärare själva tycker att de förstår målen och kriterierna och att de är tydliga. Det finns ingen anledning att betvivla att de som säger detta menar vad de säger. Problemet är snarare att det inte är säkert att olika lärare som tycker att mål och kriterier är tydliga sinsemellan är eniga om tolkningen. Så länge den egna tolkningen inte konfronteras med alternativa tolkningar framstår den lätt som tämligen oproblematisks.

Det samtalsbaserade system vi har, där tolkningen av mål och kriterier förutsätter ett ständigt pågående professionellt samtal, ställer stora krav på lärarna och skolorna, inte minst med avseende på den tid som måste avsättas. Ett system som alltså ska förena kraven på lokal målutformning med nationell likvärdighet kräver förutsättningar i form av såväl arenor som tid för ett ständigt fortlöpande professionellt och kollektivt tolkningsarbete.⁷⁸

I Skolverkets utvärdering av de senaste skolreformerna dras liknande slutsatser. Även för de nya kunskapskrav som började gälla i läroplanen 2011 beskrev lärarna med tiden en relativ stor trygghet i betygssättningen av de egna eleverna – men en osäkerhet om tolkningarna stämde med andra lärare på skolan och framförallt med lärare utanför den egna skolan:

I fallstudien framkom att den ökade tryggheten bland annat berodde på att lärarna gjorde sina bedömningar av elevkunskaperna på grundval av erfarenhet av olika elever. Efter en tids bruk av de nya styrdokumentet, hade de sålunda samlat på sig mer erfarenhet av variationen i elevprestationer och hur de kunde komma till uttryck. Många lärare

⁷⁶ Korrelationen mellan skolors genomsnittliga provbetyg (tio ämnen) och skolors socioekonomiska sammansättning (socioekonomiskt index) är ungefär 0,75–0,77 för åren 2013–2015.

⁷⁷ Skolverket (2007) och (2009). Liknande slutsatser dras i Skolverket (2015c) s.12 och Skolverket (2016a) s.6f.

⁷⁸ Skolverket (2004) s.97.

påpekade samtidigt att denna erfarenhet vanligen var begränsad till den egna skolan, och att betygssättningen därmed blev relativ de egna eleverna. De hade själva bildat sig en uppfattning om vad olika betyg krävde, men var inte alltid säkra på att den stämde med övriga Sverige.⁷⁹

Rapportens resultat om stora skolskillnader i betygssättning och relativ betygssättning understryker dessa svårigheter.

Avslutningsvis vill vi betona att rapportens resultat visar på en systemproblematik i betygssystemets utformning. Rapporten visar att lärarna i hög grad förmår att betygssätta eleverna utan att gynna eller missgynna någon elevgrupp. Det är ett krav som är rimligt att ställa och det klarar lärarna väl. Det är däremot orimligt att ställa krav på att en lärare, utifrån kursplanernas kunskapskrav, ska tolka vilka elevprestationer som motsvarar olika betyg på ett sätt som på ett likvärdigt sätt stämmer överens med alla andra lärares tolkningar i alla andra skolor i hela landet. Betygssystemet vilar på kvalitativa bedömningar av kunskapskrav och tolkningar av de värdeord som finns för att särskilja de olika betygsstegen. Det är inte lätt att formulera kunskapskrav om vad eleverna ska kunna i de olika ämnena som alla lärare förstår på ett likartat sätt.⁸⁰ Men även om lärarna skulle uppnå en hög grad av gemensam förståelse av de kunskaper som beskrivs i kunskapskraven ska de ändå särskilja *hur mycket* eleverna kan för att nå olika betygsnivåer. De olika betygsstegen särskiljs genom så kallade värdeord, till exempel: ”enkla” (betyg E), ”utvecklade” (betyg C) och ”välutvecklade” (betyg A) kunskaper inom ett område.⁸¹ Våra resultat tyder på att lärarna i hög grad gör denna bedömning utifrån den egna elevgruppens generella prestationsnivå.

En konsekvens av rådande betygssystem blir därför att lärare sätter betyg utifrån olika förståelser och tolkningar. Ansvar för den bristande likvärdigheten vid betygssättning som rapporten påvisar kan därför inte tillskrivas lärarkollektivet.

Detta utesluter inte en fortsatt strävan att ge lärare förutsättningar att utveckla en ökad samsyn om hur kunskapskraven ska förstås och på vilka grunder betyg sätts. Här har lärarutbildningen ett ansvar. Lärarstudenterna behöver få grundläggande kunskaper i betyg och bedömning. Här har också skolans huvudmän och rektorer ett ansvar. Lärarna behöver inom ramen för det kollegiala lärandet ges utrymme att diskutera betyg och bedömning. Även Skolverket har ett ansvar. Lärarna ska från

⁷⁹ Skolverket (2015c) s.12. Se samma rapport s.256f för enkätsvar från lärarna år 2004 och år 2014 om möjligheten till en likvärdig betygssättning inom skolan, inom kommunen och i hela landet. Se samma rapport bilaga 4 för en utförlig beskrivning av intervjusvar om svårigheterna att uppnå gemensamma tolkningsgrunder av kursplaner och kunskapskrav. Se även Skolverket (2016a) för liknande slutsatser som delvis bygger på samma underlag. I enkätundersökningen från 2004 konstateras att svensklärarna, som har nationella prov, inte tyckte att likvärdighetsproblemen i betygssättningen var mindre än lärare i ämnen som inte har nationella prov, Skolverket (2005) s.16f.

⁸⁰ Skolverket håller för närvarande på med en översyn av kursplaner och kunskapskrav.

<https://www.skolverket.se/undervisning/sa-arbetar-vi-med-att-revidera-kurs--och-amnesplanerna>

⁸¹ Tillämpades för första gången i årskurs 9 vårterminen 2013. Innan dess fanns betygskriterier. För en beskrivning av kunskapskraven och värdeorden se Skolverket (2018b) s.34f och <https://www.skolverket.se/undervisning/grundskolan/betyg-i-grundskolan/satta-betyg-i-grundskolan#h-Kunskapskraven>

Skolverket kunna ta del av ett användarvänligt och behovsanpassat bedömningsstöd som kan stödja den enskilda läraren vid betygssättning och vara ett underlag i det kollegiala lärandet. En ökad och i större utsträckning gemensam bedömar-kompetens i lärarkollektivet kan förhoppningsvis reducera den bristande likvärdig-heten vid betygssättning, men kan knappast rå på de grundläggande systematiska skillnader i betygssättning mellan skolor som rapporten visar. För att nå en markant ökad likvärdighet i betygssättningen krävs systemförändringar i betygs-systemet, exempelvis att de nationella proven på något vis styr betygssättningen på skolnivå. Sådana systemförändringar behöver utredas noggrant och vägas mot tänkbara negativa effekter som förändringarna kan få.

Bilaga B1

B1.1 Datakällor och exkluderingar

Betygsdata och uppgifter om provbetyg för grundskolan är hämtad från Skolverkets databaser som innehåller uppgifter om elevers betyg, provbetyg, kön, ålder, migrationsbakgrund och föräldrarnas utbildningsnivå. Dessa databaser har sedan länkats med SCB:s RAKS-databas där uppgifter om föräldrars inkomst och bidragstagande har hämtats. För länkningen används elevens löpnummer, vilket kan ses som en anonymiserad version av personnumret.⁸²

Datarensningar och bortfall

Elever som är nyanlända och ännu inte beviljats permanent uppehållstillstånd har inget personnummer och därmed inte heller något löpnummer. Detta innebär att även om vi har tillgång till dessa elevers betyg så kan vi inte koppla på andra bakgrundsdata varför dessa elever inte ingår i de flesta av rapportens analyser. Elever med skyddad identitet har inte heller något löpnummer och dessa ingår därför inte heller i analyserna. Dessutom har elever som inte erhåller konventionella betyg också exkluderats från analyserna.

Huvudresultaten i denna rapport bygger på data från åren 2013–2015 och en deskriptiv bild av variablerna som används i rapporten redovisas endast för dessa år i bilagan. Vi har även gjort ett fåtal analyser utifrån en längre tidserie men av utrymmesskäl redovisas inte dessa data i bilagan.

Observera att för matematik, engelska och svenska skriver samtliga elever som läser ämnena det nationella provet och dessa resultat samlas in. För no- och so-ämnena skriver varje skolas elever endast ett av no-proven och ett av so-proven. Detta innebär att uppgifter om provbetyg endast finns för ca en tredjedel av årskullen i vart och ett av no-ämnena och för ca en fjärdedel av årskullen i vart och ett av so-ämnena.

Elever i skolor med mindre än 15 elever i respektive ämne har exkluderats. Anledningen är att vi analyserar variabler på skolnivå och vill att resultaten ska vara så stabila som möjligt för varje enskild skola.⁸³

⁸² Nyckeln mellan löpnummer och personnummer har endast SCB tillgång till.

⁸³ Ungefär 1500 elever per årskurs har exkluderats på grund av skolstorlek, vilket motsvarar ungefär 1,8 procent.

Tabell B1.1 Antal elever som ingår i analyserna

	2013	2014	2015
Elever med betyg i betygsdatabas	96 607	97 285	97 743
varav med löpnummer	96 235	96 813	96 703
varav med provbetyg i ämne:			
matematik	87 537	87 087	87 573
engelska	88 491	88 351	88 334
svenska*	81 961	80 658	80 320
fysik	29 305	29 699	26 527
biologi	29 091	25 694	30 272
kemi	26 439	28 872	28 952
geografi	22 970	22 027	21 105
historia	22 290	21 192	22 238
religion	21 890	20 799	23 124
samhällskunskap	20 310	22 817	21 178
varav med provbetyg och bakgrundsvariabler**:			
matematik	83 989	84 109	84 210
engelska	81 214	81 883	81 463
svenska*	74 889	74 750	73 946
fysik	27 050	27 752	24 853
biologi	26 924	24 110	28 121
kemi	24 439	27 235	26 936
geografi	21 263	20 507	19 527
historia	20 524	19 962	20 591
religion	19 692	19 396	21 407
samhällskunskap	18 727	21 188	19 638
Bortfall på grund av provbetyg***			
matematik	6,7%	6,4%	7,1%
engelska	5,4%	4,9%	6,1%
svenska*	3,7%	3,7%	4,0%
Bortfall på grund av bakgrundsvariabler****			
matematik	4,1%	3,4%	3,8%
engelska	8,2%	7,3%	7,8%
svenska*	8,6%	7,3%	7,9%
fysik	7,7%	6,6%	6,3%
biologi	7,4%	6,2%	7,1%
kemi	7,6%	5,7%	7,0%
geografi	7,4%	6,9%	7,5%
historia	7,9%	5,8%	7,4%
religion	10,0%	6,7%	7,4%
samhällskunskap	7,8%	7,1%	7,3%

*Observera att elever som läser svenska som andraspråk är exkluderade.

