

Skolverket

Provbetyg – Slutbetyg- Likvärdig bedömning?

*En statistisk analys av sambandet mellan nationella prov och
slutbetyg i grundskolans årskurs 9, 1998-2006*

Rapport 300

Förord

Hur likvärdigt sätts betygen i dagens svenska skola? Ger samma kunskaper samma betyg oavsett lärare, skola, kommun eller huvudman?

Skolverket gjorde år 2000 en nationell granskning av betygsättningen som visade på betydande brister när det gäller likvärdigheten. Denna granskning bidrog till att det under senare års vidtagits en rad statliga insatser för en mer likvärdig betygsättning.

Har dessa insatser lett till att betygsättningen idag är mer likvärdig jämfört med hur det stod till i slutet av 1990-talet när betygssystemet var relativt nytt?

Dessa två grundläggande frågeställningar diskuteras i denna rapport. Det verktyg som används är de insamlade provresultaten i svenska, matematik och engelska i årskurs 9. Provresultaten används som en indikator på likvärdig betygsättning, både över tid och mellan olika skolor och kommuner.

Resultaten visar att det finns mer att göra för att betygsättningen skall bli tillfredsställande ur ett likvärdighetsperspektiv. I rapporten diskuteras, och i viss mån prövas, ett antal tänkbara *förklaringar* till skillnader mellan skolor när det gäller provresultat och betyg. Redogörelsen av förklaringarna är tänkt som ett stöd för en diskussion om likvärdig betygsättning för alla som är verksamma inom skolområdet på både lokal och nationell nivå. Rapporten visar att alla verksamma inom skolan kan och behöver göra insatser för att åstadkomma en mer likvärdig betygsättning.

Det finns en sammanfattning av rapport 300 med samma namn i tryckt form samt att ladda ner från Skolverkets webbplats: www.skolverket.se

Rapporten är skriven av Anders Auer och Jonas Sandqvist vid Skolverkets enhet för resultatbedömning. I projektgruppen ingick även Daniel Gustafsson, Kristian Ramstedt, Tommy Lagergren, Karin Hector-Stahre, Jan Engstedt, Wolfgang Dietrich, Mikael Henningsson, Cecilia Bergström, Lars Ullén och Jesper Antelius. Ytterligare ett antal personer både inom och utom Skolverket har läst och bidragit med synpunkter på manuskript mm.

Stockholm i juni 2007

Ann Carlsson-Ericsson
Avdelningschef
Avdelningen för läroplansfrågor

Jonas Sandqvist
Projektledare
Enheten för resultatbedömning

Innehållsförteckning

Kapitel 1. Inledning	8
1.1 Inledning och bakgrund	8
1.2 Dagens betygssystem	9
1.3 Statens verktyg för främjande av en likvärdig betygsättning	9
1.4 Det nationella provsystemet	10
<i>Syftet med de nationella proven</i>	
<i>De nationella proven funktion och begränsningar</i>	
1.5 Hur kan en studie av relationen provbetyg – slutbetyg ge information om likvärdighet betygsättning?	12
1.6 Rapportens syfte och avgränsningar	14
1.7 Datamaterial	15
1.8 Nettoavvikelser – ett centralt begrepp	15
1.9 Disposition	16
1.10 Förklaringar till avvikelser mellan provbetyg och slutbetyg på skolnivå	18
Kapitel 2. Relationen provbetyg – slutbetyg i ett nationellt perspektiv	22
2.1 Nationell betygsfördelning för provbetyg respektive slutbetyg	22
2.2 Sammanfattning	26
Kapitel 3. Hur mycket styr proven på elevnivå?	27
3.1 Konsekvenser för likvärdighet i betygsättning	30
3.2 Sammanfattning	30
Kapitel 4. Relationen provbetyg – slutbetyg på lokal nivå: Skolor och Kommuner	31
4.1 Kopplingen mellan provbetyg och slutbetyg på skolnivå	31
<i>Har spridningen i skolors nettoavvikelse ökat eller minskat över tid?</i>	
<i>Är det samma skolor som har höga nettoavvikelser från år till år?</i>	
4.2 Kopplingen mellan provbetyg och slutbetyg på kommunnivå	38
4.3 Speglar skillnader i skolors nettoavvikelser faktiska skillnader i kunskaper?	40
<i>Hur klarar sig elever från grundskolor med höga nettoavvikelser senare på gymnasiet?</i>	
<i>Hur klarar sig elever från grundskolor med höga nettoavvikelser enligt PISA?</i>	
4.4 Hur ser relationen slutbetyg – provbetyg ut mellan klasser?	43
4.5 Sammanfattning	44

Kapitel 5. Vilka är det som får ett högre slutbetyg än provbetyg?	45
5.1 Betygsprofil – nationell bild	45
5.2 Betygsprofiler för skolor med hög respektive låg nettoavvikelse	46
5.3 Sammanfattning	50
Kapitel 6. Konkurrens, fristående skolor och relationen provbetyg – slutbetyg	51
6.1 Konkurrensutsatta skolor	51
6.2 Fristående skolor	54
6.3 Sammanfattning	57
Kapitel 7. Relationen provbetyg-slutbetyg och skolans genomsnittliga prestationsnivå	59
<i>Nettoavvikelse och skolors genomsnittliga provbetyg – en första inblick</i>	
<i>Nettoavvikelse och skolors genomsnittliga provbetyg – efter kontroll för andra bakgrundsfaktorer</i>	
7.1 Sammanfattning	62
Kapitel 8. Resultat och diskussion	63
8.1 Inledning	63
8.2 Rapportens viktigaste resultat	63
8.3 Förklaringar av resultaten	66
<i>Varför ingen förändring över tid?</i>	
<i>Vilka förklaringar till skillnader mellan skolor är mest trovärdiga?</i>	
<i>Förklaringar som bedöms som mindre troliga</i>	
<i>Förklaringar som bedöms som mer troliga</i>	
8.4 Slutdiskussion	79
Referenser	83

Kapitel 1. Inledning

1.1 Inledning och bakgrund

Det är en självklarhet att alla aktörer inom skolväsendet ska verka för att Sveriges elever ska få likvärdiga *förutsättningar* att utveckla sina kunskaper. I Skollagen (1 kap 2 §) klargörs att:

”Utbildningen skall inom varje skolform vara likvärdig, varhelst den anordnas i landet.”

Men även om det skulle gå att nå en fullständig likvärdighet när det gäller förutsättningarna för kunskapsutvecklingen så kvarstår frågan om att få till stånd en likvärdig bedömning och betygsättning. En likvärdig och rättvis betygsättning är en förutsättning för att betyg skall vara jämförbara och visa vad eleverna har för kunskaper. Att betygsättningen görs korrekt och i enlighet med bestämmelserna är också viktigt för varje elevs rättssäkerhet.

Att eftersträva en likvärdig och rättvis betygsättning är en skyldighet och en uppgift för *skolan* och *skolhuvudmannen*. På nationell nivå ligger det övergripande ansvaret på *staten* att tillhandahålla ett skolsystem med mål och betygskriterier som ger förutsättningar för skolhuvudmän, rektorer och i slutändan lärare att ge rättvisa och likvärdiga betyg.

Men hur likvärdiga är betygen i dagens svenska skola? Ges samma prestationer och kunskaper av elever samma betyg oavsett lärare, skola eller kommun (huvudman)?

Skolverket gjorde år 2000 en nationell granskning av betygsättningen som visade på betydande brister när det gäller likvärdigheten.¹ Denna granskning bidrog till att det under senare års vidtagits en hel rad statliga insatser för en mer likvärdig betygsättning (se vidare 1.3 nedan)².

Har dessa insatser lett till att betygsättningen idag är mer likvärdig jämfört med hur det stod till i slutet av 1990-talet när betygssystemet var relativt nytt?

Dessa två grundläggande frågeställningar ska diskuteras i denna rapport. Det verktyg som används är de insamlade provresultaten i svenska, matematik och engelska i årskurs 9. Provresultaten används som en indikator på likvärdig betygsättning, både över tid och mellan olika skolor och kommuner.

Definition av likvärdig betygsättning

Skolverket formulerade år 2004 ett handlingsprogram för en likvärdig betygsättning.³ Där definieras vad Skolverket menar med likvärdig betygsättning:

En rättvis betygsättning innebär att det betyg en elev fått i ett visst ämne eller i en viss kurs ska visa elevens kunskaper och färdigheter i det som ska ingå i kursen enligt kursplanen och som motsvarar betygskriterierna för betyget. Med likvärdig menas att måttstocken för bedömningen är densamma för alla elever. Ett betyg i en klass ska motsvara ett likadant betyg i en annan klass.

Grundläggande för en rättssäker och likvärdig betygssättning är att de som sätter betygen och som utfärdar betygsdokumentet, dvs. lärare och rektorer, har tillräckliga kunskaper. De måste känna till och kunna tillämpa kursplaner, betygskriterier och andra

¹ Skolverket 2000.

² Riksrevisionen 2004.

³ Skolverket 2004d.

bestämmelser. Men enbart att känna till reglerna räcker inte, det krävs också ett aktivt lokalt kursplanearbete där betygskriterierna tolkas och anpassas utifrån det stoff och undervisningsmetoder som används. Detta förutsätter en bred diskussion mellan lärare såväl inom som mellan olika skolor om hur man ser på målpuppfyllelse och kunskapskvaliteter. Samrådet krävs för att trygga en likvärdig och rättvis betygssättning.

1.2 Dagens betygssystem

Hösten 1994 infördes dagens målrelaterade betygssystem där lärarna skall betygssätta varje elev i förhållande till betygskriterier som motsvarar fastställda kunskapsmål. De första elever som fick slutbetyg enligt det nya betygssystemet lämnade gymnasieskolan 1997 och grundskolan 1998.

Det målrelaterade betygssystemet ersatte det ”relativa” betygssystemet.⁴ I det relativa betygssystemet styrdes betygssättningen genom obligatoriska standardprov (grundskolan) i svenska, matematik och engelska. Standardproven skulle vara en riktningsgivare för betygsssteg tre. Vid betygssättningen skulle läraren ta hänsyn till hur klassens genomsnittliga resultat på standardproven låg till i förhållande till rikssnittet.

I dagens målrelaterade betygssystem skall lärarna vid betygssättningen utgå från kursplanernas förutbestämda uppnåendemål och kriterier för varje betygsssteg. Rent teoretiskt behövs då egentligen inget nationellt provsystem för att kalibrera betygssättningen. I praktiken finns dock ett tolkningsutrymme i mål och betygskriterier som gjort att staten ansett det nödvändigt att även inom ramen för vårt nuvarande betygssystem tillhandahålla nationella prov i svenska, svenska som andraspråk, matematik och engelska. Likvärdigheten i betygssättningen är beroende av i vilken utsträckning lärare gör en samstämmig tolkning och tillämpning av mål och betygskriterier. De nationella proven finns som ett stöd för lärarnas tolkning.

1.3 Statens verktyg för främjande av en likvärdig betygssättning

För att öka förutsättningarna för att betygssättningen skall bli så samstämmig som möjligt har staten ett antal verktyg till sitt förfogande. De mest centrala statliga verktygen för en likvärdig bedömning är de nationella kursplanerna med dess uppnåendemål och betygskriterier som varje lärare är skyldig att följa vid betygssättningen.⁵ Staten bidrar också till en likvärdig bedömning genom att ställa krav på ”en god kunskap i betygssättning” för en lärarexamen⁶ och genom fortbildning av lärare när det gäller betygssättning.⁷ Skolverket har också givit ut allmänna råd och kommentarer till betygsbestämmelserna, svar på betygsfrågor samt bedömningsexempel som stöd för betygssättningen.⁸ En provbank håller på att

⁴ För en kort historisk genomgång av betygssystemen, se skolverkets hemsida <http://www.skolverket.se/betyg>. Mer utförliga beskrivningar finns i Tholin 2006 kap 5 och Selghed 2004 kap 3.

⁵ Lpo 94 punkt 2.7 samt grundskoleförordningen kap 7 § 8.

⁶ Högskoleförordningen Bilaga 2.

⁷ http://www.skolutveckling.se/kunskap_bedomning/bedomning_betyg/

⁸ <http://www.skolverket.se/betyg>

byggas ut som bland annat har till syfte stödja betygsättningen i andra ämnen än de som har nationella prov.⁹

Skolverket har även en utbildningsinspektion vars uppgift är att ta reda på hur väl en verksamhet fungerar i förhållande till bestämmelserna i skollagen, skolförordningarna och de nationella läroplanerna. I inspektionen ingår också att granska betygsättningen.¹⁰

Ett ytterligare viktigt verktyg för en likvärdig betygsättning är det nationella provsystemet. För grundskolan finns sedan 1998 ämnesprov i matematik, engelska, svenska och svenska som andraspråk i årskurs 9 till stöd för betygsättningen.¹¹ Det är denna del av de statliga insatserna för en likvärdig betygsättning som denna rapport utgår ifrån.