**Här ingår också endast elever i skolor med minst 15 elever i det aktuella ämnet

***Kan ej beräknas för no- och so-ämnena på grund av att ämnesproven endast görs av ett urval.

****Bortfallet är beräknat utifrån antal elever som har betyg i ämnet.

B1.2 Variabler som ingår i analyserna

Utfallsvariabel

Avvikelse mellan elevens betyg och provbetyg utifrån betygspoäng

$$[\text{äa_diff_poäng}] = \text{Avvikelse i ämne} = \text{betyg i ämne} - \text{provbetyg i ämne}$$

Där såväl betyg som provbetyg kodats som betygspoäng, dvs F=0, E=10, D=12,5, C=15, B=17,5 och A=20 poäng.⁸⁴

Tabell B1.2a Deskriptiv bild av de elever som har löpnummer samt såväl slutbetyg som provbetyg.

	2013			2014			2015		
	Antal	Medelv.	SD	Antal	Medelv.	SD	Antal	Medelv.	SD
Betyg									
matematik	87 537	13,10	4,40	87 087	12,75	4,36	87 573	12,75	4,53
engelska	88 491	14,63	4,12	88 351	14,80	4,12	88 334	14,99	4,25
svenska*	81 961	13,85	3,64	80 658	14,10	3,66	80 320	14,19	3,71
fysik	29 305	13,11	3,92	29 699	13,13	3,97	26 527	13,31	4,24
biologi	29 091	12,91	4,04	25 694	13,56	3,88	30 272	13,58	4,11
kemi	26 439	12,76	4,09	28 872	12,96	4,24	28 952	13,75	4,05
geografi	22 970	13,77	3,89	22 027	13,96	3,82	21 105	13,79	4,04
historia	22 290	13,57	4,09	21 192	13,87	4,16	22 238	13,57	4,13
religion	21 890	13,44	3,81	20 799	13,83	3,89	23 124	14,40	3,69
samhällskunskap	20 310	14,03	3,51	22 817	13,72	3,85	21 178	14,01	3,96
Provbetyg									
matematik	87 537	12,24	5,30	87 087	11,40	5,25	87 573	10,74	5,92
engelska	88 491	14,89	4,11	88 351	15,02	4,05	88 334	15,07	4,26
svenska*	81 961	13,36	3,86	80 658	13,63	3,88	80 320	13,50	3,98
fysik	29 305	11,95	4,47	29 699	11,44	5,11	26 527	11,80	5,38
biologi	29 091	10,93	5,17	25 694	12,53	4,61	30 272	11,96	5,41
kemi	26 439	11,23	4,96	28 872	11,44	5,35	28 952	13,33	4,77
geografi	22 970	13,01	4,63	22 027	13,10	4,63	21 105	12,63	4,82
historia	22 290	11,21	5,68	21 192	12,44	5,58	22 238	10,74	5,38
religion	21 890	11,29	5,03	20 799	12,87	4,60	23 124	13,83	4,27
samhällskunskap	20 310	13,79	3,28	22 817	12,68	4,24	21 178	12,93	4,79
Avvikelse mellan betyg och provbetyg									
matematik	87 537	0,86	2,67	87 087	1,35	2,88	87 573	2,00	3,48
engelska	88 491	-0,26	1,70	88 351	-0,22	1,75	88 334	-0,07	1,85
svenska*	81 961	0,49	2,27	80 658	0,46	2,27	80 320	0,69	2,37
fysik	29 305	1,16	2,79	29 699	1,69	3,23	26 527	1,51	3,24
biologi	29 091	1,99	3,37	25 694	1,03	2,75	30 272	1,63	3,30
kemi	26 439	1,53	3,11	28 872	1,52	3,22	28 952	0,42	2,65
geografi	22 970	0,76	2,78	22 027	0,86	2,60	21 105	1,17	2,69
historia	22 290	2,36	3,84	21 192	1,43	3,35	22 238	2,83	3,61
religion	21 890	2,14	3,53	20 799	0,96	2,48	23 124	0,57	2,41
samhällskunskap	20 310	0,24	2,18	22 817	1,04	2,41	21 178	1,08	2,74

*Observera att elever som läser svenska som andraspråk ej ingår i analyserna av svenska.

⁸⁴ Vi har även genomfört analyserna utifrån symmetrisk betygsskala, *betygssteg*, dvs F=1, E=2, D=3, C=4, B=5 samt A=6. Resultat från dessa körningar redovisas i excelfil.

Tabell B1.2b Deskriptiv bild av de elever som analyserna bygger på.**

	2013			2014			2015		
	Antal	Medelv.	SD	Antal	Medelv.	SD	Antal	Medelv.	SD
Betyg									
matematik	83 989	13,14	4,37	84 109	12,79	4,33	84 201	12,79	4,50
engelska	81 214	14,76	3,98	81 883	14,95	3,97	81 463	15,16	4,06
svenska*	74 889	13,98	3,48	74 750	14,23	3,51	73 946	14,33	3,54
fysik	27 050	13,21	3,82	27 741	13,22	3,88	24 852	13,43	4,13
biologi	26 914	13,04	3,90	24 110	13,67	3,78	28 117	13,72	4,00
kemi	24 438	12,87	3,99	27 234	13,07	4,15	26 936	13,88	3,93
geografi	21 263	13,93	3,72	20 507	14,12	3,66	19 527	13,96	3,87
historia	20 524	13,75	3,90	19 962	14,00	4,03	20 588	13,74	3,97
religion	19 692	13,56	3,66	19 396	13,94	3,78	21 407	14,56	3,55
samhällskunskap	18 727	14,18	3,38	21 188	13,87	3,71	19 638	14,17	3,81
Provbetyg									
matematik	83 989	12,28	5,27	84 109	11,45	5,22	84 201	10,79	5,91
engelska	81 214	15,02	3,98	81 883	15,16	3,91	81 463	15,23	4,10
svenska*	74 889	13,47	3,74	74 750	13,75	3,76	73 946	13,63	3,85
fysik	27 050	12,04	4,38	27 741	11,55	5,04	24 852	11,94	5,27
biologi	26 914	11,05	5,08	24 110	12,64	4,52	28 117	12,12	5,31
kemi	24 438	11,35	4,90	27 234	11,56	5,29	26 936	13,49	4,64
geografi	21 263	13,19	4,47	20 507	13,31	4,45	19 527	12,83	4,64
historia	20 524	11,41	5,56	19 962	12,59	5,48	20 588	10,94	5,25
religion	19 692	11,41	4,95	19 396	12,99	4,50	21 407	14,00	4,12
samhällskunskap	18 727	13,91	3,17	21 188	12,84	4,09	19 638	13,12	4,63
Avvikelse mellan betyg och provbetyg									
matematik	83 989	0,86	2,65	84 109	1,34	2,86	84 201	2,00	3,48
engelska	81 214	-0,25	1,65	81 883	-0,21	1,69	81 463	-0,06	1,79
svenska*	74 889	0,52	2,19	74 750	0,48	2,20	73 946	0,71	2,29
fysik	27 050	1,17	2,77	27 741	1,68	3,20	24 852	1,49	3,19
biologi	26 914	1,99	3,33	24 110	1,03	2,70	28 117	1,60	3,26
kemi	24 438	1,52	3,07	27 234	1,50	3,19	26 936	0,39	2,59
geografi	21 263	0,74	2,70	20 507	0,82	2,50	19 527	1,13	2,60
historia	20 524	2,34	3,79	19 962	1,41	3,31	20 588	2,80	3,55
religion	19 692	2,16	3,50	19 396	0,95	2,43	21 407	0,56	2,35
samhällskunskap	18 727	0,26	2,09	21 188	1,03	2,34	19 638	1,05	2,67

*Observera att elever som läser svenska som andraspråk ej ingår i analyserna av svenska.

**Endast elever som har löpnummer och uppgift om slutbetyg, provbetyg och samtliga bakgrundsvariabler. Dessutom ingår endast elever som går i skolor med minst 15 elever i respektive ämne.

Bakgrundsvariabler

Variabel	Förklaring	Ingår i analysmodell för faktiska data	Ingår i modell för att generera fiktiva data
[sex]	Kön, 0=pojke, 1=flicka	ja	ja
[migfor]	Utlandsfödd, 0=född i Sverige, 1=född utomlands*	ja	nej
[sociox]	Ämnesspecifikt socioekonomiskt index (baserat på [edu13], [pink], [bidrag])*	ja	ja
[edu13]	Föräldrarnas genomsnittliga utbildningsnivå, utifrån 7 nivåer (13 steg)*	ja (ingår i index)	ja (ingår i index)
[pink]	Föräldrars inkomst, percentilrankad*	ja (ingår i index)	ja (ingår i index)
[bidrag]	Föräldrars grad av bidragstagande*	ja (ingår i index)	ja (ingår i index)
[hman]	Typ av huvudman, 0=kommunal, 1=fristående	ja	ja
[pb_skola]	Skolans genomsnittliga provbetyg*	ja	ja
[elev_tidig], [elev_sen]	Ålder i förhållande till typålder för årskurs, 1=elev yngre respektive äldre än typåldern, 0=typåldern*	nej	ja
[ar_swe]	Antal år i Sverige*	nej	ja
[HDI]	Elevers ursprungsland utifrån geografisk närhet och utvecklingsindex (7 grupper)*	nej	ja

*Variabeln förklaras i detalj nedan

Migrationsbakgrund / utlandsfödd

Anges som 1 om eleven är född utomlands (oavsett var föräldrarna är födda), annars 0. Till skillnad från SCB antas här alla elever vara utlandsfödda om uppgift saknas om födelseland.⁸⁵

Föräldrarnas utbildningsnivå

Utgångspunkten är den 7-gradiga variabeln [SUN2000_old] där:

- 1 = Förgymnasial utbildning kortare än 9 år
- 2 = Förgymnasial utbildning 9 år
- 3 = Gymnasial utbildning högst 2 år
- 4 = Gymnasial utbildning 3 år
- 5 = Eftergymnasial utbildning kortare än 3 år
- 6 = Eftergymnasial utbildning 3 år eller längre
- 7 = Forskarutbildning

⁸⁵ SCB placerar istället alla elever som inte kan definieras på grund av bortfall i födelseland i kategorin födda i Sverige då det är den största kategorin. Vi anser att det är suboptimalt då gruppen utlandsfödda då underskattas med motiveringen att det är högst sannolikt att en individ som saknar uppgift om födelseland är född i utlandet.

Variabeln [edu13] har sedan skapats genom att ta genomsnittet av båda föräldrarnas 7-gradiga utbildningsnivå, vilket genererar en variabel med 13 steg (1: 1,5: 2: 2,5: 3: 3,5: 4: 4,5: 5: 5,5: 6: 6,5: 7). Om uppgift saknas för en förälder baseras variabelvärdet på den förälder som det finns uppgift för. Den 13-gradiga variabeln [edu13] har sedan kodats om till 12 dikotoma (0, 1) variabler för att få en bättre modellpassning för framtagandet av det socioekonomiska indexet.

Föräldrarnas inkomst

Måttet bygger på ett genomsnitt av föräldrarnas årsvisa percentilrankade samlade inkomst för de 7 sista åren fram till och med det år eleven går ut åk 9. Genomsnittet bygger på de år som det finns observationer för. För vissa elever kan därmed föräldrarnas inkomst bygga på färre än 7 års uppgifter.⁸⁶ Om uppgift finns för en förälder men inte den andra ett år har inkomsten för den förälder som det saknas uppgift från antagits vara 0.

Att percentilranka inkomsten innebär lite förenklat att rangordna inkomsterna och dividera med antalet observationer (och sedan multiplicera med 100). Därmed fås en så kallad percentilranking som ligger mellan 0 och 100 och som sedan används i analyserna. Ett värde på t.ex. 97 innebär då att inkomsten för elevens föräldrar är högre än 97 procent av övriga elevers föräldrars inkomst.

Observera att inkomstvariabeln [pink] bygger på genomsnittet över de 7 senaste åren av föräldrarnas sammanlagda inkomst där inkomsten för varje enskilt år är percentilrankad. Detta är inte samma sak som att percentilranka genomsnittet över 7 år. Därför är de årsvisa medelvärdena för [Pink] inte identiska för varje enskilt år, vilket de annars skulle varit.

Föräldrarnas grad av bidragstagande

Variabeln mäter i vilken utsträckning någon förälder har eller har haft bidrag som huvudsaklig inkomstkälla under de senaste 7 åren.⁸⁷ För vart och ett av de 7 åren fram till eleven gått ut årskurs 9 anges om minst en av föräldrarna varit bidragstagare (haft bidrag som huvudsaklig inkomstkälla). Variabeln är ett genomsnitt av dessa 7 år vilket innebär att den inte är en dikotom variabel utan varierar mellan 0 och 1. Ett värde på 1 innebär att minst en av föräldrarna haft bidrag som huvudsaklig inkomstkälla under samtliga 7 år. Ett värde på t.ex. 0,42 (3/7) innebär att någon av föräldrarna varit bidragstagare under 3 av dessa 7 år.

Socioekonomiskt index

För flera av analyserna används ett index som bygger på ovanstående tre redovisade variabler: *föräldrarnas utbildningsnivå*, *föräldrarnas inkomst* och *föräldrarnas*

⁸⁶ Inkomstuppgifter finns bara tillgängliga fram tom. 2013. För 2014 är den senaste inkomstuppgiften därmed 1 år gammal och för 2015 är den 2 år gammal. Detta innebär sannolikt att sambandet mellan betyg/provbetyg och inkomst underskattas något för dessa två år (jämfört med övriga år).

⁸⁷ Uppgifter om bidragstagande finns bara fram tom. 2013. Se föregående fotnot.

grad av bidragstagande. Indexet är konstruerat så att varje komponent (variabel) har en del i indexet som motsvarar dess relativa betydelse för att förklara resultaten. Exempelvis utgör föräldrarnas genomsnittliga utbildningsnivå störst andel av indexet då denna variabel är den som förklarar mest av elevers resultat, av de tre variablerna. Grad av bidragsberoende är istället den variabel som kan förklara minst av elevers resultat och därmed utgör den också den minsta andelen av det socioekonomiska indexet. Inkomst intar en mellanposition. För att maximera prediktionskraften i indexet har separata index skapats för varje enskilt ämne och år.⁸⁸

Skolans genomsnittliga provbetyg

Variabeln är beräknad som det aritmetiska medelvärdet av elevernas provbetyg. Här ingår även elevers provbetyg som av olika anledningar inte ingår själva i analysen, på grund av bortfall i någon bakgrundsvariabel. Men eftersom syftet med denna variabel är att fånga skolans genomsnittliga prestationsnivå så inkluderas dessa elever för att få en så väntevärdesriktig skattning som möjligt.

Ålder i förhållande till typåldern för årskursen

[elev_tidig] respektive [elev_sen] är dikotoma variabler som anger om eleven är yngre respektive äldre än typåldern för årskursen (baserat på kalenderår). Att en elev är yngre än typåldern kan indikera att eleven startat skolan ett år tidigare eller hoppat över ett år. Att en elev är äldre än typåldern kan indikera att eleven gått om ett år eller börjat skolan senare. Men det kan också indikera (och det verkar vara den vanligaste orsaken) att eleven är utlandsfödd och invandrat efter ordinarie skolstart och placerats i en årskurs lägre än normalt för den elevens ålder.

Antal år i Sverige

För elever som är födda utomlands beräknas variabeln [ar_swe] utifrån elevens invandningsår. För elever födda i Sverige sätts värdet till 16. [ar_swe2] är variabeln [ar_swe] kvadrerad.

Elevens ursprungsland utifrån geografisk närhet och landets utvecklingsindex

Variabeln består av 7 kategorier som kodats som 6 dummyvariabler. HDI-grupp 1=elever födda i Sverige, 2=födda i Norden (utanför Sverige), 3=födda i EU (utanför Norden), 4=födda i övriga Europa, 5=födda utanför Europa från land med högt utvecklingsindex, 6=födda utanför Europa i land med medelindex, 7=födda utanför Europa i land med lågt utvecklingsindex.

⁸⁸ Rent formellt skapas indexet utifrån en multipel regression med de tre ovan nämnda socioekonomiska variabler som oberoende variabler (x-variabler) och provbetyget i det aktuella ämnet som utfallsvariabel (y). Utifrån denna regression tas de predicerade (förväntade) värdena fram utifrån det skattade regressionssambandet. Därmed fås ett sammanvägt värde på vilket resultat en elev kan förväntas prestera, givet sin socioekonomiska bakgrund enligt de tre variablerna. Detta förväntade resultat utgör indexvärdet för den eleven.

Tabell B1.3 Deskriptiv bild av kön, elev tidig, elev sen, utlandsfödd, antal år i Sverige, utbildningsnivå, inkomst, grad av bidragstagande samt typ av huvudman.

	2013			2014			2015		
	Antal	Medelv.	SD	Antal	Medelv.	SD	Antal	Medelv.	SD
kön	91 029	0,488	0,500	90 998	0,485	0,500	91 449	0,489	0,500
elev tidig	91 029	0,015	0,122	90 998	0,016	0,125	91 449	0,016	0,125
elev sen	91 029	0,052	0,222	90 998	0,051	0,221	91 449	0,051	0,220
utlandsfödd	91 029	0,092	0,289	90 998	0,098	0,298	91 449	0,104	0,305
antal år i Sverige	91 029	15,21	2,77	90 998	15,14	2,91	91 449	15,10	2,97
utbildningsnivå	91 029	3,99	1,26	90 998	4,01	1,28	91 449	4,04	1,29
inkomst	91 029	49,98	27,61	90 998	49,90	28,00	91 449	49,62	28,14
grad av bidragstagande	91 029	0,147	0,297	90 998	0,147	0,306	91 449	0,147	0,306
elev i fristående skola	91 029	0,150	0,357	90 998	0,159	0,366	91 449	0,169	0,375

Tabell B1.4 Deskriptiv bild av de socioekonomiska indexen.

	2013			2014			2015		
	Antal	Medelv.	SD	Antal	Medelv.	SD	Antal	Medelv.	SD
Ämnesspecifikt socioekonomiskt index									
matematik	83 989	3,22	0,65	84 109	2,94	0,62	84 210	2,87	0,65
engelska	81 214	4,09	0,52	81 883	4,14	0,52	81 463	4,18	0,55
svenska*	74 889	3,48	0,42	74 750	3,59	0,45	73 946	3,55	0,46
fysik	27 050	3,04	0,49	27 752	2,96	0,53	24 853	3,13	0,60
biologi	26 924	2,81	0,49	24 110	3,26	0,54	28 121	3,18	0,58
kemi	24 439	2,87	0,50	27 235	3,00	0,57	26 936	3,57	0,57
geografi	21 263	3,46	0,56	20 507	3,50	0,59	19 527	3,34	0,60
historia	20 524	3,01	0,54	19 962	3,38	0,57	20 591	2,83	0,58
religionskunskap	19 692	2,90	0,45	19 396	3,37	0,51	21 407	3,73	0,46
samhällskunskap	18 727	3,61	0,46	21 188	3,28	0,53	19 638	3,46	0,55

Tabell B1.5 Deskriptiv bild av skolans genomsnittliga provbetyg.

	2013			2014			2015		
	Antal	Medelv.	SD	Antal	Medelv.	SD	Antal	Medelv.	SD
Skolans genomsnittliga provbetyg									
matematik	83 989	12,25	1,84	84 109	11,41	1,87	84 210	10,76	2,18
engelska	81 214	14,90	1,43	81 883	15,04	1,44	81 463	15,09	1,59
svenska*	74 889	13,14	1,28	74 750	13,39	1,35	73 946	13,22	1,41
fysik	27 050	11,92	1,87	27 752	11,43	2,17	24 853	11,80	2,23
biologi	26 924	10,91	2,12	24 110	12,53	1,86	28 121	11,96	2,35
kemi	24 439	11,24	2,03	27 235	11,45	2,33	26 936	13,35	1,84
geografi	21 263	13,03	1,77	20 507	13,15	1,67	19 527	12,63	1,94
historia	20 524	11,22	2,11	19 962	12,45	2,09	20 591	10,76	2,15
religion	19 692	11,26	1,87	19 396	12,86	1,68	21 407	13,85	1,57
samhällskunskap	18 727	13,80	1,25	21 188	12,70	1,74	19 638	12,95	1,75

*Elever som läser svenska som andraspråk ingår ej i analyserna av svenskämnet.