1.4 Det nationella provsystemet

Syftet med de nationella proven

Allt sedan den nya läroplanen och det nya betygssystemet infördes 1994 står i grundskoleförordningen (7 kap 10 §) att:

Ämnesprov i svenska, engelska och matematik skall användas i slutet av årskurs 9 för att bedöma elevernas kunskapsutveckling och som stöd för betygsättning.

I det första provuppdraget till Skolverket 1994 skriver regeringen att: ”Syftet med ämnesproven skall vara att kontrollera om eleven uppnått kursplanernas kunskapskrav” och vidare: ”ge underlag för lärarens betygsättning av elevernas kunskaper.”¹² I provuppdraget från 1999 används liknande formuleringar. Ämnesproven ska: ”ge läraren stöd vid bedömning av om eleven nått uppställda mål och ge underlag för lärarens betygsättning av elevernas kunskaper.”¹³ Även i det nuvarande provuppdraget från 2004¹⁴ formuleras de nationella provens syften på ett liknande sätt. De nationella provens olika syften sammanfattas i provuppdraget på följande sätt:

- bidra till ökad måluppfyllelse för eleverna,
- förtydliga målen och visa på elevers starka och svaga sidor (diagnostisk funktion),
- konkretisera kursmål och betygskriterier (kommentarmaterial)
- stödja en likvärdig och rättvis bedömning och betygssättning (betygsstöd) samt
- ge underlag för en analys av i vilken utsträckning kunskapsmålen nås på skolnivå, på huvudmannanivå och på nationell nivå (uppföljning).

Formuleringen i grundskoleförordningen (ovan) visar att det som starkast betonats när det gäller de nationella proven är det individuella stödet till läraren vid betygsättningen. Ett ytterligare exempel är att när det första provuppdraget formulerades till Skolverket 1994 så fanns en tanke om att de nationella proven

⁹ Utbildningsdepartementet 2004.

¹⁰ Skolverket 2005b.

¹¹ <http://www.skolverket.se/sb/d/284>

¹² Utbildningsdepartementet 1994.

¹³ Utbildningsdepartementet 1999.

¹⁴ Utbildningsdepartementet 2004.

kanske bara behövdes tillfälligt under några år tills lärarna vant sig vid att använda och tolka de nationella målen och kriterierna.¹⁵

De nationella provens funktion och begränsningar

De insamlade provresultaten är alltså även tänkta att ge underlag för en analys av i vilken utsträckning kunskapsmålen nås på skolnivå, på huvudmannanivå och på nationell nivå. Dessutom kan de utgöra ett underlag för analys av bedömning och betygsättning utifrån ett likvärdighetsperspektiv, både på lokal och på nationell nivå. Det betyder dock inte att proven ska ha en styrande effekt på det viset att en lärare förväntas sätta betyg på sina elever eller sin klass utifrån en viss relation till de nationella provresultaten. Staten har inte uttryckt någon uppfattning om hur nära slutbetygen ska ligga provbetygen - varken för enskilda elever eller för klasser, skolor eller kommuner. Det finns ur prov- och betygssystemets perspektiv flera goda skäl för en lärare att sätta ett annat slutbetyg på en enskild elev än vad resultatet blev på det nationella provet. En total överensstämmelse mellan provbetyg och slutbetyg för varje enskild elev i en klass eller på en skola skulle betyda att proven fungerar som examensprov, vilket de inte är tänkt att göra.

Skolverket har i flera sammanhang uttryckt stark oro för tendensen att de nationella proven får en sådan vikt för betygsättningen av enskilda elever att de kommit att likna examensprov.¹⁶ Skolverket har bland annat påtalat att de nationella proven inte är utformade för att fungera som examensprov. De nationella proven prövar inte alla mål i kursplanen och de prövar inte alla mål lika mycket. Det är helt enkelt för komplicerat att konstruera och genomföra ett enda prov som skulle ge en perfekt bild av en enskild elevs kunskaper i ett ämne.

De nationella proven är inte heller helt tillförlitliga på aggregerad nivå som ett mått på kunskaperna i ett ämne.¹⁷ De är inte exakta mätinstrument från år till år. Bland annat måste provfrågorna bytas ut varje år. Provresultaten varierar också mellan ett år till ett annat i högre utsträckning än slutbetygen, vilket sannolikt indikerar att betygsgränserna i proven varierar, snarare än att den totala kunskapsnivån för årskullen skulle variera i sådan utsträckning.

Det är också viktigt att påpeka att de nationella proven enbart omfattar tre skolämnen. Det gör att de nationella provens betygsstödjande funktion är begränsad. Även den mest centrala funktionen hos de nationella proven: att vara betygsstödjande genom att konkretisera mål och betygskriterier - gäller enbart de tre provämnena. När en elev söker till gymnasiet räknas ett meritvärde ut för de 16 bästa betygen.¹⁸ De nationella proven har, även om de skulle tillämpas som examensprov, därmed även en ganska begränsad påverkan på elevernas meritvärden. Men samtidigt kan de nationella proven spela en annan viktig roll för många elever genom att de finns just i de tre behörighetsgivande ämnena till gymnasiet nationella program.¹⁹ Utan ett godkänt betyg i matematik, svenska och engelska är en elev hänvisad till gymnasiets individuella program. De nationella proven kan därför ha en stor betydelse för att stödja lärares bedömning av om en elev är behörig att gå på gymnasiet.

¹⁵ Utbildningsdepartementet 1994.

¹⁶ Skolverket 2003 och Skolverket 2004a.

¹⁷ Skolverket 2006g s.24ff.

¹⁸ Gymnasieförordningen 6 kap 4 §

¹⁹ Skollagen 5 kap 5§

1.5 Hur kan en studie av relationen provbetyg – slutbetyg ge information om likvärdighet betygsättning?

Om nu de nationella proven är osäkra som mätinstrument och staten inte har fastslagit någon specifik relation mellan prov och betyg, varken för lärare i sin individuella bedömning av elever eller för skolor när de tittar på samlade prov- och betygsresultat – hur kan då en studie motiveras som just jämför prov och betyg mellan skolor?

Utgångspunkten för rapporten är att även om det inte finns någon fastslagen policy för hur skolor och huvudmän skall förhålla sig till resultaten från de nationella proven, och det därmed inte finns några möjligheter att fastställa om en enskild skola sätter för höga eller för låga betyg, så ger *variationen* i hur skolor sätter betyg i förhållande till sina provresultat en *indikation* på hur mycket tolkning och tillämpning av betygsriterier varierar på systemnivå.

En parallell är hur resultaten från internationella undersökningar, så som PISA, kan användas till att ta fram indikatorer på likvärdighet genom att studera variationen mellan skolor i förhållande till den totala variationen i elevresultat.²⁰ Detta anses fungera även om dessa instrument inte mäter exakt vad som ingår i de svenska kursplanerna samt att kunskaperna inte mäts på en skala som kan översättas till, eller direkt jämföras med, de svenska betygsstegen.²¹ Denna rapport argumenterar därför på liknande sätt när det gäller de nationella proven. Även med en medvetenhet om att de nationella proven inte prövar precis alla mål som ingår i kursplanerna och att ett nationellt prov aldrig kan ge en fullständig bild av en elevs kunskaper så bör ändå *variationen* i skolors relation mellan genomsnittliga provbetyg och slutbetyg kunna fungera som en *indikator* på systemnivå över i vilken grad betygsättningen är likvärdig.

Att se proven som en indikator på likvärdig bedömning är inget nytt. Det har redan tidigare gjorts av både Skolverket och utbildningsdepartementet i olika sammanhang. I en tidigare studie som baserades på data från 2001 och 2002 fann Skolverket relativt stora avvikelser mellan nationella prov och slutbetyg.²² Regeringen kommenterade rapportens resultat i det nya provuppdraget 2004 till Skolverket:

Provsystemet bör ge lärarna stöd vid bedömning och betygsättning, bl.a. genom att konkretisera kursmål och betygsriterier, och därmed bidra till en mer likvärdig betygsättning över landet. Detta är ytterst en fråga om elevernas rättssäkerhet. Att det finns brister vad gäller bedömning och betygsättning konstaterades redan i Skolverkets kvalitetsgranskning och tillsyn av bedömnings- och betygsfrågor år 2000.

I rapporten "Det nationella provsystemet – vad, varför och varthän?" (dnr. 01-2003:2038) bekräftar Skolverket denna bild. Där framgår att variationerna i relationen mellan provbetyg och slutbetyg mellan skolor, och för gymnasieskolan dessutom mellan olika program, är stora. Även om de nationella proven inte är avsedda som examensprov, och det således inte skall finnas en absolut överensstämmelse mellan provresultat och betyg, indikerar detta en bristande likvärdighet när det gäller bedömning och betygsättning.²³

²⁰ Skolverket 2004g

²¹ Skolverket 2006e, Skolverket 2006f.

²² Skolverket 2003 och Skolverket 2004a.

²³ Utbildningsdepartementet 2004.

Skolverket kommenterade samma resultat i Allmänna råd och kommentarer om Likvärdig bedömning och betygsättning 2004 ²⁴:

Skolverket har även gjort jämförelser av betygen och resultaten på nationella prov. Dessa jämförelser visar att vissa skolor eller lärare tenderar att sätta betyg som ligger betydligt högre eller lägre än vad som borde vara rimligt i förhållande till provresultaten. Sådana skillnader kan vara helt förklarliga när det gäller enskilda elever, men knappast när det gäller betygen i ett visst ämne i en hel skola. Jämförelserna indikerar att det finns en bristande likvärdighet i betygsättningen.

Både regeringen och Skolverket understryker alltså att de nationella provresultaten kan användas som en indikator för att belysa likvärdigheten i betygsättningen.

Under *Frågor och svar om det nationella provsystemet*²⁵ beskrivs på ett åskådligt vis skillnaden mellan provens betygsstödjande roll på individ- och gruppnivå:

Flera lärare på vår skola tycker de nationella proven är för lätta och sätter egna gränser för betygen. Är det riktigt?

Svar: De förslag till bedömning som finns i det bedömningsunderlag som följer med ett prov har tillkommit genom en bred utprövning. Elever på olika skolor runt om i Sverige har gjort proven året innan de genomförs över hela landet. Syftet med dessa utprövningar är att undersöka om proven fungerar på avsett sätt, om eleverna förstår materialet, om det kan missuppfattas, om det är av rimlig omfattning osv. Utprövningen ger också information som kan vara användbar vid bestämningen av förslag till gränser för de olika betygen.

Förslag till betygsgränser på ett nationellt prov bestäms utifrån de mål och betygskriterier som gäller för ämnet eller kursen av de institutioner som konstruerar proven. Detta görs i samverkan med en referensgrupp i vilken ingår aktiva lärare från olika skolformer, lärarutbildare, forskare och, när det gäller språkprov, personer med språket som modersmål. De vägledande förslagen till gränser för bedömningen av elevers prestationer på ett nationellt prov är således väl underbyggda och förankrade.

När lärarna och arbetslaget i den skola frågan gäller har tolkat betygskriterierna har de uppenbarligen gjort en strängare tolkning än provkonstruktörerna när det gäller kraven för ett visst betyg på provet. Det innebär att de egna eleverna får lägre betyg på provet än elever på andra skolor med likvärdiga prestationer. Detta är inte önskvärt. Betygen ska vara likvärdiga och rättvisa och ett av syftena med de nationella proven är att bidra till detta. Det kan då inte anses riktigt att enskilda skolor eller lärare frångår den nationella bedömningen därför att de har egna uppfattningar om vilka betygskriterier som ska gälla för olika betyg. Rättvisa mot eleverna kräver att ambitionen ska vara att alla lärare bedömer på ett likvärdigt sätt.

Om man som lärare eller som lärarlag anser att gränserna i de nationella proven ligger fel har man naturligtvis rätt att framföra dessa uppfattningar på olika sätt. I själva verket är det grundläggande för det system vi har att sådana frågor förs upp till offentlig diskussion inom lärarprofessionen. Men de egna eleverna ska inte betala priset i form av lägre betyg än andra elever.

Till detta ska läggas att ett prov inte förmår pröva alla mål som läraren har att ta hänsyn till vid sin betygsättning. Nationella och andra prov är bara en del i allt som läraren har att bedöma och väga in i ett slutbetyg. De nationella proven ska inte avgöra enskilda elevers betyg. En elevs slutbetyg kan vara högre eller lägre än betyget på det nationella provet. Däremot ska inte hela elevgrupper systematiskt ges lägre eller högre betyg än vad de nationella proven anger.

²⁴ Skolverket 2004e s.5.

²⁵ <http://www.skolverket.se/sb/d/298>

Även om många enskilda elever kan få höjda eller sänkta betyg i jämförelse med provresultaten bör avvikelserna på klass-, skol-, kommun och nationell nivå inte vara alltför stora. Stora skillnader mellan prov- och betygsresultat på aggregerad nivå är en indikation på att betygsättningen inte är likvärdig mellan olika skolor, klasser eller elevgrupper – att lärare inte tolkar mål och kriterier på liknande sätt eller använder olika grunder för betygsättningen. Det är just på detta sätt - som en *indikator* på likvärdig bedömning - som de nationella proven kommer att användas i denna rapport.

1.6 Rapportens syfte och avgränsningar

Syftet med rapporten är att bidra till en diskussion om likvärdig betygsättning utifrån en analys av relationen mellan provresultat och slutbetyg på nationell så väl som lokal nivå. Hur ser relationen mellan prov och betyg ut mellan olika skolor och kommuner? Vad kan det finnas för orsaker till olika avvikelser? Hur ser avvikelserna mellan provbetyg och slutbetyg ut över tid, sedan de infördes 1998 och fram till 2006? Vad ska man förvänta sig ha hänt och hur ska man i så fall tolka förändringar eller en brist på förändringar?

I jämförelse med den tidigare Skolverksrapporten ”*Det nationella provsystemet i den målstyrda skolan*” som också analyserade relationen provbetyg – slutbetyg, ges i föreliggande rapport framförallt ett tidsperspektiv på hur relationen mellan provbetyg och slutbetyg eventuellt har förändrats. Genom att genomföra i stort sett alla analyser för hela tidsperioden 1998-2006 kan en del resultat som föregående rapport funnit också bekräftas och därmed bli mer pålitliga. Dessutom görs en hel del fördjupade analyser för att med större säkerhet kunna dra slutsatser samt bekräfta eller förkasta vissa samband mellan skolors relation till provresultaten.

Samtliga analyser i denna rapport bygger på statistiska bearbetningar av data om slutbetyg och nationella provresultat från åren 1998-2006 för grundskolan. Analyserna är således av generell karaktär och söker finna övergripande mönster som eventuellt kan säga något om likvärdigheten i betygsättningen på systemnivå. Att analyserna baseras på kvantitativa data har sina begränsningar då det är svårt att få mera djupgående förklaringar till vissa resultat. I viss utsträckning har dock bilden kompletterats med resultat från studier som använt andra metoder. Skolverket har genomfört en enkätundersökning med specificerade frågor om betygsättning till lärare, rektorer och huvudmän.²⁶ Även en fallstudie på enskilda grundskolor med djupintervjuer har använts för att bidra till en ökad förståelse av hur betygsättningen går till och hur den relateras till provbetygen på en enskilda skola.²⁷

Rapporten gör inte på något vis anspråk att ge en heltäckande bild av likvärdighet i betygsättning. Inte ens när det gäller relationen mellan prov- och slutbetyg ges här en fullständig bild. En begränsning är att på samma sätt som rapporten visar att det finns variation i hur skolor sätter betyg finns det säkerligen också variation mellan klasser inom skolor, dvs. hur olika lärare inom samma skola sätter betyg och även kanske variation i hur samma lärare bedömer olika elever inom samma klass. Detta analyseras inte, av olika skäl, i rapporten.²⁸ Även eventuella systematiska skillnader i

²⁶ Skolverket 2004a kap 2.

²⁷ Skolverket 2004b. För gymnasieproven finns en motsvarande studie Skolverket 2005f och Korp 2006.

²⁸ Varför klassvariationen inte behandlas redogörs för i kap 3.4.

betygsättning mellan olika undergrupper av elever, såsom – kön, utländsk bakgrund, social bakgrund är inte behandlade i denna rapport. Dessa aspekter är naturligtvis inte ointressanta på något sätt men har av utrymmes- och resursskäl inte tagits upp. En begränsning är också att proven i svenska som andraspråk inte kommer att beröras i rapporten.²⁹

1.7 Datamaterial

Data för slutbetyg i grundskolan gäller totaldata för åren 1998-2006. För provresultat från ämnesproven årskurs 9 kommer data från ett urval av skolor för åren 1998-2002 och för de senaste fyra åren, 2003 – 2006, finns i princip totaldata, dvs. uppgifter från samtliga elever som gick ut grundskolan de åren. I praktiken förekommer dock ett visst bortfall när det gäller resultat från ämnesproven.

Samtliga analyser i rapporten bygger endast på de elever där uppgift om både provbetyg och slutbetyg finns tillgängliga. Dessutom har skolor med uppgifter från mindre än 15 elever rensats bort ur datasetet. För 1998 baseras analyserna därför på cirka 23 500 elever från cirka 280 skolor och för åren 1999-2002 på cirka 8 000 – 9 000 elever från cirka 100-120 skolor. För år 2003 finns som tidigare nämnts totaldata men bortfallet i provresultaten var ovanligt stort och analyserna gäller därför endast cirka 60 000 – 80 000 elever beroende på ämne. För 2004-2006 baseras analyserna på cirka 96 000 – 112 000 elever beroende på ämne och år.

Provinsamlingarna omfattas av ett bortfall som till en del beror på att skolor inte skickat in uppgifter, men också på att elever helt eller delvis inte genomfört proven. Våren 2006 saknade cirka 10 procent av eleverna i årskurs 9 ett sammanvägt provbetyg i svenska respektive i engelska, och 13 procent saknade provbetyg i matematik. Bortfallet har minskat över tid.

Det resultatmått som genomgående används från de nationella proven är provbetyget, som är ett sammanfattande betyg utifrån samtliga delprov. För åren 2002, 2003 och 2004 gavs dock inget provbetyg för svenska och engelska utan endast delprovsbetyg. För att möjliggöra analyserna i denna rapport har därför ett sammanvägt provbetyg beräknats i efterhand utifrån uppgifter om delprovsbetygen.³⁰

Slutbetygen från varje år baseras på det betyg som sattes i början av juni på vårterminen i årskurs 9. Därmed ingår alltså inte betyg som satts efter eventuella ”sommarkurser” och prövningar i slutet av sommaren. Sommarkurser och efterföljande prövningar skulle annars kunna vara en förklaring till avvikelser mellan prov och slutbetyg.

1.8 Nettoavvikelser – ett centralt begrepp

Den mest centrala jämförelsen i rapporten gäller den mellan hur skolors slutbetyg i svenska, matematik och engelska förhåller sig till de nationella provbetygen. I rapporten används begreppet *nettoavvikelse* i många sammanhang. På en enskild skola kan det finnas elever som får högre slutbetyg eller lägre slutbetyg än vad de fått på

²⁹ 2006 deltog 5 751 elever i nationellt prov i svenska som andraspråk jämfört med 120 837 elever i nationellt prov i svenska. Avgränsningen görs enbart av tidsskäl. Det vore även intressant att göra analyser av det nationella provet i svenska som andraspråk.

³⁰ Analyser av 2005 data visar att det sammanvägda provbetyget har hög överensstämmelse med det provbetyg som lärarna satt.

det nationella provet. Det är helt naturligt eftersom det nationella provet inte är tänkt att vara ett examensprov som helt styr slutbetyget. Eleven har givetvis visat sitt kunnande även i andra sammanhang än på det nationella provet vilket kan ge läraren skäl att sätta ett annat slutbetyg än eleven fick på det nationella provet. En elev kan också ha råkat ha en dålig eller bra dag vid provtillfällena eller ha ovanligt lätt eller svårt att hantera provsituationer.

Nettoavvikelsen är den genomsnittliga avvikelsen mellan elevernas slutbetyg och provbetyg i ett givet ämne. På en skola kan t ex 40 procent av eleverna ha fått ett högre betyg än provbetyget. 50 procent kan ha fått samma betyg som på provet och 10 procent kan ha fått ett lägre betyg. Nettoavvikelsen blir då 30 procent, dvs. 40 procent ”höjda” betyg minus 10 procent ”sänkta” betyg. Fördelen med detta mått är att det kan ge en generell bild av hur mycket slutbetygen avviker från provbetygen. Men det är viktigt att vara medveten om att måttet *nettoavvikelse* inte kan ge information om alla frågeställningar. Det visar t ex *inte* hur nära lärarna sätter betyg i förhållande till provbetygen på individnivå, dvs. hur mycket läraren låter resultaten från det nationella provet avgöra den enskilde elevens slutbetyg. För att ta ett extremt exempel så är en skolas nettoavvikelse noll om 50 procent av eleverna får ett höjt betyg och 50 procent ett sänkt betyg. Men en skolas nettoavvikelse är också noll om alla elever får exakt samma slutbetyg som provbetyg. En nettoavvikelse på noll innebär således inte *nödvändigtvis* att skolan använder proven som examensprov även om så skulle kunna vara fallet.

Nettoavvikelsen på skolnivå ger därmed heller inget mått på variationen mellan olika klasser inom en skola. På en skola med två klasser som betygsätts av två olika lärare kan det i teorin vara så att den ena läraren ger alla elever ett höjt betyg och den andre läraren ger alla elever ett sänkt betyg i förhållande till proven. Enligt den modell som använts här skulle en sådan skola få nettoavvikelsen noll. Nettoavvikelsen ger enbart en generell bild av om en skola har slutbetyg som ligger högt eller lågt i förhållande till de nationella proven.

1.9 Disposition

Innan de statistiska analyserna presenteras ska i den resterande delen av detta kapitel några av de tänkbara förklaringarna tas upp till varför resultaten på de nationella proven kan tänkas avvika från slutbetygen i svenska, matematik och engelska. I rapporten tas endast upp förklaringar som har att göra med avvikelser mellan klasser, skolor, kommuner och på en total (nationell) nivå.

I kapitel 2 presenteras en generell deskriptiv bild av relationerna provbetyg-slutbetyg på nationell nivå och hur utvecklingen ser ut över tid. *I kapitel 3* visas hur mycket proven styr på elevnivå. Frågor som besvaras är t ex hur stor andel elever som inte får godkänt provbetyg respektive slutbetyg och hur vanligt det är att elever får samma, lägre respektive högre slutbetyg i förhållande till provbetyg på nationell nivå.

I kapitel 4 flyttas fokus från den nationella nivån till den lokala nivån. Det som framförallt undersöks är hur relationen provbetyg – slutbetyg varierar, dels mellan skolor och dels mellan kommuner. Hur stora är dessa skillnader? En viktig analys i kapitlet är hur variationen mellan skolor eventuellt har förändrats över tid. Ytterligare frågor som besvaras är om det är samma skolor/kommuner som avviker på ett likartat sätt från år till år samt om det finns tendenser till överensstämmelse mellan hur de tre ämnena avviker på en given skola. För att pröva om de observerade

nettoavvikelserna också står för faktiska kunskapsskillnader mellan olika skolor har även mer oberoende kunskapsmått använts. Elevers gymnasiebetyg samt resultat från den internationella undersökningen PISA har använts för att pröva i vilken utsträckning skillnader i skolors betygsättning i förhållande till provbetygen verkligen indikerar skillnader i likvärdighet på systemnivå.

I kapitel 5 undersöks mer i detalj vilka elever som får höjda betyg i förhållande till provbetygen. En analys görs också av på vilket sätt betygsättningen skiljer sig mellan de skolor som avviker mest respektive minst i fråga om relationen provbetyg – slutbetyg.

I kapitel 6 undersöks och testas till viss del några förklaringar som har förekommit i skoldebatten när det gäller avvikelser mellan skolors prov- och slutbetyg. Den första gäller huruvida ökad konkurrens bland skolor har lett till betygsinflation och större skillnader mellan skolor. Den andra förklaringen som undersöks är om betygsättningen i förhållande till de nationella proven skulle vara högre i fristående skolor. *I kapitel 7* undersöks om det är svårare att få ett visst betyg om man går på en skola där den generella prestationsnivån är relativt hög.

I kapitel 8 slutligen sammanfattas resultaten. Det förs en diskussion om vilka förklaringar till resultaten som kan vara mest rimliga och det diskuteras vilka slutsatser som kan dras utifrån rapportens resultat.

1.10 Förklaringar till avvikelser mellan provbetyg och slutbetyg på skolnivå

Innan de statistiska analyserna presenteras ska några av de tänkbara förklaringarna till varför resultaten på de nationella proven kan tänkas avvika från slutbetygen i svenska, matematik och engelska tas upp. Som redan nämnts handlar det här enbart om förklaringar som kan förklara skillnader på aggregerad nivå: t ex skillnader mellan olika skolor, men inte mellan enskilda individer. Syftet med detta är att ge en tolkningsram för de statistiska analyserna. Förklaringarna har utvecklats utifrån en liknande sammanställning i den tidigare skolverksrapporten *Det nationella provsystemet i den målstyrda skolan*³¹. Det görs ännu ingen värdering av hållbarheten och styrkan i förklaringarna. Först i det avslutande kapitel 8 diskuteras rimligheten i de olika argumenten. I allmänhet går argumenten ut på att förklara varför det kan finnas *variationer* mellan olika skolor när det gäller avvikelser mellan skolors genomsnittliga provbetyg och slutbetyg. Ofta gäller förklaringar varför slutbetygen är högre än provbetygen för en skola men det kan även gälla förklaringar till lägre slutbetyg än provbetyg. .

Förklaringarna har delats upp i tre kategorier, dels utifrån om de ligger inom ramen för betygsbestämmelserna och dels om förklaringarna påverkar likvärdigheten i betygsättningen.

Kategori 1: Förklaringar som ligger inom ramen för betygsbestämmelserna och som inte påverkar likvärdigheten i betygsättningen.

(a) Läraren har att ta hänsyn till fler mål än de som ingår i nationella proven.