**Elever i skolor med färre än 15 elever i respektive ämne är exkluderade.

Tabell B1.6 Deskriptiv bild av typ av ursprungsland utifrån HDI-grupp

	2013	2014	2015
HDI-grupp 1	90,8%	90,2%	89,6%
HDI-grupp 2	0,6%	0,6%	0,6%
HDI-grupp 3	1,6%	1,7%	1,7%
HDI-grupp 4	1,0%	1,0%	1,0%
HDI-grupp 5	0,5%	0,5%	0,5%
HDI-grupp 6	4,2%	4,3%	4,4%
HDI-grupp 7	1,3%	1,7%	2,1%
Totalt	100,0%	100,0%	100,0%

HDI-grupp 1=elever födda i Sverige, 2=födda i Norden (utanför Sverige), 3=födda i EU (utanför Norden), 4=födda i övriga Europa, 5=födda utanför Europa från land med högt utvecklingsindex, 6=födda utanför Europa i land med medelindex, 7=födda utanför Europa i land med lågt utvecklingsindex.

B1.3 Metoder

Skattning av betygssättningseffekter för elevgrupper och skolkategorier

Här ges en översiktlig redovisning av hur betygssättningseffekterna som presenteras i figur 1–3 respektive 6–7 i huvudtexten skattas. En mer detaljerad genomgång finns i den tekniska rapporten (kommande).

1. Separata linjära regressioner enligt modell (1) nedan skattas för varje enskilt provbetyg, för varje enskilt ämne samt för varje enskilt år 2013, 2014 respektive 2015.

$$\begin{aligned} [Diff_{\ddot{a}\ddot{a}_{po\ddot{a}ng}}] = & \beta_0 + \beta_1[sex] + \beta_2[migfor] + \beta_3[sociox_{\ddot{a}\ddot{a}}] + \\ & \beta_4[hman] + \beta_5[pb_{\ddot{a}\ddot{a}_{skola}}] \end{aligned} \quad (1)$$

där:

[diff_ä_ä_poäng] = [betyg_ä_ä_poäng] – [pb_ä_ä_poäng], dvs. avvikelserna mellan elevens betyg i ämne ää och elevens provbetyg i motsvarande ämne, uttryckt i betygspoäng.⁸⁹

[sex] = elevens kön.

[migfor] = utlandsfödd.

[sociox_ä_ä] = socioekonomisk bakgrund utifrån ämnesspecifikt index.

[hman] = huvudman för elevens skola.

[pb_ä_ä_skola] = skolans genomsnittliga provbetyg i ämne ää.

Skattningarna av b1, b2, b3, b4 och b5 utgör *bruttoeffekter* på betygssättning. Dessa bruttoeffekter behöver dock rensas för *fördelningseffekter*.

2. Utifrån samma faktiska dataset genereras ett fiktivt/simulerat dataset som har samma antal elever, skolor och strukturer vad gäller samband mellan provbetygsresultat och bakgrundsvariabler som faktiska data, men där betygssättningen sker helt mekaniskt objektivt.⁹⁰
3. Regressionsmodell (1) skattas nu även på det simulerade/fiktiva datasetet från steg 2, enligt samma princip, dvs med separata regressioner för varje provbetyg. Effektskattningarna f1, f2, f3, f4 och f5 från denna simulerade data utgör mått på *fördelningseffekterna*.
4. Den slutliga skattningen av betygssättningseffekten för varje provbetyg fås genom att subtrahera fördelningseffekterna från bruttoeffekterna enligt: Slutlig betygssättningseffekt S1=b1–f1, S2=b2–f2, ..., S5=b5–f5.

⁸⁹ Vi har även skattat motsvarande modeller utifrån betygssteg, dvs (1–6)-skala istället för (0–20)-skala. Se Excel-filen för resultat från dessa analyser.

⁹⁰ Se den tekniska rapporten för hur det fiktiva datasetet genererats.

5. De slutliga betygssättningseffekterna för varje provbetyg vägs ihop till en samlad betygssättningseffekt per ämne utifrån andelen elever med respektive provbetyg.

Beräkning av skolors nettoavvikelser samt skolskillnader (figur 4–5)

Figur 4a–c baseras på skolornas nettoavvikelser utifrån betygssteg (F=1, E=2, D=3, C=4, B=5, A=6). För varje skola har ett genomsnitt beräknats utifrån elevernas avvikelser. Eftersom avvikelserna är baserade på betygssteg kan man tolka en skolas nettoavvikelse som andelen elever som får ett betygssteg högre betyg än provbetyget medan den resterande andelen får samma betyg som provbetyg. Skolorna har sedan rangordnats efter sin nettoavvikelse.

För figur 5 har standardavvikelsen beräknats utifrån skolornas nettoavvikelser. Dessutom har resultaten justerats för provbetygs- och fördelningseffekter, enligt samma princip som i föregående avsnitt. Standardavvikelsen i skolornas nettoavvikelse kan tolkas som hur mycket en skolors nettoavvikelse i genomsnitt skiljer sig från riksgenomsnittet i skolors nettoavvikelse. En högre standardavvikelse innebär att det är större skillnader mellan skolor i hur de sätter betyg i förhållande till de nationella proven. Observera dock att om alla skolor avvek lika mycket från det nationella provresultatet (åt samma håll), skulle standardavvikelsen bli noll. Det skulle motsvaras av att alla staplar i figur 4a–c är precis lika höga.

Se vidare i den tekniska rapporten för en utförligare beskrivning av metoderna.

Beräkning av andel av skolskillnader som kan förklaras av relativ betygssättning/skolans prestationsnivå (figur 8)

För att kunna beräkna hur stor andel av skolskillnaderna i betygssättning som kan förklaras av relativ betygssättning används variansdekomposition. Variansdekomposition innebär i detta sammanhang att man först beräknar andelen av den totala variationen i elevers avvikelser mellan betyg och provbetyg som kan förklaras av skillnader mellan skolor, utan några förklarande variabler, en så kallad nollmodell. Sedan tillförs variablerna på elevnivå (kontrollvariabler), i detta fall [sex], [migför] och [sociox]. Nu beräknas den återstående variation som finns mellan skolorna. Sedan läggs variabeln [pb_skola] till modellen och vi beräknar hur mycket av den variationen som försvinner på grund av att vi lägger till denna variabel. Detta utgör den andel av mellanskolvariationen i betygssättning som kan förklaras av relativ betygssättning.⁹¹

Med denna metod kan vi inte justera för provbetygseffekter eller fördelningseffekter. Därmed riskerar vi att under- eller överskatta den ”sanna” andelen förklarad variation. För att få skattningar som är approximativt väntevärdesriktiga används två olika metoder där den ena ger ett undre golv för den sanna andelen

⁹¹ I ett ytterligare steg tillförs variabeln [hman] och vi kan beräkna hur mycket den kan förklara av den totala mellanskolvariationen. Ordningen som man lägger till variablerna kan påverka förklaringsgraden och vi har provat att byta plats i ordningen mellan [pb_skola] och [hman]. Förklaringsgraden av [pb_skola] påverkas endast marginellt.

och den andra ett övre tak för den sanna andelen. Detta görs genom att först utföra variansdekompositionen utifrån betygssteg. Med betygssteg (1–6) kommer betygssättningseffekten av skolans prestationsnivå att underskattas när man inte justerar för provbetygs- och fördelningseffekter. Sedan utförs variansdekompositionen utifrån betygspoäng. Med betygspoäng (0–20) kommer betygssättningseffekten av skolans prestationsnivå att överskattas när man inte justerar för provbetygs- och fördelningseffekter. Därmed erhålls ett undre golv respektive ett övre tak för den sanna andelen av skolskillnaderna som kan förklaras av relativ betygssättning (skolans prestationsnivå).⁹²

Alternativa modeller – känslighetsanalyser

Det finns flera olika sätt att skatta betygssättningseffekter. Nedan följer en genomgång av några alternativa modeller och/eller skattningsmetoder som övervägts i framtagandet av denna rapport.

Gemensamt för samtliga modeller som testats är att de baseras på samma kontrollvariabler, nämligen [sex], [migfor], [sociox], [hman] samt [pb_skola].

Utfallsvariabeln, avvikelsen mellan slutbetyg och provbetyg kan däremot variera utifrån två aspekter/dimensioner:

1. Huruvida vi uttrycker avvikelsen utifrån en asymmetrisk poängskala, dvs betygspoäng (0, 10, 12,5, 15, 17,5, 20) eller utifrån en symmetrisk skala, dvs betygssteg (1, 2, 3, 4, 5, 6).
2. Om vi väljer att uttrycka avvikelser utifrån betygssteg kan vi dessutom välja att mäta avvikelsen som en dikotom/nominell utfallsvariabel (0/1) eller en multinominell utfallsvariabel (t.ex. -2, -1, 0, 1, 2). (Med betygspoäng används multinominellt utfall).

Fördelen med att använda betygssteg är att resultaten blir mer lättolkade och att vi får en symmetri i skattningarna som gör att vi lättare kan se hur t.ex. fördelningseffekterna påverkar resultaten på ett mer principiellt sätt. Fördelen med att använda betygspoäng är att de reflekterar verkligheten på ett bättre sätt genom att skillnaden mellan att få F eller E är 10 poäng medan skillnaden mellan att få E eller D endast är 2,5 poäng, en skillnad på faktor 4. Genom att använda betygspoäng istället för betygssteg ges således en större vikt på betygssättningseffekter för elever med F i provbetyg.

En dikotom eller nominell utfallsvariabel innebär att vi endast beaktar om en elev får ett högre slutbetyg än provbetyg (=1) eller om eleven får samma eller lägre slutbetyg (=0). Då spelar det ingen roll om vi utgår från betygssteg eller

⁹² Vi vet att vi får dessa under- respektive överskattningar utifrån våra analyser med en konventionell OLS-regression som sedan jämförts med våra mått när vi rensat för provbetygs- och fördelningseffekter. Effektstorlekarna i variansdekompositionen (som bygger på en 2-nivåmodell) kan likställas med effektstorlekarna som erhålls genom konventionell OLS (utan uppdelning på provbetyg). Därmed vet vi att även variansdekompositionen leder till motsvarande under- och överskattningar.

betygspoäng eftersom alla utfall kodas till 0 eller 1. Ett multinominellt utfall innebär att vi även tar hänsyn till att en elev kan få ett lägre betyg än provbetyg (-1/-2,5 etc.) samt att graden av höjning, t.ex. en höjning eller sänkning med flera steg (-2, +2 eller -2,5, +2,5 eller -10, +10 ETC). Fördelen med att använda ett dikotomt utfall är att vi då kan använda en logistisk (icke linjär) regressionsansats där vi kan beräkna sannolikheter att få ett högre betyg. Det kan också argumenteras för att en icke linjär modell passar data bättre än en linjär modell. Fördelen med en multinominell utfallsvariabel är att den behåller betydligt mer information som annars kastas bort. Men vi hänvisas då till att använda en linjär modell.