De nationella proven prövar inte alla mål i kursplanen och vissa mål provas enbart i begränsad utsträckning. Om eleverna är duktigare eller sämre på de mål som *inte* provas, eller mål som provas i begränsad utsträckning, inom ramen för de nationella proven så borde slutbetygen bli högre eller lägre än provbetygen. Det går också att tänka sig att olika lärare och skolor gör olika prioriteringar när det gäller om man prioriterar de mål som provas inom ramen för de nationella proven. Därför kan skillnader uppstå mellan skolor.

(b) Läraren har ett bredare underlag än provresultatet för att bedöma elevernas måluppfyllelse.

Det nationella provet är bara ett av en mängd olika bedömningsunderlag inför betygsättningen av en elev. Läraren kan i de andra bedömningsunderlagen ha sett att en elev klarade de mål som inte uppnåddes på det nationella provet. Alternativt kan läraren i andra bedömningsunderlag se att eleven normalt inte klarar en typ av uppgifter som eleven lyckats med på det nationella provet. Därför kan också slutbetyget bli högre eller lägre än vad eleven visat i det nationella provet. Lärare och skolor kan göra olika prioriteringar när det gäller hur stor vikt man lägger vid det nationella provet jämfört med andra bedömningsunderlag varför det kan uppstå skillnader mellan skolor.

(c) Särskilda insatser sätts in för elever som inte klarat proven.

De nationella proven görs på vårterminen i årskurs 9. Även om tiden är knapp finns det möjlighet att sätta in stödinsatser efter proven. För vissa elever kanske det handlar om att man bara saknat en mindre del för att nå godkänt betyg. Även elever som nått godkänt betyg på det nationella provet kan göra ansträngningar och framsteg efter provet som gör att slutbetygen blir högre än provbetygen. Lärare och

³¹ Skolverket 2004a s.70ff.

elever kan göra olika mycket insatser och ansträngningar tiden efter de nationella proven. Insatserna kan också vara olika mycket effektiva vilket kan förklara skillnader mellan skolor när det gäller avvikelser till provbetygen.

(d) Eleverna har arbetat med ett annat stoff än det provet behandlar.

Kursplanerna lämnar ett stort utrymme för lärarna att välja vilket stoff man ska arbeta med för att nå målen. Man kan tänka sig att elever på skolor som tränat på ett stoff som ligger nära det som kommer på det nationella provet når bättre resultat på provet. Det kan vara en tänkbar förklaring till avvikelser mellan skolor i förhållande till det nationella provet.

(e) Lärare planerar undervisningsmomenten olika

Förklaringen är en variant av den ovanstående. Kursplanerna reglerar inte när man ska ta upp olika moment i ett ämne. Även om de nationella proven görs på våren i årskurs 9 kan planeringen av undervisningen ha betydelse. Det går att tänka sig att det finns skolor som väntar med att undervisa om, eller kanske snarare repetera vissa moment till efter att de nationella proven är gjorda. Man kan också tänka sig att moment som är centrala i ett nationellt prov behandlas vid olika tidpunkter på olika skolor. Var det länge sedan eleverna läste ett moment som är centralt på proven så kan resultaten bli sämre. Detta skulle kunna vara en förklaring till skillnader mellan skolor när det gäller relationen prov och betyg.

Samtliga förklaringar i kategori 1 påverkar inte likvärdigheten i betygsättningen eftersom de ligger inom ramen för lärarnas friutrymme när det gäller att planera undervisningen och pröva elevernas kunskaper. De här förklaringarna finns inte något skäl för staten att försöka påverka inom ramen för nuvarande bestämmelser. Det finns inte heller något som säger att dessa förklaringar betyder att eleverna vid betygsättningen bedöms efter olika måttstockar. Har t ex eleverna genom lyckosamma stödinsatser på en skola tillägnat sig mer kunskaper efter provtillfället ska detta naturligtvis vägas in när slutbetygen sätts. I viss utsträckning kommer rapporten att pröva om stora ”höjningar” mellan proven och slutbetygen på en skola också betyder att dessa elever har bättre kunskaper (kap 3.3).

Kategori 2: Förklaringar som ligger inom ramen för betygsbestämmelserna men som påverkar likvärdigheten i betygsättningen.

(f) Lärare på olika skolor gör olika tolkningar av mål och betygs-kriterier, dels sinsemellan och dels i förhållande till de nationella proven.

Kursplanernas uppnåendemål och betygs-kriterier, som ska ligga till grund för både de nationella proven och slutbetygen, är utformade på ett sådant sätt att de måste tolkas av lärare. Detta kan i så fall leda till att vissa skolor gör en strängare tolkning av kursplanens mål och kriterier och andra skolor gör en generösare tolkning i jämförelse med den tolkning som görs i det nationella provet. Om lärarna på olika skolor gör olika tolkningar uppstår skillnader mellan skolorna när det gäller avvikelser i förhållande till slutbetygen.

Dessutom fastställs de betygsgränser som finns i bedömningshandledningarna till de nationella proven av olika universitets- och högskoleinstitutioner och referensgrupper av aktiva lärare och i samråd med Skolverket och provuppgifterna prövas ut på skolor i olika delar av landet. Det finns dock ingen garanti för att dessa

tolkningar av mål och kriterier helt överensstämmer med den samlade lärarkårens tolkningar. Om lärarkåren, som helhet, gör en generösare tolkning jämfört med provgruppen blir slutbetygen på nationell nivå högre än provbetygen. Det kan också vara så att tolkningarna skiljer sig åt mellan lärare på olika skolor varför olika skillnader i förhållande till de nationella proven kan uppstå.

(g) Lärare tolkar kursplanerna relativt elevgruppens förutsättningar

Denna förklaring är en variant och fördjupning av ovanstående förklaring (f) då den försöker förklara *varför* lärares tolkningar skiljer sig. Kursplanernas uppnåendemål och kriterier är utformade på ett sådant sätt att de förutsätter en professionell tolkning av lärarna. Detta kan i så fall leda till att vissa skolor har strängare tolkningar av kursplanens mål och kriterier och andra skolor har generösare tolkningar i jämförelse med den tolkning som görs i det nationella provet. Denna förklaring kommer delvis att kunna prövas och belysas med statistiska data i kapitel 5.2.

(h) Lärarna bedömer proven olika

Lärarna bedömer de nationella proven själva med hjälp av centralt utformade bedömningsanvisningar. Det går att tänka sig att lärare tolkar och tillämpar bedömningsanvisningarna på olika sätt. Det kan i så fall vara en förklaring till skillnader mellan skolor när det gäller relationen mellan prov och betyg.

Förklaringarna i kategori 2 handlar alla om tolkningar. Lärarna är suveräna i sin tolkning av både kursplanernas mål och betygsriterier och provens bedömningshandledningar. Även om förklaringarna ligger inom ramen för betygsbestämmelserna är detta ändå faktorer som påverkar likvärdigheten i betygsättningen. En samstämmig tolkning av mål och kriterier är eftersträfvansvärt och nödvändigt för staten att understödja.

Kategori 3: Förklaringar som inte ligger inom ramen för betygsbestämmelserna och som påverkar likvärdigheten i betygsättningen.

(i) Betygen sätts på andra grunder än provbetygen.

Det kan finnas även andra faktorer än kursplanernas mål och betygsriterierna som ligger till grund för lärares betygsättning. Till exempel kan sådant som att elevernas allmänna beteende och attityder vägs in i betygsättningen. Eller så kan vissa kursplanemål väljas bort av lärarna av olika skäl. En förklaring till en variation i nettoavvikelser mellan olika skolor kan bero på olika benägenhet att använda andra kriterier än de som anges i kursplanen

(j) Behörighetskraven gör att lärare inte gärna vill underkänna elever

Får en elev inte godkänt betyg på ett av de behörighetsgivande ämnena kommer eleven inte in på gymnasiets nationella- eller specialutformade program utan hänvisas till det individuella programmet. Lärare som står inför betygsättning av elever som ligger på gränsen till godkänt befinner sig därför under stor press från elever och deras föräldrar. Lärare är medvetna om att även rektor och kommunledning önskar så få elever som möjligt som inte blir godkända. Andelen elever som är behöriga till gymnasiet är ett viktigt mått i skolornas och kommunernas kvalitetsredovisningar. Läraren kan även se det som att andelen godkända elever är ett mått på den egna arbetsinsatsen. Det är därför förståeligt, om än inte försvarbart, om lärarna ger vika

och sätter godkänt betyg även i fall där läraren bedömer att en elev inte når alla uppnåendemål. Det kan vara en förklaring till att slutbetygen är högre än provbetygen. Det går också att tänka sig att benägenheten att godkänna elever som egentligen inte nått alla mål skiljer sig åt mellan skolor.

(k) Konkurrensen ger betygsinflation och variation mellan skolor.

Man kan tänka sig att den ökande konkurrensen mellan skolor skulle kunna leda till att betygen fungerar som ett konkurrensmedel. En skola med höga betyg drar till sig nya elever och därmed nya pengar och skapar en positiv spiral. Frestelsen att medvetet sätta ”för höga” betyg – dvs. öka den positiva avvikelsen gentemot ämnesproven – kan vara stor. Ett ökat användande av betygsantagning till gymnasieskolorna kan också skapa ett tryck från eleverna att få höga betyg.³² Detta skulle kunna ge en generell effekt, en ”betygsinflation”, så att den totala avvikelsen mellan prov och betyg ökar.³³ Det skulle också kunna leda till ökade skillnader mellan skolor beroende på hur pass konkurrensutsatt en skola är och i vilken utsträckning en skola ”ger efter” för konkurrensstrycket när det gäller betygsättningen. Denna förklaring kommer delvis att kunna prövas empiriskt med statistiska data i rapporten i kapitel 5.1.

Förklaringarna i kategori 3 innebär alla att lärarna betygsätter eleverna på andra grunder än vad som ska göras enligt betygsbestämmelserna. Det är självklart en angelägenhet för staten att försöka vidta åtgärder för att motverka dessa olika förklaringar.

Ingen av de olika förklaringarna i de tre kategorierna till skillnaderna mellan slutbetyg och provbetyg torde ensam vara tillräcklig. Med de statistiska analyser som görs i denna rapport går det inte att på något säkert sätt avgöra vilka förklaringar som är mest trovärdiga eller ger starkast utslag. I slutkapitlet kommer ändå en diskussion att föras om rimligheten i de olika förklaringarna.

³² Söderström 2006 s.18.

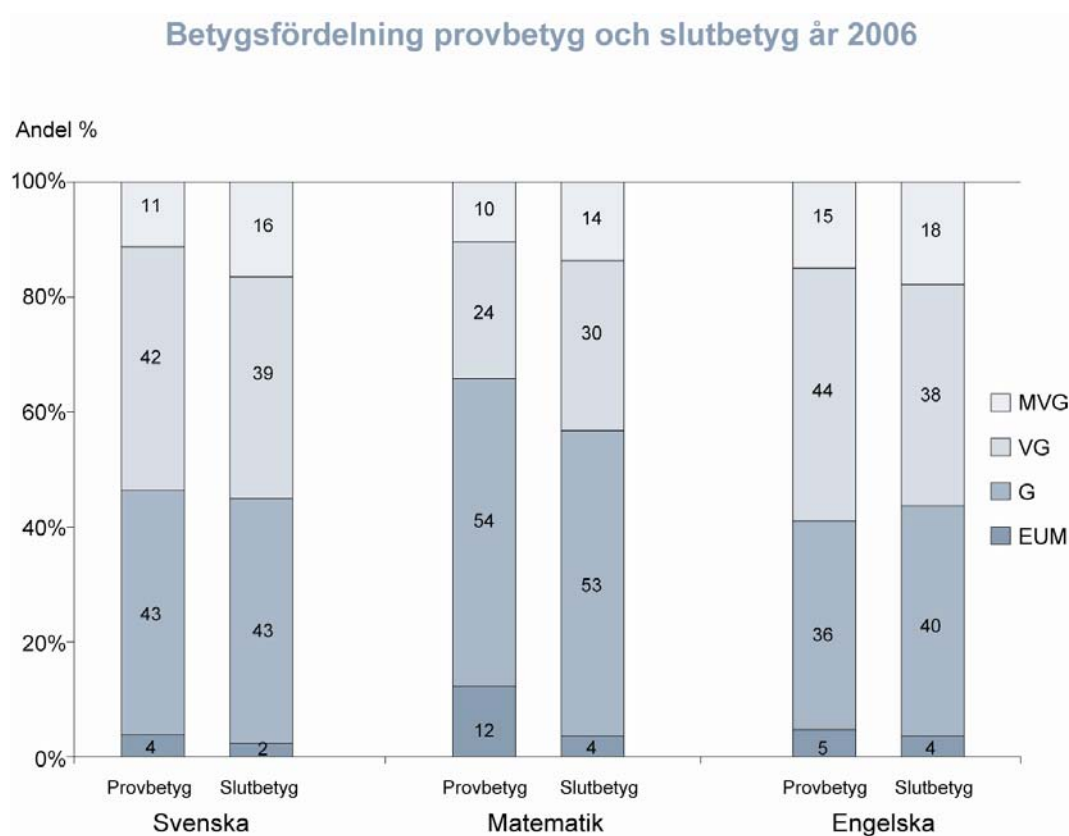
³³ Frågan om betygsinflation har framförallt uppmärksamats när det gäller gymnasieskolan. Se Söderström 2006, Wikström 2005, Cliffordsson 2004, Skolverket 2004c.

Kapitel 2. Relationen provbetyg – slutbetyg i ett nationellt perspektiv

I föregående kapitel diskuterades utifrån teoretiska grunder hur provbetyg och slutbetyg kan tänkas förhålla sig till varandra, både på individnivå och på aggregerad nivå. Ett antal tänkbara förklaringar presenterades om varför det kan förekomma avvikelser mellan provbetyg och slutbetyg även på aggregerad nivå, dvs. skolnivå, kommunnivå och nationell nivå. I detta kapitel presenteras en empirisk bild av hur relationen mellan provbetyg och slutbetyg ser ut på nationell nivå. I kapitel 3 ges en motsvarande bild gällande relationen mellan prov och betyg på skol- och kommunnivå.

2.1 Nationell betygsfördelning för provbetyg respektive slutbetyg.

I figur 2.1 nedan presenteras betygsfördelningen för provbetyg respektive slutbetyg för vart och ett av de tre ämnena svenska, matematik och engelska för år 2006.



Figur 2.1 Fördelning av provbetyg och slutbetyg efter ämne, år 2006.³⁴

³⁴ Observera att dessa fördelningar inte helt stämmer överens med Skolverkets officiella statistiktabeller för t ex andelen elever som inte uppnått målen i slutbetyg. http://www.skolverket.se/content/1/c4/74/88/Tabell1_4.xls Anledningen är att i Skolverkets officiella statistik ingår **alla** elever som det finns uppgift om slutbetyg för, även de elever som det inte finns uppgift om provbetyg för. Eftersom det finns ett visst bortfall i provbetygen och eftersom

Figur 2.1 visar att andelen elever som inte uppnått målen (EUM) i det nationella provet är högre än motsvarande andel elever som inte uppnår målen enligt slutbetyget. Det gäller alla ämnen. I matematik är denna skillnad som störst med hela 12 procent som inte uppnår målen i provbetyget mot endast 4 procent som inte uppnår målen enligt slutbetyget. I engelska är skillnaden minst, cirka en procentenhet. Om man tittar i andra änden av fördelningen så är förhållandet det motsatta, dvs. andelen som får MVG enligt slutbetyget är klart högre än andelen som får MVG i provbetyget. I svenska erhåller 11 procent av eleverna MVG i provbetyg medan 16 procent har fått MVG i slutbetyg, en skillnad på 5 procentenheter. I matematik och engelska är skillnaderna något mindre, 4 respektive 3 procentenheter.

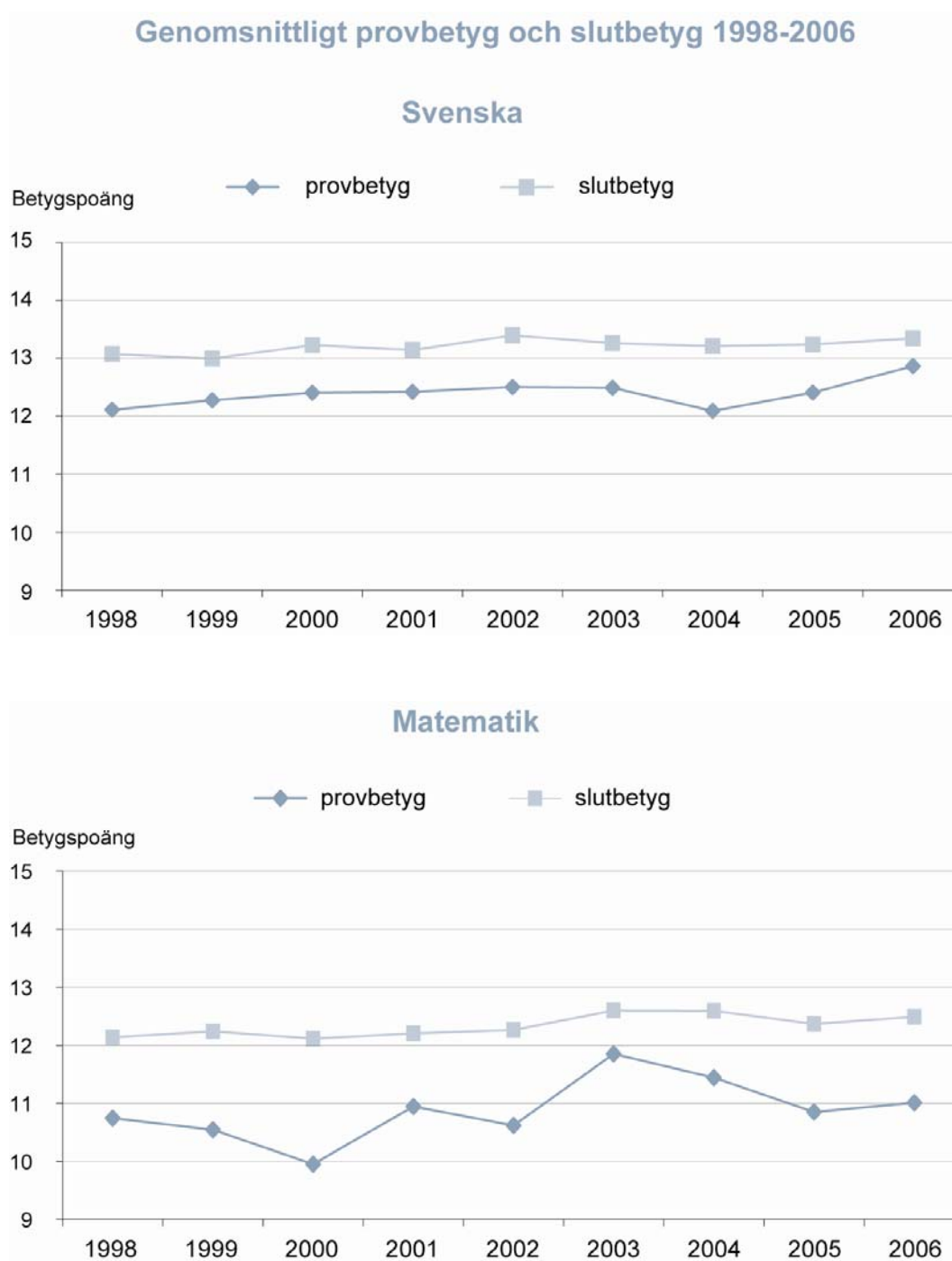
För att få en mer övergripande bild av hur provbetyg och slutbetyg förhåller sig till varandra på nationell nivå kan man beräkna en genomsnittlig betygspoäng för provbetyg respektive slutbetyg³⁵. I figur 2.2 a-c presenteras den genomsnittliga betygspoängen för prov- respektive slutbetyg för varje ämne och för åren 1998-2006, dvs. så långt tillbaka i tiden som ämnesproven har funnits.

För svenska visar figur 2.2a att år 2006 var den genomsnittliga betygspoängen för provbetyget cirka 12,9 poäng medan den genomsnittliga betygspoängen för slutbetyget var cirka 13,3 poäng. Dessa två siffror sammanfattar därmed den genomsnittliga skillnad i betygsfördelning som kunde ses i figur 2.1 för svenska.

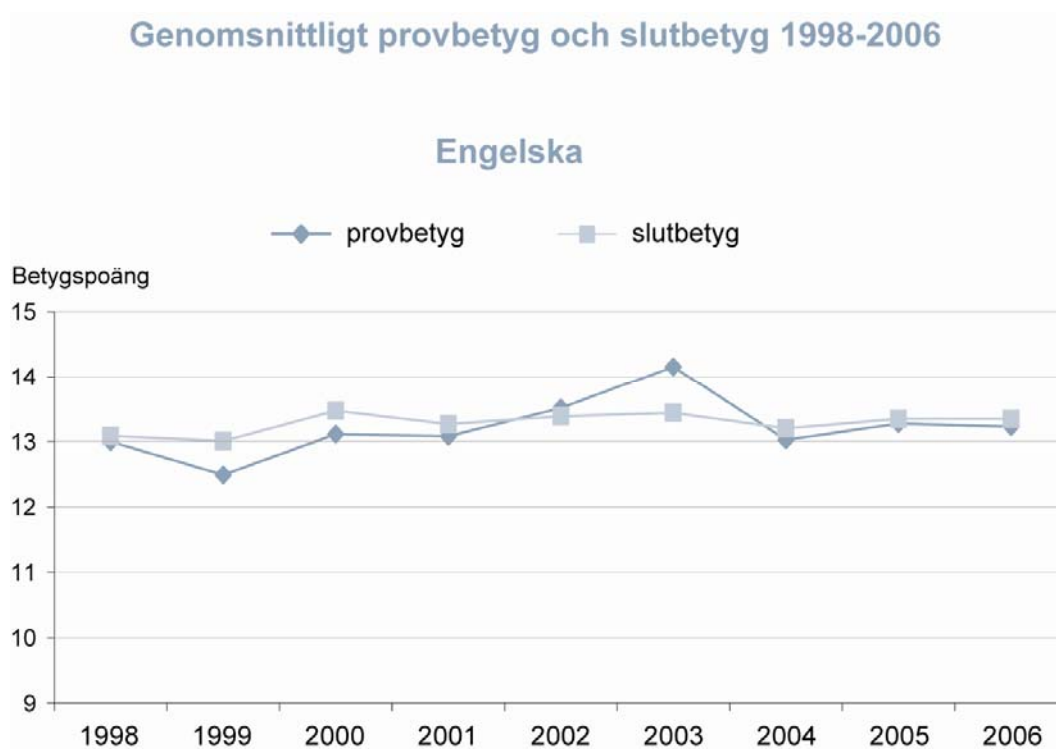
En viktig iakttagelse är att den genomsnittliga betygspoängen för slutbetyget konsekvent ligger över motsvarande betygspoäng för provbetyget i både svenska och matematik för hela tidsperioden. I engelska ligger provbetyg och slutbetyg närmare varandra och för 2003 var till och med det genomsnittliga provbetyget i engelska högre än det genomsnittliga slutbetyget. Men varför avviker då slutbetyget från provbetyget systematiskt i svenska och matematik på den nationella nivån?

”svaga” elever är överrepresenterade bland dem som det inte finns uppgift om provbetyg för, kommer andelen elever med EUM i slutbetyg antagligen att vara lägre enligt figur 2.1 (där dessa elever inte finns med) jämfört med Skolverkets officiella statistiktabeller. Men för att meningsfulla analyser skall kunna göras av relationen provbetyg – slutbetyg är det viktigt att jämförelser görs med exakt samma elevpopulation. Därför måste elever där det inte finns uppgift om både provresultat och slutbetyg exkluderas i analyserna i denna rapport.

³⁵ Genomsnittlig betygspoäng för provbetyg respektive slutbetyg beräknas enligt modellen: Ej uppnått målen = 0 poäng, G = 10 p, VG = 15 p och MVG = 20 p.



Figur 2.2 a-b. Genomsnittlig betygspoäng för provbetyg respektive slutbetyg per ämne 1998-2006.



Figur 2.2c. Genomsnittlig betygspoäng för provbetyg respektive slutbetyg, engelska 1998-2006.

Under antagandet att de nationella proven och betygen på det stora hela mäter samma kunskaper och utgår från samma betygskriterier, skulle man ur ett rent statistiskt perspektiv kunna förvänta sig att den genomsnittliga betygspoängen för provbetyget bör vara på ungefär samma nivå som den genomsnittliga betygspoängen för slutbetyget. Däremot kan man inte förvänta sig att också spridningen för provbetygen respektive slutbetygen är lika stora. Rent statistiskt är det helt logiskt att spridningen för provbetygen är större jämfört med motsvarande spridning för slutbetygen. Anledningen är att provbetygen bygger på ett enskilt mättilfälle³⁶ och därmed genererar större osäkerhet (mindre reliabilitet) jämfört med slutbetygen som ur ett rent statistiskt perspektiv skulle kunna ses som någon form av medelvärde utifrån en rad olika mätningar på elevernas förmåga som läraren har tillgång till.³⁷ En

³⁶ Även om de nationella proven består av tre delprov.