Vi har testat och jämfört både en logistisk som linjär modell med dikotomt utfall och det visar sig att skillnaderna är mycket små. Där skillnader syns är för extrembetygen F och A, men de är inga stora skillnader.

Eftersom linjär och logistisk regression i praktiken ger samma resultat har vi valt att använda en linjär modell därför att vi då också kan använda en multinominell utfallsvariabel som bevarar all information.

Att använda en multinominell utfallsvariabel istället för en dikotom innebär i praktiken att vi får något större betygssättningseffekter. De flesta resultat blir helt enkelt tydligare eftersom vi tar hänsyn till såväl höjningar som sänkningar samt hur många stegs höjning och sänkning.

Konventionell OLS-ansats

Samtliga ovanstående modellvarianter har alla det gemensamt att de justerats för provbetygs- och fördelningseffekter. Att ta hänsyn till provbetygseffekter innebär att vi delar upp analysen på varje enskilt provbetyg istället för att skatta ett samband för alla provbetyg (poolad). Att ta hänsyn till fördelningseffekter innebär att vi skattar dessa effekter med hjälp av datasimuleringar och sedan subtraherar dessa effekter från de initiala effekterna (bruttoeffekterna) vi erhåller när vi skattar effekten per provbetyg.

Med en konventionell OLS-ansats menas här att modellen skattas för samtliga provbetyg tillsammans (poolad), dvs tar ej hänsyn till provbetygseffekter. Inte heller tas någon hänsyn till fördelningseffekter. Modellen är linjär och vi använder en multinominell utfallsvariabel uttryckt i betygspoäng. I Excel-filen redovisas resultaten från den konventionella OLS-ansatsen. De visar att den övergripande resultatbilden är densamma men att den konventionella OLS-ansatsen ger något starkare resultat i de flesta fall. Skillnaden påverkas också om man använder betygsslag eller betygspoäng. Sammantaget så blir dock den övergripande mönstret detsamma.

B1.4 Osäkerhet i skattningarna av systematiska avvikelser i betygssättning

Osäkerheten i skattningarna av betygssättningseffekter består av 2 komponenter:

1. Osäkerhet i samband med skattning av de så kallade bruttoeffekterna, vilka härrör från en regression på faktiska data. Denna osäkerhet är beräknad med hjälp av mplus och tar på ett korrekt sätt hänsyn till datastrukturen (t.ex. klustereffekter).
2. Osäkerhet i samband med skattning av fördelningseffekter. Denna osäkerhet skattas genom att utnyttja variationen i datasimuleringarna (10 oberoende simuleringar). Denna osäkerhet är beräknad utifrån varje provbetyg och ett medelvärde måste därför beräknas. I denna medelvärdesskattning har vi inte lyckats fånga in eventuell samvariation och det finns därmed risk för att osäkerheten är underskattad. Denna komponent är dock betydligt mindre än den första komponenten och bedömningen är att eventuell underskattning knappast kan ha en avgörande inverkan på den totala osäkerheten.

Den totala osäkerheten fås genom att summera de två komponenterna enligt:

Approximativt standardfel = $\sqrt{(se1)^2 + (se2)^2}$, där (se1) respektive (se2) är standardfelet för komponent 1 respektive komponent 2.

I tabell B1.7a-e redovisas approximativa standardfel för de betygssättningseffekter som redovisats i respektive figur 1–3 samt 6–7.

Tabell B1.7a-e Systematiska skillnader i betygssättning i förhållande till de nationella proven med approximativa standardfel samt test för statistisk signifikans.

Tabell B1.7a (figur 1: flickor och pojkar)

Ämne	Betygsättningseffekt	Approximativt standardfel	Signifikans på 5%nivå*
ma	0,274	0,014	0,247
en	0,354	0,009	0,338
sv	0,045	0,012	0,022
fy	0,290	0,027	0,238
bi	0,072	0,025	0,023
ke	-0,072	0,027	0,019
ge	0,288	0,025	0,240
hi	-0,131	0,033	0,066
re	-0,278	0,030	0,220
sh	-0,081	0,023	0,035

Tabell B1.7b (figur 2: migrationsbakgrund)

Ämne	Betygsättningseffekt	Approximativt standardfel	Signifikans på 5%nivå*
ma	0,097	0,032	0,034
en	0,197	0,018	0,162
sv	0,109	0,030	0,051
fy	0,334	0,055	0,226
bi	0,256	0,055	0,148
ke	0,169	0,053	0,066
ge	0,134	0,054	0,027
hi	0,211	0,068	0,077
re	-0,175	0,046	0,084
sh	0,132	0,049	0,037

Tabell B1.7c (figur 3: socioekonomisk bakgrund)

Ämne	Betygsättningseffekt	Approximativt standardfel	Signifikans på 5%nivå*
ma	-0,050	0,009	0,033
en	0,002	0,005	-0,008
sv	0,022	0,006	0,011
fy	-0,025	0,015	-0,004
bi	-0,070	0,016	0,040
ke	0,015	0,016	-0,016
ge	-0,083	0,017	0,049
hi	-0,129	0,018	0,093
re	0,000	0,015	-0,029
sh	0,023	0,011	0,000

*Positiva värden i den högra kolumnen, vilket motsvaras av gröna fält, anger att betygsättningseffekten är statistiskt signifikant på 5 % nivå.

Tabell B1.7d (figur 6: huvudman)

Ämne	Betygsättningseffekt	Approximativt standardfel	Signifikans på 5%nivå*
ma	0,278	0,036	0,207
en	0,109	0,028	0,054
sv	0,118	0,028	0,063
fy	0,332	0,073	0,189
bi	0,314	0,074	0,168
ke	0,311	0,076	0,163
ge	0,218	0,069	0,081
hi	0,202	0,083	0,040
re	0,272	0,071	0,132
sh	0,330	0,080	0,173

Tabell B1.7e. (figur 7: relativ betygsättning)

Ämne	Betygsättningseffekt	Approximativt standardfel	Signifikans på 5%nivå*
ma	-0,506	0,015	0,476
en	-0,168	0,019	0,131
sv	-0,280	0,011	0,258
fy	-0,858	0,033	0,793
bi	-0,852	0,032	0,790
ke	-0,810	0,031	0,748
ge	-0,530	0,029	0,474
hi	-0,894	0,038	0,819
re	-0,690	0,034	0,622
sh	-0,507	0,029	0,450

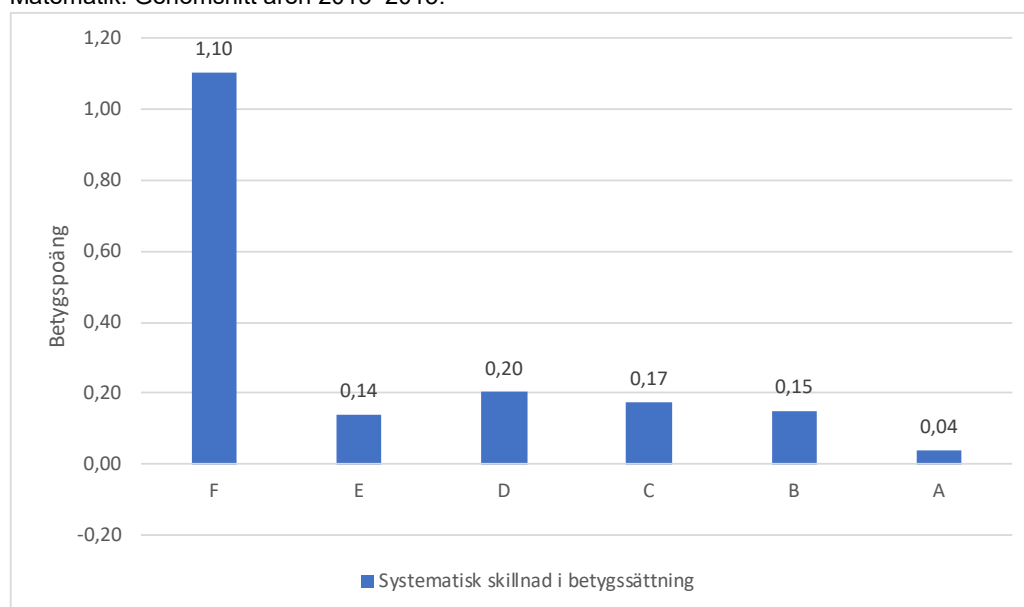
*Positiva värden i den högra kolumnen, vilket motsvaras av gröna fält, anger att betygsättningseffekten är statistiskt signifikant på 5 % nivå.

Bilaga 2 Kompletterande analyser

B2.1 Skillnader i betygssättning uppdelat på provbetyg

För varje ämne och år (2013–2015), har skillnaderna i betygssättning skattats för varje enskilt provbetyg, som sedan vägts ihop till ett samlat mått på skillnad i betygssättning mellan elevgrupperna/skolorna. Nedan redovisas som exempel skillnader i betygssättning mellan fristående och kommunala skolor, uppdelat på varje enskilt provbetyg för ämnet matematik. I excelfilen redovisas motsvarande skillnader i betygssättning för samtliga ämnen och elevgrupper/skolgrupper.

Figur B2.1. Skillnader i betygssättning i förhållande till de nationella proven mellan elever i fristående och kommunala skolor. Uppdelat på provbetyg. Uttryckt som betygspoäng. Matematik. Genomsnitt åren 2013–2015.