³⁷ Som exempel på detta kan man tänka sig en pilkastningstävling där först alla deltagare får kasta endast en pil var. Några kommer att träffa mitt i prick (=10 poäng) och några kommer att missa tavlan helt och hållet (=0 poäng). Men det stora flertalet kommer sannolikt att få resultat någonstans mitt emellan. Om nu tävlingen istället gäller det genomsnittliga resultatet av tre pilar per deltagare kommer fördelningen att se annorlunda ut. Till exempel kan det tänkas att inte lika många lyckas träffa mitt i prick med alla tre pilarna (= 10 poäng i genomsnittsresultat). Dessutom är det troligt att inte lika många missar tavlan helt och hållet med alla tre pilarna (= 0 poäng i genomsnittsresultat). Därmed kommer resultaten från tävlingen som bygger på ett genomsnitt från tre pilkast att vara mer samlade kring mitten av medelvärdet för alla deltagare jämfört med tävlingen som byggde på resultatet från ett enda pilkast. Det blir helt enkelt fler resultat ute svansarna av fördelningen (fler extrema resultat) när resultaten bygger på en enda mätning per individ jämfört med om slutresultatet bygger på ett genomsnitt av flera mätningar per individ. Man skulle kunna säga att tur och otur tar ut varandra till

större förväntad spridning för provbetyget innebär att en större andel av eleverna kan förväntas få det lägsta provbetyget EUM men också att en större andel av eleverna kan förväntas få det högsta provbetyget MVG, jämfört med slutförbetyget. Enligt figur 2.1 stämmer detta för andelen EUM, som mycket riktigt också är högre för provbetyget jämfört med slutförbetyget. Däremot stämmer det inte för andelen MVG som istället är *mindre* för provbetyget jämfört med slutförbetyget snarare än större som det statistiska perspektivet förutsäger. Det måste således finnas ytterligare förklaringar till de avvikelser på nationell nivå som figur 2.1 och 2.2 visar.

En ytterligare observation är att det genomsnittliga provbetyget i ett givet ämne varierar ganska mycket mellan olika år, i vart fall betydligt mer än det genomsnittliga slutförbetyget gör. Detta gäller framförallt i matematik och engelska men inte lika mycket i svenska. Om proven gav en perfekt indikation på den nationella kunskapsnivån skulle denna observerade variation innebära att kunskapsnivån mellan olika årskullar varierar ganska påtagligt, något som inte förefaller troligt. Mer troligt är att proven inte är lika lätta eller svåra från år till år, eller att betygsgränserna varierar lite från år till år. Sett ur detta perspektiv skulle man kunna tolka det som att lärarna till tar hänsyn till om det nationella provet var ovanligt svårt eller lätt när de sätter sina slutförbetyg. Detta kan dock knappast förklara varför slutförbetygen systematiskt, år efter år, skulle vara högre än provbetygen i matematik och svenska. Att provens kravnivåer verkar variera mellan åren är däremot naturligtvis inte oproblematiskt för provens betygsstödjande syfte.³⁸

2.2 Sammanfattning

För att sammanfatta kapitel 2 kan sägas att det finns avvikelser på nationell nivå mellan provbetyg och slutförbetyg och att dessa avvikelser är som störst i matematik och minst i engelska.

En ytterligare viktig observation är att provbetygen på nationell nivå uppvisar större variation från år till år jämfört med slutförbetygen och att detta sannolikt indikerar att bedömningskriterier och poänggränser för olika provbetyg varierar något från år till år. Detta är problematiskt för provens betygsstödjande funktion.

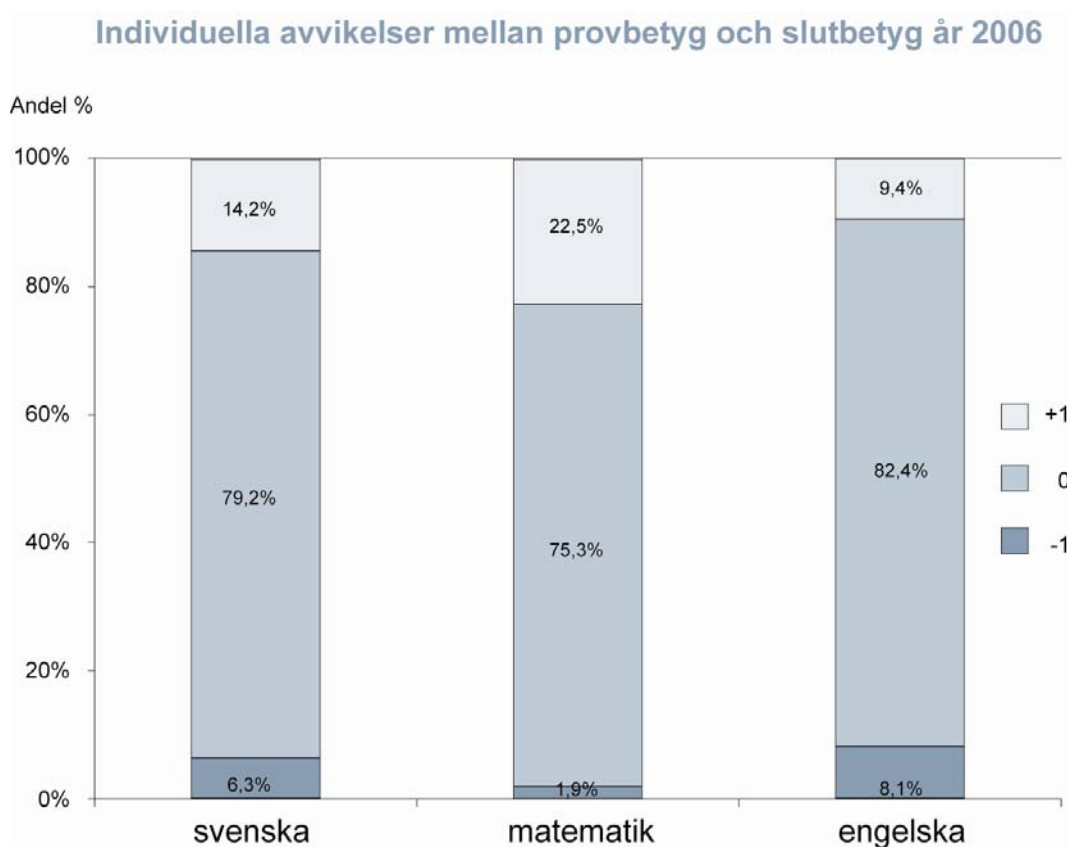
viss del när man får fler chanser att visa vad man går för och därför minskar andelen resultat som till viss del berodde på ovanligt mycket tur eller ovanligt mycket otur.

³⁸ Som nämndes ovan så är det helt naturligt att variationen i provbetygen är större än variationen i slutförbetygen. Detta kan därför bidra till att även det genomsnittliga nationella provbetyget varierar lite mer än det genomsnittliga nationella slutförbetyget. Men denna statistiska effekt kan inte förklara hela den variation som syns i det genomsnittliga provbetyget på nationell nivå.

Kapitel 3. Hur mycket styr proven på elevnivå?

Om proven fungerade som examensprov skulle det innebära att det provbetyg som en elev erhåller med automatik blir lika med elevens slutbetyg. Så är proven inte tänkta att fungera. Om det skulle visa sig att det genomsnittliga provbetyget *på nationell nivå* helt skulle överensstämma med det genomsnittliga slutbetyget så behöver det inte betyda att proven styr för mycket på *individuell nivå*. För enskilda elever kan relationen mellan prov och betyg fortfarande variera genom att lika många elever får högre som lägre slutbetyg i förhållande till provbetygen.

Men det kan vara intressant att se hur vanligt det är att elever får samma, högre eller lägre slutbetyg jämfört med provbetyg. Figur 3.1 illustrerar detta för år 2006. Om en elev har samma slutbetyg som provbetyg är avvikelsen lika med noll. En avvikelse på +1 innebär att slutbetyget är ett steg högre än provbetyget och -1 innebär att elevens slutbetyg är ett steg lägre än provbetyget.



Figur 3.1 Avvikelse på individnivå mellan provbetyg och slutbetyg för svenska, engelska och matematik år 2006. Observera att andelen elever som avviker mer än ett steg (-3, -2, +2 och +3) är så liten att de inte syns i figuren.

Figur 3.1 visar att i svenska så fick 79,2 procent av alla elever samma slutbetyg som provbetyg (avvikelse = 0) medan 14,2 procent fick ett slutbetyg som är ett steg högre

än deras provbetyg. 6,3 procent av eleverna i svenska fick ett slutbetyg som är ett steg lägre än deras provbetyg. I matematik fick hela 22,5 procent ett slutbetyg som är ett steg högre än provbetyget samtidigt som bara 1,9 procent fick ett slutbetyg som är ett steg lägre än provbetyget. I engelska är andelen elever som får ett steg högre respektive ett steg lägre slutbetyg än provbetyg ungefär lika stora, 9,4 respektive 8,1 procent. Andelen elever som fick ett slutbetyg som är antingen två steg högre eller två steg lägre än provbetyget är försvinnande liten för samtliga ämnen och syns därför inte i figuren.³⁹

I jämförelse med figur 2.1 åskådliggör figur 3.1 information om hur varje enskild elev avviker eller inte avviker mellan slutbetyg och provbetyg. Figur 2.1 däremot visade bara nettoandelarna för respektive provbetyg och slutbetyg. På så sätt gav figur 2.1 information om den aggregerade avvikelserna på nationell nivå medan figur 3.1 ger information om hur relationen mellan provbetyg och slutbetyg ser ut på elevnivå. Det är således fullt möjligt att fördelningarna i figur 2.1 mellan provbetyg och slutbetyg är exakt likadana, dvs. inga skillnader på aggregerad nationell nivå medan figur 3.1 med samma underlag skulle kunna visa på betydande avvikelser för olika elever.⁴⁰

Det finns, som redan diskuterats i kapitel 1, inte några direktiv i förordningar, författningar eller allmänna råd som anger hur lärare eller skolor skall förhålla sig till relationen mellan slutbetyg och provbetyg på aggregerad nivå. Inte heller finns några angivelser för exakt hur stor betydelse en enskild elevs resultat på det nationella provet skall ha för dennes slutbetyg. Men figur 3.1 säger ändå något om hur mycket vikt lärarna ger de nationella proven på individnivå.

Om t ex alla elever får exakt samma slutbetyg som provbetyg kan det antingen tolkas som att proven styr för mycket för den enskilde eleven, dvs. att lärarna använder det nationella provet som ett examensprov, vilket som redan nämnts inte är förenligt med gällande förordningar om betygsättning och provsystem. Men rent teoretiskt skulle det också kunna tolkas som att proven är så perfekt konstruerade, både vad gäller validitet (vad de mäter) och reliabilitet (hur säkert de mäter), att de helt enkelt fångar upp elevens kunskaper på ett hundra procentigt säkert sätt. Om istället en mycket liten andel elever får samma slutbetyg som provbetyg skulle det kunna tolkas som att proven har mycket låg validitet och/eller reliabilitet, dvs. att de inte mäter de kunskaper de är avsedda att mäta och/eller att de inte mäter dem med tillräckligt stor säkerhet.

Utifrån figur 3.1 är det svårt att hävda att proven skulle styra extremt mycket på individnivå eller fungera extremt dåligt som instrument med avseende på validitet och reliabilitet. Men att mer precist uttala sig om vad som kan anses vara en "lagom" nivå är inte möjligt utifrån dagens regelsystem.

Ett sätt att summera i vilken utsträckning elever får samma slutbetyg som provbetyg är att titta på korrelationen mellan slutbetyg och provbetyg på individnivå.⁴¹ En

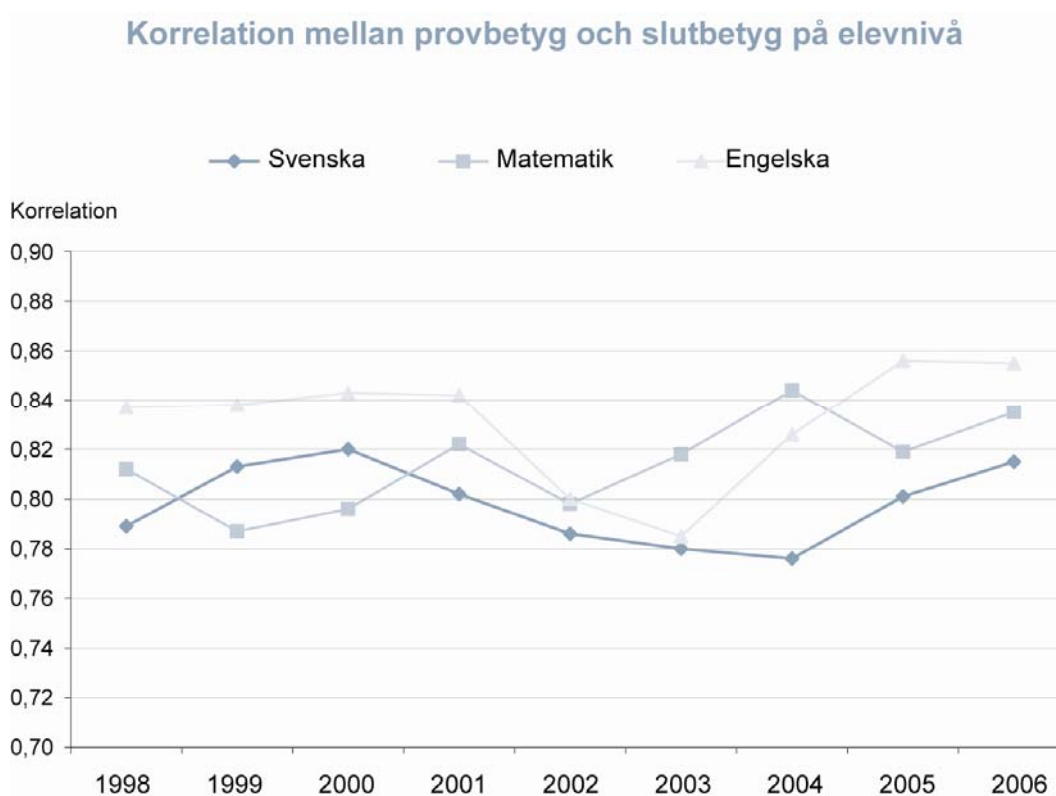
³⁹ Som exempel kan nämnas att endast 0,3 procent av eleverna i matematik får ett slutbetyg som är två steg över provbetyget.

⁴⁰ Om t ex exakt lika många elever får ett steg högre slutbetyg som elever som får ett steg lägre slutbetyg blir andelen elever med ett visst provbetyg exakt lika stor som andelen elever med ett visst slutbetyg, men de representerar inte samma elever.

⁴¹ Korrelationen mellan provbetyg och slutbetyg kan tolkas som hur mycket provbetyg och slutbetyg samvarierar.

korrelation på 1 kan tolkas som att alla elever får exakt samma slutbetyg som provbetyg och en korrelation nära noll innebär att slutbetygen sätts helt slumpmässigt i förhållande till vilket provbetyg eleven har.⁴²

I figur 3.2 åskådliggörs hur korrelationen mellan provbetyg och slutbetyg på elevnivå sett ut för varje ämne och år från 1998 till 2006.



Figur 3.2 Korrelation mellan provbetyg och slutbetyg på elevnivå 1998-2006 för svenska, matematik och engelska. Observera att skalan för den vertikala axeln inte startar på noll.

Den viktigaste informationen som figur 3.2 visar är att korrelationen på elevnivå har varit relativt stabil över tiden 1998-2006, drygt 0,80 i genomsnitt för alla ämnen för alla år, med 0,78 som den lägsta uppmätta korrelationen (svenska år 2004) och 0,86 som den högsta (engelska år 2005). Som redan nämnts kan en mycket låg korrelation på elevnivå tolkas som att proven inte mäter vad de avser att mäta.⁴³ En mycket hög korrelation på elevnivå däremot, kan tolkas som att ämnesproven fungerar som examensprov, dvs. att de har en för stor inverkan på den enskilda elevens betyg, eller annorlunda uttryckt, att elever i hög utsträckning får samma slutbetyg som provbetyg. Att bedöma vad en korrelation på t ex 0,80 innebär utifrån en balans

⁴² Det skall nämnas att en korrelation på 1 också kan fås om t ex alla elever fick exakt ett steg högre slutbetyg än provbetyg. Men även en sådan teoretiskt möjlig relation skulle kunna tolkas som att proven styr mycket eller att proven är perfekta mätinstrument, förutom att lärarna korregerar slutbetygen likformigt uppåt.

⁴³ Detta under förutsättning att slutbetygen mäter vad de avser att mäta.

mellan betygstöd och ”för mycket” individuell styrning på elevnivå är däremot inte möjlig att göra utifrån nuvarande regelverk.

Men även om nivån på korrelationen är svårtolkad så är en viktig slutsats, utifrån konstaterandet att det inte finns någon påtaglig trend i hur korrelationen på elevnivå har utvecklats över tid, att proven inte blivit mer eller mindre styrande på elevnivå över denna tidsperiod, för något ämne. Möjligen går det att spåra en svagt och relativt marginellt tilltagande trend för matematik. Sammantaget innebär detta att det *inte* har blivit vanligare att proven i större utsträckning skulle användas som examensprov. En intressant observation är dock att korrelationen i svenska och engelska är som lägst för åren 2002, 2003 och 2004, dvs. de år då inget sammanvägt provbetyg i dessa ämnen gavs.⁴⁴

Ovanstående resultat skall däremot inte tolkas som att det inte skulle finnas enskilda skolor där proven har examenskaraktär, dvs. att det betyg eleven får på det nationella provet blir det betyg eleven får i slutbetyg. För år 2005 i matematik t ex var det 13 av 1289 skolor där samtliga elever fick samma slutbetyg som provbetyg, vilket innebär att på dessa skolor var korrelationen mellan provbetyg och slutbetyg lika med 1.

3.1 Konsekvenser för likvärdighet i betygsättning

Det behöver inte vara problematiskt ur ett likvärdighetsperspektiv att lärarkollektivets tolkning av vad som krävs för ett betyg skiljer sig från den tolkning som ligger bakom provens betygsgränser så länge alla lärare och skolor gör samma tolkning. I vilken utsträckning så är fallet går inte att säga utifrån de analyser som redovisats i detta kapitel. I kapitel 3 kommer skillnader mellan hur skolor och kommuner sätter betyg i förhållande till provresultaten att analyseras. Detta kan fungera som en indikator på i vilken utsträckning lärarkollektivets tolkning av mål och betygskriterier är likartad, dvs. i vilken utsträckning bedömer alla lärare (eller skolor) efter samma ”måttstock”?

3.2 Sammanfattning

Kopplingen mellan provbetyg och slutbetyg på elevnivå har varit ganska konstant under hela tidsperioden 1998-2006 vilket kan tolkas som att proven *inte* styr mer eller mindre på individnivå idag än de gjorde när de infördes i slutet av 1990-talet. Detta motsäger dock inte att det finns enskilda skolor där de nationella proven används som examensprov, vilket det uppenbarligen gör, men det sker inte i någon större utsträckning idag än det gjort tidigare.

⁴⁴ Som redan nämnts så har provbetyg för åren 2002-2004 i svenska och engelska beräknats i efterhand centralt för att möjliggöra analyserna i denna rapport.

Kapitel 4. Relationen provbetyg – slutbetyg på lokal nivå: Skolor och kommuner

I kapitel 2 och 3 gavs en bild av hur provbetyg och slutbetyg förhåller sig på både elevnivå och nationell nivå. I detta kapitel skall redogöras för i vilken utsträckning skillnaden mellan slutbetyg och provbetyg varierar mellan olika skolor och mellan olika kommuner, dvs. hur relationen mellan det genomsnittliga provbetyget och slutbetyget ser ut mellan olika skolor och kommuner. Provbetygen används alltså som en indikator på hur mycket tolkningen av betygsriterier varierar mellan olika skolor och olika kommuner och därmed hur likvärdig betygsättningen är mellan skolor och kommuner.

4.1 Kopplingen mellan provbetyg och slutbetyg på skolnivå

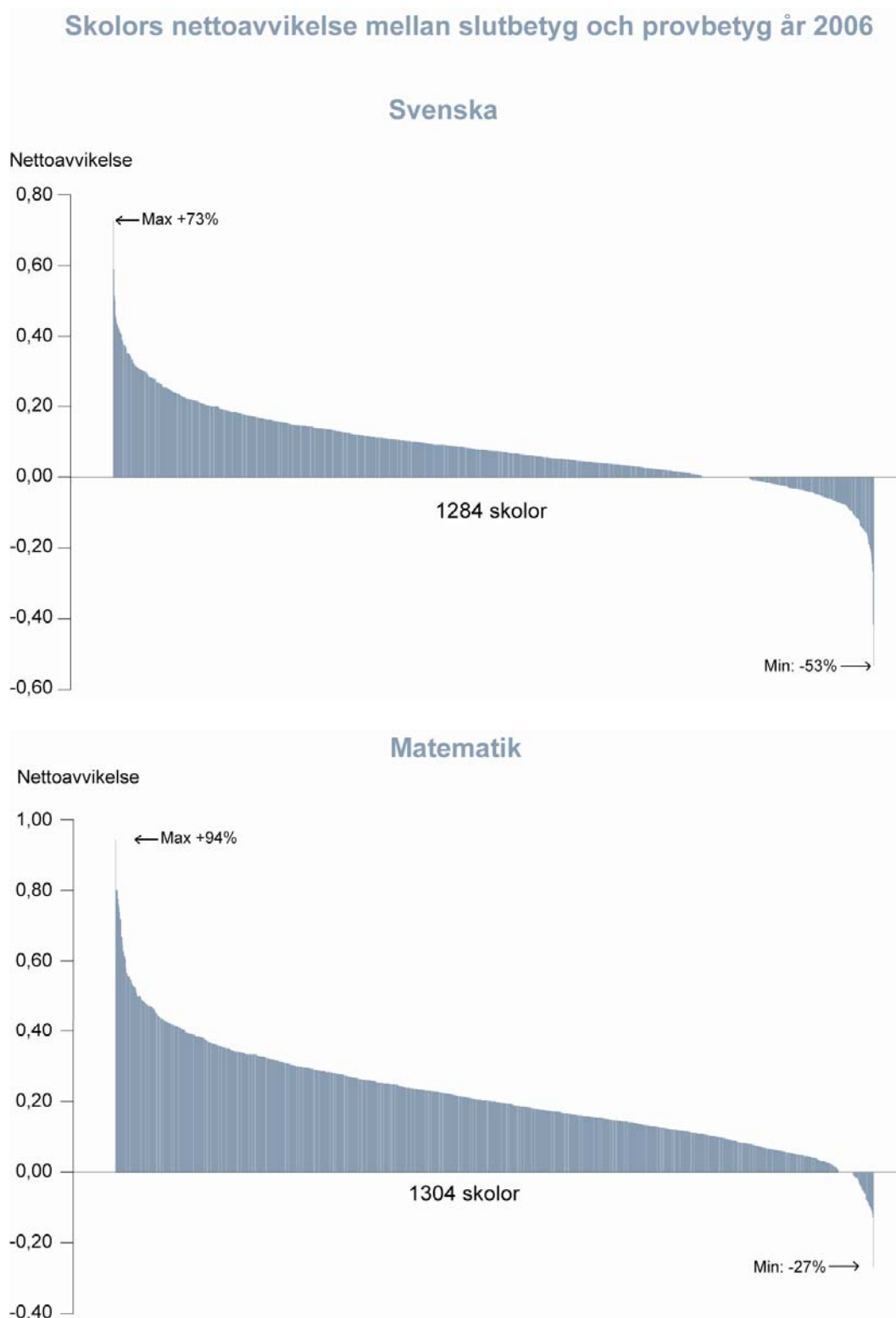
Som redan nämnts i kapitel 1 och visats empiriskt i kapitel 2 och 3 är det helt naturligt att provbetyg och slutbetyg inte överensstämmer för den enskilde eleven. Provbetyg och slutbetyg behöver med andra ord inte vara lika, framförallt beroende på att en elev vid ett enstaka mättillfälle både kan underprestera och överprestera eller prestera på sin förväntade nivå i förhållande till dennes ”sanna” eller ”egentliga” prestationsnivå. Men när en hel skolas resultat aggregeras kan man förvänta sig att sådana avvikelser upp och ner, i alla fall till viss del, tar ut varandra och följaktligen kan det förväntas att det finns en starkare koppling mellan provbetyg och slutbetyg på skolnivå än vad som är fallet på elevnivå.

Figur 4.1 a-c visar hur mycket enskilda skolor avviker i sitt genomsnittliga slutbetyg i förhållande till sitt genomsnittliga provbetyg. För varje skola har beräknats en så kallad *nettoavvikelse* mellan provbetyg och slutbetyg, vilket innebär att ett genomsnitt av alla elevers avvikelse mellan slutbetyg och provbetyg har beräknats (se kap 1.8).