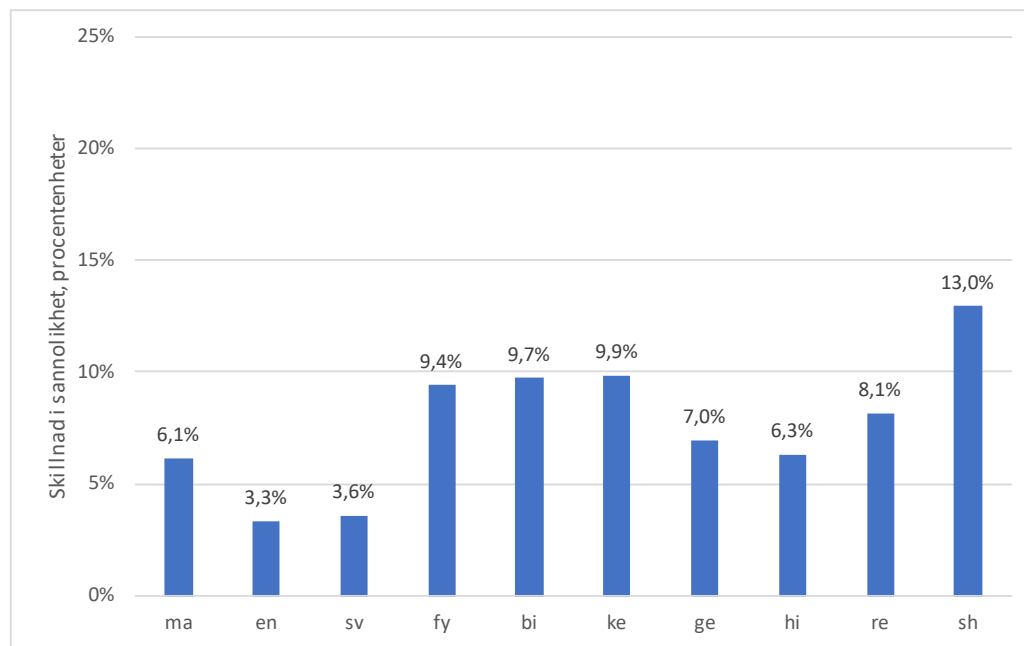
Positiva värden innebär att elever i fristående skolor får högre betyg i förhållande till de nationella proven, givet att alla övriga bakgrundsfaktorer är lika. Betygspoängen kan variera mellan 0-20. Justerad för provbetygs- och fördelningseffekter.

B2.2 Skillnader i betygssättning utifrån betygssteg

I rapporten uttrycks avvikelsen mellan betyg och provbetyg i betygspoäng (0, 10, 12,5, 15, 17,5, 20). Alla skattningar av betygssättningseffekter har också gjorts utifrån betygssteg (1–6). Nedan redovisas som exempel skillnader i betygssättning mellan kommunala och fristående skolor i matematik utifrån betygssteg. Se excel-filen för en redovisning av motsvarande resultat för samtliga ämnen och elevgrupper/skolkategorier.

Observera att betygssättningseffekterna (värdena) inte är direkt jämförbara med motsvarande effekter (värden) utifrån betygspoäng eftersom utfallsvariablerna uttrycks på olika skalor. I figuren nedan kan man istället tolka effekterna som skillnad i sannolikhet (i procentenheter) att erhålla ett steg högre slutbetyg (givet samma provbetyg).

Figur B2.2. Skillnader i betygssättning i förhållande till de nationella proven mellan elever i fristående och kommunala skolor. Uttryckt som sannolikheten att få ett steg högre betyg givet provbetyget. Efter ämne. Genomsnitt åren 2013–2015.

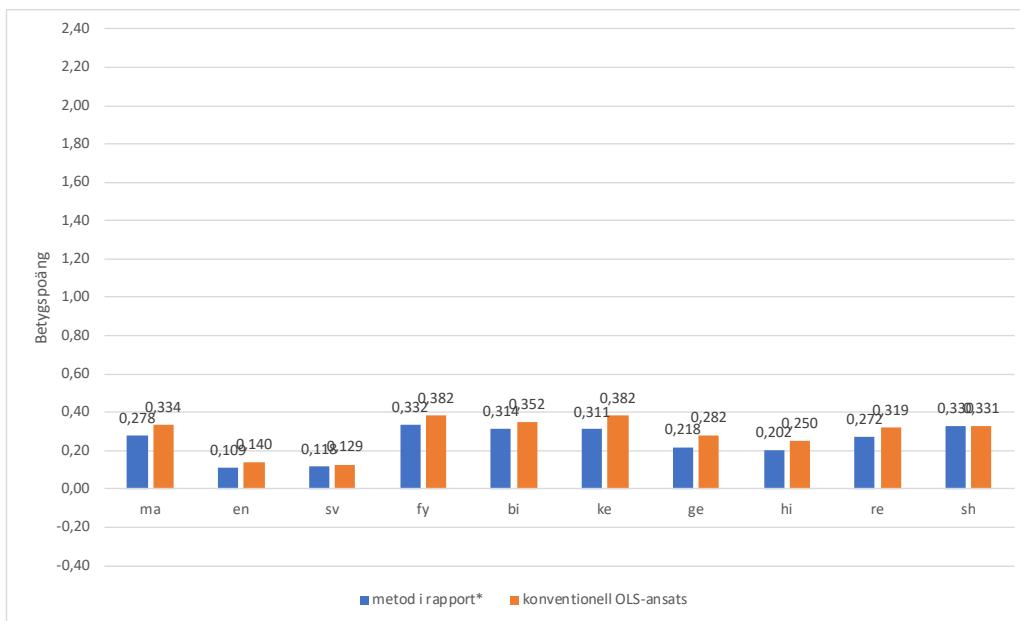


Positiva värden innebär att elever i fristående skolor får högre betyg i förhållande till de nationella proven. Betygsstegen kan variera mellan 1–6 i varje ämne. Justerad för provbetygs- och fördelningseffekter.

B2.3 Skillnader i skattning av betygssättning beroende på val av statistisk metod

I figur B2.3 redovisas som exempel skillnader i betygssättning mellan elever i fristående och kommunala skolor utifrån såväl den skattningsmetod som tillämpas i rapporten som en konventionell OLS-ansats, som ej tar hänsyn till provbetygs- och fördelningseffekter. I excelfilen redovisas motsvarande alternativa skattningar för övriga analysgrupper (kön, migrationsbakgrund, socioekonomisk bakgrund samt skolans genomsnittliga provbetyg).

Figur B2.3 Skillnader i betygssättning i förhållande till de nationella proven mellan elever i fristående och kommunala skolor. Enligt metod i rapport (= figur 6) samt enligt konventionell OLS-ansats.*



*Skillnaden mellan metod i rapport och konventionell OLS-ansats är att för metoden som används i rapporten, är betygssättningseffekterna justerade för provbetygs- och fördelnings-effekter. Positiva värden innebär att elever i fristående skolor systematiskt erhåller högre betyg i förhållande till provbetygen i jämförelse med elever i kommunala skolor.

B2.4 Spridning i genomsnittlig avvikelse mellan kommunala skolor respektive mellan fristående skolor

Tabell B2.1a. Spridning (standardavvikelse) i skolors genomsnittliga avvikelse, uttryckt som procentenheter. Mellan kommunala respektive mellan fristående skolor. Genomsnitt 2013–2015. Efter ämne.

	2013-2015 genomsnittlig spridning (standardavvikelse) i skolors genomsnittliga avvikelse	
	kommunala skolor	fristående skolor
ma	21,1%	24,5%
en	18,4%	20,9%
sv	19,3%	22,5%
fy	31,3%	35,6%
bi	31,5%	36,5%
ke	30,5%	34,5%
ge	25,3%	31,3%
hi	34,7%	37,0%
re	27,7%	30,4%
sh	26,6%	33,1%

Tabell B2.1b. Spridning (standardavvikelse) i skolors genomsnittliga avvikelse, uttryckt som procentenheter. Mellan kommunala respektive mellan olika skolkoncerner samt övriga fristående skolor. Genomsnitt 2013–2015. Efter ämne.

	2013-2015 genomsnittlig spridning i skolors genomsnittliga avvikelse (standardavvikelse)				
	kommunala skolor	Skolloncern 1	Skolkoncern 2	Skolkoncern 3	övriga fristående
ma	21,1%	24,8%	18,9%	26,5%	24,6%
en	18,4%	17,7%	16,3%	23,3%	21,3%
sv	19,3%	20,7%	18,6%	24,3%	23,2%
fy	31,3%	32,3%	43,9%	47,9%	32,0%
bi	31,5%		28,9%	35,1%	36,9%
ke	30,5%		17,6%		34,4%
ge	25,3%		23,6%	24,8%	33,4%
hi	34,7%		43,1%		37,6%
re	27,7%		25,7%		31,5%
sh	26,6%		28,7%	21,6%	30,6%

B2.5 Relativ betygssättning mellan kommunala respektive mellan fristående skolor

I tabell B2.2 redovisas effektstorleken i betygssättning för skolans genomsnittliga resultat på det nationella provet, på samma sätt som i figur 7, men här uppdelat på kommunala respektive fristående skolor. En skillnad jämfört med figur 7 är att effektstorlekarna inte justerats för provbetygs- eller fördelningseffekter.

Tabell B2.2a. Skillnader i betygssättning i förhållande till de nationella proven för elever i skolor som skiljer sig med en standardavvikelse i genomsnittligt provbetyg. Uppdelat på kommunala respektive fristående skolor. Genomsnitt 2013–2015.

ämne	kommunala skolor	fristående skolor
ma	-0,64	-0,66
en	-0,19	-0,07
sv	-0,26	-0,22
fy	-0,86	-0,73
bi	-0,88	-0,81
ke	-0,84	-0,78
ge	-0,53	-0,56
hi	-0,92	-0,90
re	-0,71	-0,59
sh	-0,42	-0,50

Rödfärgade värden innebär att den systematiska skillnaden i betygssättning inte är signifikant på 5 % nivå. Alla övriga värden innebär att den systematiska skillnaden i betygssättning är signifikant på 5 % nivå. Effekterna är inte rensade för provbetygs- eller fördelningseffekter.

Tabell B2.2b. Skillnader i betygssättning i förhållande till de nationella proven för elever i skolor som skiljer sig med en standardavvikelse i genomsnittligt provbetyg. Uppdelat på kommunala respektive olika fristående skolkoncerner. Genomsnitt 2013–2015.

	Kommunala skolor	Skolloncern 1	Skolkoncern 2	Skolkoncern 3	Övriga fristående skolor
ma	-0,35	-0,38	-0,41	-0,37	-0,36
en	-0,12	0,07	-0,05	-0,03	-0,07
sv	-0,21	0,04	-0,19	-0,20	-0,17
fy	-0,44	-0,34	-0,51	-0,47	-0,39
bi	-0,44		-0,38	-0,40	-0,43
ke	-0,43		-0,26		-0,40
ge	-0,33		-0,42	-0,43	-0,34
hi	-0,47		-0,57		-0,45
re	-0,43		-0,46		-0,36
sh	-0,28		-0,33	0,03	-0,34

Rödfärgade värden innebär att den systematiska skillnaden i betygssättning inte är signifikant på 5 % nivå. Alla övriga värden innebär att den systematiska skillnaden i betygssättning är signifikant på 5 % nivå. Skolgrupper med färre än 5 skolor i något av åren 2013, 2014 eller 2015, redovisas ej. Effekterna är inte rensade för provbetygs-, eller fördelningseffekter.

B2.6 Andel av variationen i betygssättning mellan skolor som kan förklaras av huvudman

Tabell B2.3. Undre golv respektive övre tak för andel förklarad variation av skolans prestationsnivå respektive huvudman. Genomsnitt 2013–2015.

ämne	Undre golv* för andel som kan förklaras av:			Övre tak** för andel som kan förklaras av:		
	[pb_skola]	[hman]	[pb_skola] och [hman]	[pb_skola]	[hman]	[pb_skola] och [hman]
ma	23%	1,3%	24%	41%	1,9%	43%
en	4%	0,5%	5%	8%	1,0%	9%
sv	11%	0,4%	12%	24%	0,9%	25%
fy	31%	1,4%	33%	50%	1,8%	52%
bi	33%	1,0%	34%	52%	1,3%	53%
ke	35%	1,7%	37%	52%	1,7%	54%
ge	25%	0,9%	26%	39%	1,3%	40%
hi	31%	0,7%	32%	51%	0,7%	51%
re	31%	1,6%	32%	48%	2,0%	50%
sh	20%	3,6%	24%	29%	3,6%	33%
Genomsnitt	24%	1,3%	26%	39%	1,6%	41%

*beräknad utifrån betygsssteg

**beräknad utifrån betygspoäng

B2.7 Andelen behöriga lärare på skolan och betygssättning

Hur andelen behöriga lärare på skolan påverkar betygssättningen har inte ingått i den ursprungliga analysmodellen som både kontrollerar för övriga variabler och tar hänsyn till provbetyg- och fördelningseffekter. Resultaten som redovisas nedan bygger istället på en tilläggsanalys på skolnivå för varje enskilt ämne och gällande 2015.

$$[\text{ää_diffpo_skola}] = b_0 + b_1[\text{pb_skola}] + b_2[\text{hman}] + b_3[\text{andel_leg_beh}] + b_4[\text{sociox_skola}] \quad (2)$$

Där:

$[\text{ää_diffpo_skola}]$ = skolans genomsnittliga avvikelse mellan betyg och provbetyg, uttryckt i betygspoäng (i respektive ämne).

$[\text{pb_skola}]$ = skolans genomsnittliga provbetyg (i respektive ämne).

$[\text{hman}]$ = kommunal eller fristående skola.

$[\text{andel_leg_beh}]$ = andelen lärare med lärarlegitimation och behörighet på skolan (i det aktuella ämnet).⁹³

$[\text{sociox_skola}]$ = motsvarande socioekonomiska index som i tidigare analyser på elevnivå men aggregerat till skolnivå (för respektive ämne).

⁹³ Variabeln är hämtad från SIRIS: <https://www.skolverket.se/skolutveckling/statistik/sok-statistik-om-forskola-skola-och-vuxenutbildning?sok=SokD&niva=S&omr=personal&exp=102&lasar=2014&uttag=1>

Genom att studera andelen förklarad variation i denna skolnivåmodell ges en approximativ skattning av hur mycket andelen behöriga lärare bidrar till att förklara skillnader mellan skolor i betygssättning (se tabell B2.4 nedan).

Resultaten visar på ett negativt samband för de flesta ämnen, dvs. att en högre andel behöriga lärare samvarierar med en mer restriktiv betygssättning. Andelen av den totala variationen mellan skolor i betygssättning som kan förklaras av skillnader i andelen behöriga lärare uppgår dock inte till mer än 0,5 procent (i genomsnitt).

Tabell B2.4 Andel av variation på skolnivå i genomsnittlig avvikelse (mellan betyg och provbetyg) som kan förklaras av [pb_skola], [hman] respektive [andel_leg_beh]

	Referensvärden (=tabell B2.3)		Regression på skolnivå utan [sociox_skola]			Regression på skolnivå med [sociox_skola]	
	[pb_skola]	[hman]	[pb_skola]	[hman]	[andel_leg_beh]	[hman]	[andel_leg_beh]
ma	41,3%	1,9%	45,5%	0,5%	0,6%	0,2%	0,8%
en	7,7%	1,0%	6,0%	0,9%	1,0%	0,9%	0,9%
sv	23,8%	0,9%	14,0%	0,9%	0,6%	0,6%	1,0%
fy	50,1%	1,8%	49,4%	0,8%	0,4%	0,2%	0,5%
bi	51,9%	1,3%	54,0%	0,6%	0,5%	0,1%	1,2%
ke	52,0%	1,7%	47,6%	0,8%	0,0%	0,3%	0,1%
ge	39,2%	1,3%	38,7%	1,5%	0,0%	1,0%	0,1%
hi	50,6%	0,7%	48,4%	0,1%	0,0%	0,0%	0,0%
re	48,0%	2,0%	30,8%	2,4%	0,2%	0,7%	0,3%
sh	29,5%	3,6%	37,6%	1,5%	0,3%	1,1%	0,2%
genomsnitt	39,4%	1,6%	37,2%	1,0%	0,4%	0,5%	0,5%

= negativt samband (lärare med behörighet mer restriktiva)

= positivt samband (lärare med behörighet mindre restriktiva)

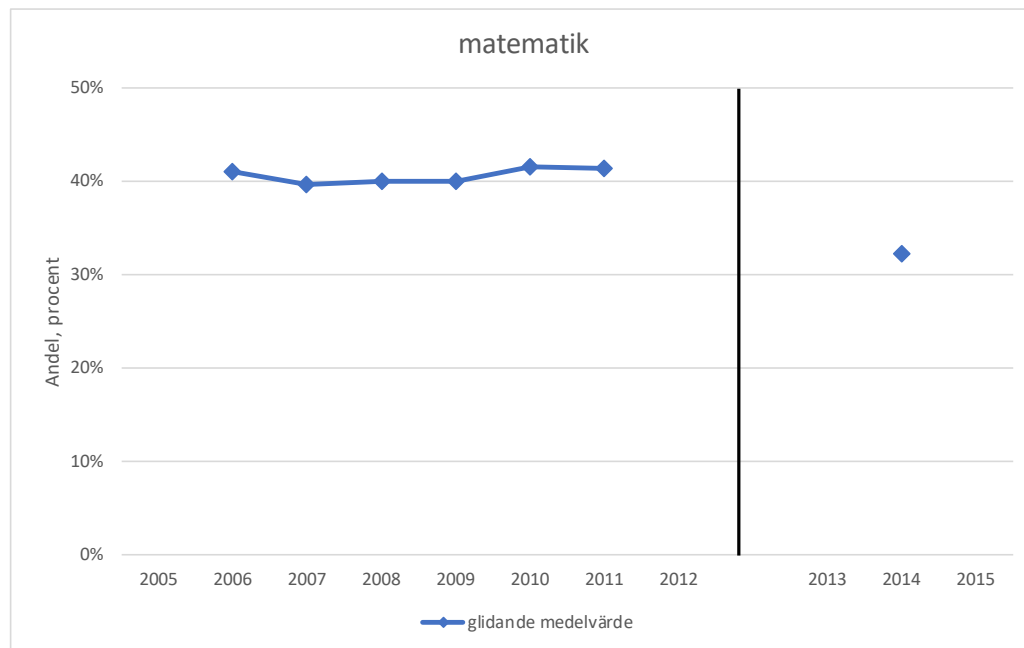
= ej signifikant samband

B2.8 Relativ betygssättning över tid

I rapportens huvudtext redovisas hur stor andel av variationen mellan skolor i betygssättning som kan förklaras av relativ betygssättning, dvs förklaras av skolornas genomsnittliga provbetyg i respektive ämne. Dessa analyser byggde på ett genomsnitt över åren 2013–2015 och gav ingen information om hur denna andel eventuellt har förändrats över tid.

I figur B2.4 redovisas den andel av variationen mellan skolor i betygssättning som kan förklaras av skolans genomsnittliga provbetyg för åren 2006–2014, utifrån ett glidande medelvärde (3 år). Detta innebär att värdena beräknade på samma sätt som i huvudtexten.⁹⁴ I excelfilen redovisas motsvarande tidserier även för engelska, svenska samt no-ämnena.

Figur B2.4 Andel av variationen i skolors betygssättning som kan förklaras av skolans genomsnittliga prestationsnivå (skolans genomsnittliga provbetyg), matematik.



Det glidande medelvärdet är baserat på de 3 år som är centrerade kring respektive år, t.ex. är värdet i figuren för år 2007 ett genomsnitt av åren 2006, 2007 och 2008. Dessa underliggande årsvärden (som ej syns i figuren) är i sin tur ett genomsnitt av en "undre golv-" respektive "övre takskattning". Observera att betygsskalan ändrades mellan 2012 och 2013, vilket innebär att det ej går att analysera förändring över tidsperioder som går över denna brytpunkt.

⁹⁴ Det som i huvudtexten bygger på åren 2013–2015 motsvarar i figur B2.4 värdet för 2014.

Referenser

- Alm Fjellborg, Andreas & Lena Molin (2018). Vilka typer av uppgifter gynnar elever som följer kursplanen i svenska som andraspråk? En undersökning med data från de nationella proven i geografi. Vol. 12, Nr. 4. Art. 5, *Acta Didactica Norge*.
- Ângelo, Catarina & Ana Balcão Reis (2018). Gender gaps in different grading systems. *FEUNL Working Paper Series* wp625, Universidade Nova de Lisboa, Faculdade de Economia.
- Calsamiglia, Caterina & Annalisa Loviglio (2018). *Grading on a Curve: When Having Good Peers is not Good*, Working Papers, CEMFI.
- Diamond, Rebecca & Petra Persson (2016). *The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests*. SIEPR Discussion Paper No. 16-003, Stanford University.
- Eklöf, Hanna & Mikaela Nyroos (2013). Pupil perceptions of national tests in science: perceived importance, invested effort, and test anxiety. *European Journal of Psychology of Education* (2013) 28:497–510.
- Fischer, F., Schult, J., and Hell, B. (2013). Sex differences in secondary school success: why female students perform better. *European Journal of Psychology of Education* 28, 529–543.
- Fredriksson Peter & Jonas Vlachos (2011). Reformer och resultat: Kommer regeringens utbildningsreformer att ha någon betydelse? Finanspolitiska rådet. *Studier i finanspolitik* 2011/3.
- Grönqvist, Hans & Susan Niknami (2017). *Ankomst och härkomst – en ESO-rapport om skolresultat och bakgrund*. ESO-rapport 2017:3.
- Gustafsson, Jan-Eric & Gudrun Ericksson (2013). To trust or not to trust? — teacher marking versus external marking of national tests. *Educational Assessment, Evaluation & Accountability*, Feb 2013, Vol. 25 Issue 1, p.69.
- Gustafsson, Jan-Eric & Gudrun Ericksson (2018). Nationella prov i Sverige – tradition, utmaning, förändring Vol. 12, Nr. 4. Art. 2, *Acta Didactica Norge*.
- Gustafsson, Jan-Eric, Christina Cliffordsson & Gudrun Erickson (2014). *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. SNS förlag.
- Heissel, Jennifer A, Emma K. Adam, Jennifer L. Doleac, David N. Figlio, Jonathan Meer (2018). *Testing, Stress, and Performance: How Students Respond Physiologically to High-Stakes Testing*. NBER Working Paper No. 25305.
- Heller Sahlgren, Gabriel & Henrik Jordahl (2016). *Information – ett verktyg för ett bättre skolsystem*. SNS förlag.
- Hinton, Daniel P & Helen Higson (2017). A large-scale examination of the effectiveness of anonymous marking in reducing group performance differences in higher education assessment. *PLoS ONE* 12(8): e0182711.

- Holmlund, Helena, Josefin Häggblom, Erica Lindahl, Sara Martinson, Anna Sjögren, Ulrika Vikman & Björn Öckert (2014). *Decentralisering, skolval och fristående skolor: resultat och likvärdighet i svensk skola*. IFAU, Rapport 2014:25.
- Kilpatrick, Jeremy & Bengt Johansson (1994). Standardized mathematics testing in Sweden: The legacy of Frits Wigforss. *Nordic Studies in Mathematics Education*, 2(1), 6-30.
- Kiss, David. (2011). Are Immigrants and Girls Graded Worse? Results of a Matching Approach. *Education Economics* 21 (5): 447–463.
- Klapp Lekholm, Alli och Christina Cliffordsson (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation*, 15(1), 1–23.
- Klapp Lekholm, Alli (2011). Effects of School Characteristics on Grades in Compulsory School, *Scandinavian Journal of Educational Research*, 55:6, 587-608
- Klapp, Alli (2018). *Betyg – deras funktioner och vad de mäter*. Skolverket
https://www.skolverket.se/download/18.6011fe501629fd150a277f4/1537272233334/klapp_7-9-gy-vux.pdf
- Lind Pantzare, Anna (2018). *Dimensions of validity Studies of the Swedish national tests in mathematics*. Avhandling. Institutionen för tillämpad utbildningsvetenskap. Umeå Universitet.
- Lindahl, Erica (2007). *Systematiska avvikelser mellan slutbetyg och provresultat – spelar elevens kön och etniska bakgrund roll?* Rapport 2007:22. IFAU – Institutet för arbetsmarknadspolitiska studier.
- Lindahl, Erica (2016). Are teacher assessments biased? Evidence from Sweden. *Education Economics*, 24(3): pp.224–238.
- Lundahl, Christian (2009). *Viljan att veta vad andra vet. Kunskapsbedömning i tidigmodern, modern och senmodern skola*. Arbetslivsinstitutet och Uppsala universitet.
- Lundahl, Christian (2009). *Varför nationella prov? – framväxt, dilemman, möjligheter*. Studentlitteratur.
- Lundahl, Christian, Magnus Hultén, Alli Klapp & Larissa Mickwitz (2015). *Betygens geografi. Forskning om betyg och summativa bedömningar i Sverige och internationellt*. Vetenskapsrådets rapporter.
- Lundahl, Christian, Magnus Hultén & Sverre Tveit (2017). *Betygssystem i internationell belysning*. Skolverket.
- Lärarnas riksförbund (2011). *Betygssättning under påverkan*.
- Löfstedt, Arne (2018). Det första nationella provet i samhällskunskap - en studie i bedömersamstämmighet. Vol. 12, Nr. 4. Art. 13, *Acta Didactica Norge*.
- Malouff, John (2008). Bias in Grading, *College Teaching*, 56:3, 191-192
- Malouff, John & Einar B Thorsteinsson (2016). Bias in grading: A meta-analysis of experimental research findings. *Australian Journal of Education* 2016 Vol. 60(3) 245–256.

- Reardon, S. F., Kalogrides, D., Fahle, E. M., Podolsky, A., & C., R. (2018). The Relationship Between Test Item Format and Gender Achievement Gaps on Math and ELA Tests in Fourth and Eighth Grades. *Educational Researcher*, 47(5), 284–294.
- Riksrevisionen (2011). *Lika betyg, lika kunskap? En uppföljning av statens styrning mot en likvärdig betygssättning i grundskolan*. Rapport 2011:23.
- Riksrevisionen (2018). *Otillåten spridning av nationella prov – vad gör Skolverket och Skolinspektionen?* Rapport 2018:34.
- Sjögren, Anna (2017). Könsskillnader i utbildning. *SNS Analys* nr 42, augusti 2017.
- Skolinspektionen (2018). *Rättvis och likvärdig betygssättning i grundskolan? Granskning av rektorers och huvudmäns arbete. Tematisk kvalitetsgranskning 2018*.
- Skolverket (1998). *Grundskola för bildning – kommentarer till läroplan, kursplaner och betygskriterier*.
- Skolverket (2004). *Det nationella provsystemet i den målstyrda skolan*.
- Skolverket (2005). *Hur används kursplanerna i skolan? - några resultat från en enkätundersökning 2004*.
- Skolverket (2006). *Könsskillnader i måluppfyllelse och utbildningsval*. Rapport 287.
- Skolverket (2007). *Provbetyg – Slutbetyg – Likvärdig bedömning?* Rapport 300.
- Skolverket (2008). *Central rättning av nationella prov*.
- Skolverket (2009). *Likvärdig betygssättning i gymnasieskolan?* Rapport 338.
- Skolverket (2011). *Planering och genomförande av undervisningen. Skolverkets allmänna råd – för grundskolan, grundsärskolan, specialskolan och sameskolan*.
- Skolverket. (2015a). *Provbetygens stabilitet, årskurs 9*. Dnr 2015:4.
- Skolverket. (2015b). *Provpoängens tillförlitlighet*. Dnr 2015:5.
- Skolverket (2015c). *Skolreformer i praktiken. Hur reformerna landade i grundskolans vardag 2011–2014*. Rapport Nr 418.
- Skolverket (2016a). *Utvärdering av den nya betygsskalan samt kunskapskravens utformning*. Regeringsuppdrag, Dnr 2014:892.
- Skolverket (2016b). Yttrande över betänkandet ”Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsbedömning” SOU 2016:25. Dnr 2016:717.
- Skolverket (2017). *Skolverkets systemramverk för nationella prov*.
- Skolverket (2018a). *Från gymnasieskola till högskola – en registerstudie*. Rapport 466.
- Skolverket (2018b). *Betyg och betygssättning*. Skolverkets allmänna råd med kommentarer.
- Skolverket (2018c). *Analys av familjebakgrundens betydelse för skolresultaten och skillnader mellan skolor*. Rapport 467.
- Skolverket (2019). *Relationen mellan provresultat och betyg i grundskolans årskurs 6 och årskurs 9 2018*.

- SOU 2010:96 *Riktiga betyg är bättre än höga betyg. Förslag till omprövning av betyg.*
- SOU 2016:25 del 1. *Likvärdigt, rättvist och effektivt – ett nytt nationellt system för kunskapsbedömning.*
- SOU 2016:25 del 2 bilagor. *Likvärdigt, rättvist och effektivt – ett nytt nationellt system för kunskapsbedömning.*
- Tengberg, Michael & Gustaf B Skar (2017). Hur tillförlitligt är det nationella provet i läsning i åk 9? *Utbildning & Demokrati* 2017, vol 26, nr 2, 113–137
- Terrier, Camille (2015). *Giving a Little Help to Girls? Evidence on Grade Discrimination and its Effect on Students' Achievement.* CEP Discussion Paper No 1341 March 2015. London School of Economics.
- Thorsen, Cecilia & Christina Cliffordson (2012). Teachers' grade assignment and the predictive validity of criterion-referenced grades, *Educational Research and Evaluation*, 18:2, 153-172
- Thorsen, Cecilia (2014). Dimensions of Norm-Referenced Compulsory School Grades and their Relative Importance for the Prediction of Upper Secondary School Grades, *Scandinavian Journal of Educational Research*, 58:2, 127-146
- Tyrefors Hinnerich, Björn & Jonas Vlachos (2012). "Systematiska skillnader mellan interna och externa bedömningar av nationella prov", bilaga 4 i <http://www.skolinspektionen.se/Documents/omrattning/omrattning-2012-resultatbilaga.pdf>.
- Tyrefors Hinnerich, Björn & Jonas Vlachos (2013). "Systematiska skillnader mellan interna och externa bedömningar av nationella prov – en uppföljningsrapport", bilaga 5 till Skolinspektionen (2013) Olikheterna är för stora: Omrättning av nationella prov i grundskolan och gymnasieskolan redovisning av regeringsuppdrag dnr U2011/6544/GV.
- Tyrefors Hinnerich, Björn, Erik Höglén och Magnus Johannesson. (2011). Are Boys Discriminated in Swedish High Schools? *Economics of Education Review* 30 (4): 682–690.
- Tyrefors Hinnerich, Björn, Erik Höglén & Magnus Johannesson (2015). Discrimination against students with foreign backgrounds: evidence from grading in Swedish public high schools- *Education Economics*, 23:6, 660-676.
- Tyrefors Hinnerich, Björn & Jonas Vlachos (2017). The impact of upper-secondary voucher school attendance on student achievement. Swedish evidence using external and internal evaluations. *Labour Economics*
- Wikström, Cristina & Wikström, Magnus (2017). Group differences in student performance in the selection to higher education: tests vs grades. *Frontiers in Education*, 2: 45