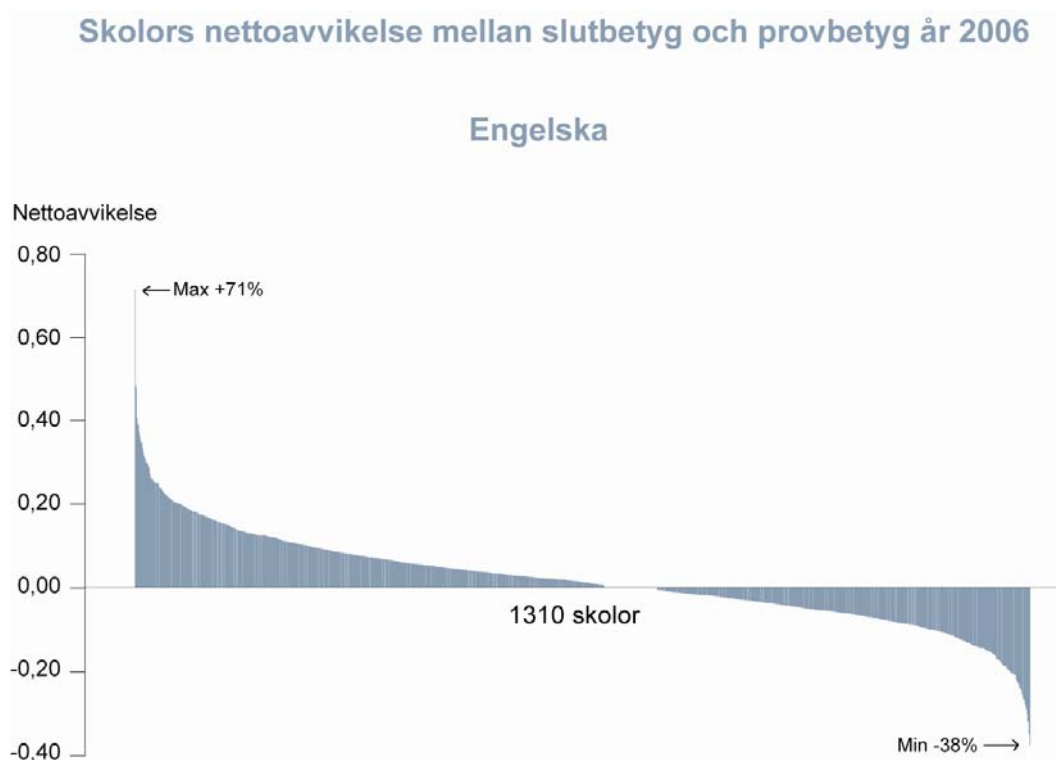
En nettoavvikelse t ex på 0,30 eller 30 procent kan tolkas som att 30 procent av alla skolans elever har fått ett slutbetyg som är ett steg högre än provbetyget medan resterande 70 procent av eleverna erhållit samma slutbetyg som provbetyg.⁴⁵ Ju högre positiv nettoavvikelse desto högre genomsnittligt slutbetyg i förhållande till provbetyg har skolan. En negativ nettoavvikelse innebär att skolan i genomsnitt har satt ett lägre slutbetyg på eleverna än vad eleverna fått i provbetyg.

I figurerna har skolorna sorterats efter hur stor nettoavvikelse de har. Den skola som har högst positiv nettoavvikelse representeras av stapeln längst till vänster medan den skola med lägst nettoavvikelse (eller störst negativ nettoavvikelse) representeras av stapeln längst till höger i figurerna.

⁴⁵ I verkligheten kan en nettoavvikelse på 30 % naturligtvis också spegla en annan fördelning, t ex att 40 % fick ett högre betyg, 50 % samma betyg och 10 % ett lägre slutbetyg jämfört med provbetyget (40 % - 10 % = 30 %). Denna kombination ger således samma nettoavvikelse. Det finns egentligen ett nästan oändligt antal kombinationer som alla ger samma genomsnittliga avvikelse (nettoavvikelse).



Figur 4.1a-b Nettoavvikelse mellan slutbetyg och provbetyg, svenska, matematik år 2006.



Figur 4.1c Nettoavvikelse mellan slutbetyg och provbetyg, engelska år 2006.

Som kan ses i figur 4.1a gällande svenska så har den skola som avviker mest en nettoavvikelse på +73 procent, vilket kan tolkas som att 73 procent av eleverna på den skolan fick ett slutbetyg som var ett steg högre än provbetyget medan övriga elever fick samma slutbetyg som provbetyg. Samtidigt har den skola med minst avvikelse en nettoavvikelse på -53 procent vilket kan tolkas som att 53 procent av eleverna på den skolan fick ett slutbetyg som var ett steg lägre än deras provbetyg medan övriga elever fick samma slutbetyg som provbetyg. Det är uppenbart att flera av skolorna skiljer sig ganska markant i nettoavvikelse men det finns också skolor där avvikelsen är noll eller mycket nära noll, vilket innebär att dessa skolor i genomsnitt sätter ett slutbetyg som är lika högt som provbetyget. Ungefär 60 procent av skolorna har en avvikelse mellan slutbetyg och provbetyg på -10 till +10 procentenheter. Den genomsnittliga nettoavvikelsen för alla skolor är för svenska år 2006 drygt 8 procentenheter.

För matematik är skillnaderna också stora mellan olika skolor (figur 4.1b). Den skola som avviker mest ”uppåt” har en nettoavvikelse på hela +94 procent medan i andra änden av spektret finns en skola med en nettoavvikelse på -27 procent. De flesta skolorna har positiva nettoavvikelser medan en minoritet har negativa avvikelser. Drygt 20 procent av skolorna har en nettoavvikelse i intervallet -10 till +10 procentenheter. I genomsnitt är avvikelsen för matematik ungefär +21 procentenheter.

För engelska (figur 4.1c) ser mönstret vid en första anblick något annorlunda ut då cirka hälften av alla skolor har positiva avvikelser och cirka hälften har negativa

avvikelser. Den genomsnittliga avvikelser är också praktiskt taget noll (0,02 procent) och cirka 72 procent av skolorna har en nettoavvikelse i intervallet -10 till $+10$ procentenheter. Men bara för att det i genomsnitt finns ungefär lika många skolor som avviker positivt som det finns skolor som avviker negativt betyder det inte att skolorna är mer lika i hur de avviker mellan prov- och slutbetyg. Skolan som avviker mest uppåt har en nettoavvikelse på $+71$ procent medan skolan som avviker mest nedåt har en nettoavvikelse på -38 procent. Detta innebär en nästan lika stor spännvidd i nettoavvikelse som för både svenska och matematik.

Spännvidden (skillnaden) mellan skolan med högst respektive lägst nettoavvikelse är dock inte det mest relevanta måttet även om denna information naturligtvis kan vara av intresse. Samtidigt är andelen skolor som ligger nära noll i nettoavvikelse inte heller nödvändigtvis ett mått på hur lika skolor är i sin betygsättning i förhållande till provbetygen.

Ett bättre mått att mäta variationen (spridningen) mellan skolors nettoavvikelse är med *standardavvikelsen* som kan tolkas som skolornas genomsnittliga avstånd till det nationella medelvärdet för nettoavvikelsen.⁴⁶ Utifrån detta mått (för år 2006) skiljer sig skolorna mest i matematik med en standardavvikelse på 14,1 procent, följt av svenska, 11,1 procent och engelska, 10,8 procent. Det är således i matematik som de största skillnaderna finns mellan hur skolor sätter slutbetyg i förhållande till provbetyg.

Olika tänkbara förklaringar till skillnaden mellan provbetygen och slutbetygen för skolorna diskuteras i kapitel 8. Här ska endast påpekas att det finns förklaringar av rent statistisk natur men även när man tar hänsyn till dessa effekter kvarstår samma mönster.⁴⁷

Samtidigt som ovanstående figurer tyder på relativt stora skillnader mellan vissa skolor så är det också värt att uppmärksamma att skillnaderna inte är lika stora mellan alla skolor. Med det menas att om man bara tittar på skolorna i mitten av fördelningen (mitten av den horisontella axeln), så är spridningen mellan dessa skolor

⁴⁶ För att förstå innebörden i detta mått ges ett exempel. Om det bara finns 2 skolor och den ena skolan har en nettoavvikelse på $+50\%$ och den andra en nettoavvikelse på $+10\%$ kommer den genomsnittliga nettoavvikelsen att vara $+30\%$ ($[50+10]/2$). Den ena skolan ligger 20 procentenheter från detta genomsnitt liksom den andra skolan. Därmed är det genomsnittliga avståndet mätt i absoluta tal till genomsnittet 20 procentenheter ($[20+20]/2$). Standardavvikelsen sägs då vara 20 procentenheter. Om alla skolor hade precis samma nettoavvikelse så hade spridningen, mätt med standardavvikelsen varit noll, dvs. skolornas genomsnittliga avstånd till den totala (nationella) genomsnittliga nettoavvikelsen är noll.

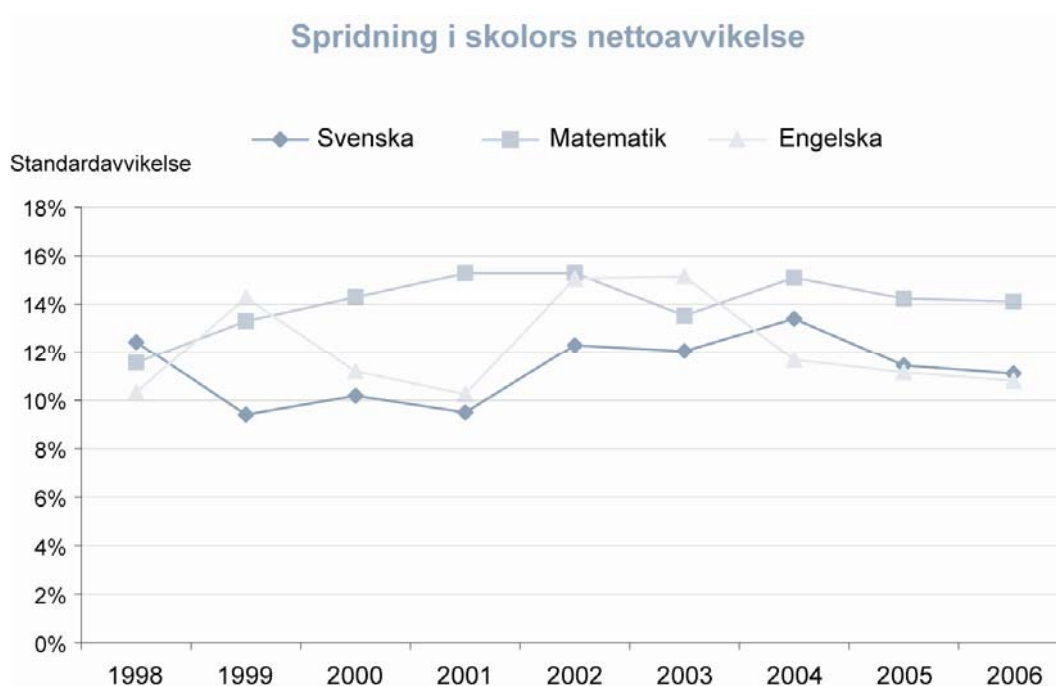
⁴⁷ En sådan statistisk förklaring är att de skolor som har störst avvikelse åt endera hållet kan vara skolor med relativt få elever. Därmed skulle de individuella avvikelserna på elevnivå inte lika lätt ta ut varandra eller annorlunda uttryckt, nettoavvikelsen för en liten skola skulle bli mer slumpmässig. Men även då endast skolor tas med i analysen som har minst 50 elever med uppgift om provbetyg och slutbetyg så är mönstret mycket likartat. En ytterligare förklaring av statistisk natur har att göra med att andelen elever med EUM i provbetyg är olika på olika skolor och att dessa elever generellt har större sannolikhet att få ett högre slutbetyg. Dessutom är elevers "sanna" underliggande prestationsförmåga inom ett givet provbetyg inte nödvändigtvis lika stor, vilket påverkar sannolikheten att få ett högre betyg. Om man försöker kontrollera för båda dessa effekter erhålls standardavvikelser i skolors nettoavvikelse på 9,6, 10,6 och 10,1 procentenheter för respektive svenska, matematik och engelska, år 2005. Variationen minskar således bara lite grann genom att ta hänsyn till dessa effekter. Se appendix och/eller teknisk rapport för mer detaljer.

naturligtvis mycket lägre.⁴⁸ Detta visar att även om det finns skillnader mellan skolor i hur de sätter betyg i förhållande till provbetyg så finns det också många skolor som förhåller sig ganska lika till provbetygen, dvs. de avviker på ett ganska likartat sätt.

Har spridningen i skolors nettoavvikelse ökat eller minskat över tid?

En ytterligare fråga är om år 2006 är ett extremt år på så sätt att skillnaderna mellan skolor i deras genomsnittliga avvikelse mellan slutbetyg och provbetyg kanske var ovanligt stora år 2006? Frågan är således hur stora skillnaderna var mellan skolor för tidigare år?

För att kunna jämföra olika år med avseende på hur mycket skolorna skiljer sig i nettoavvikelse används liksom i avsnitt 4.1 *standardavvikelsen* i skolors nettoavvikelse. I figur 4.2 visas denna spridning (*standardavvikelse*) i skolors nettoavvikelse för åren 1998-2006. Som tidigare redan nämnts så är standardavvikelsen i matematik år 2006 14,1 procent, vilket representeras av punkten överst till höger i figur 4.2. Sett över hela tidsperioden kan sägas att spridningen/standardavvikelsen genomgående är störst för matematik, inte bara år 2006 utan för de flesta åren. Figuren visar att standardavvikelsen fluktuerar lite från år till år under hela tidsperioden, beroende på ämne. Men det går inte att se någon tydlig trend för något ämne. Den viktigaste informationen som figur 4.2 ger är att de stora observerbara skillnaderna i skolors nettoavvikelse som åskådliggjordes i figur 4.1 a-c gällande år 2006 också återfinns för tidigare år, dvs. det var i stort sett lika stora skillnader i skolors nettoavvikelse i slutet av 1990-talet när proven var nya som det är nu i mitten av 2000-talet.



Figur 4.2 *Spridning i skolors nettoavvikelse, mätt som standardavvikelsen i skolors nettoavvikelse för svenska, matematik och engelska, år 1998-2006.*

⁴⁸ För matematik t ex är standardavvikelsen i nettoavvikelse för mitten 40 % skolorna endast 4 % jämfört med 14 % när alla skolor tas med.

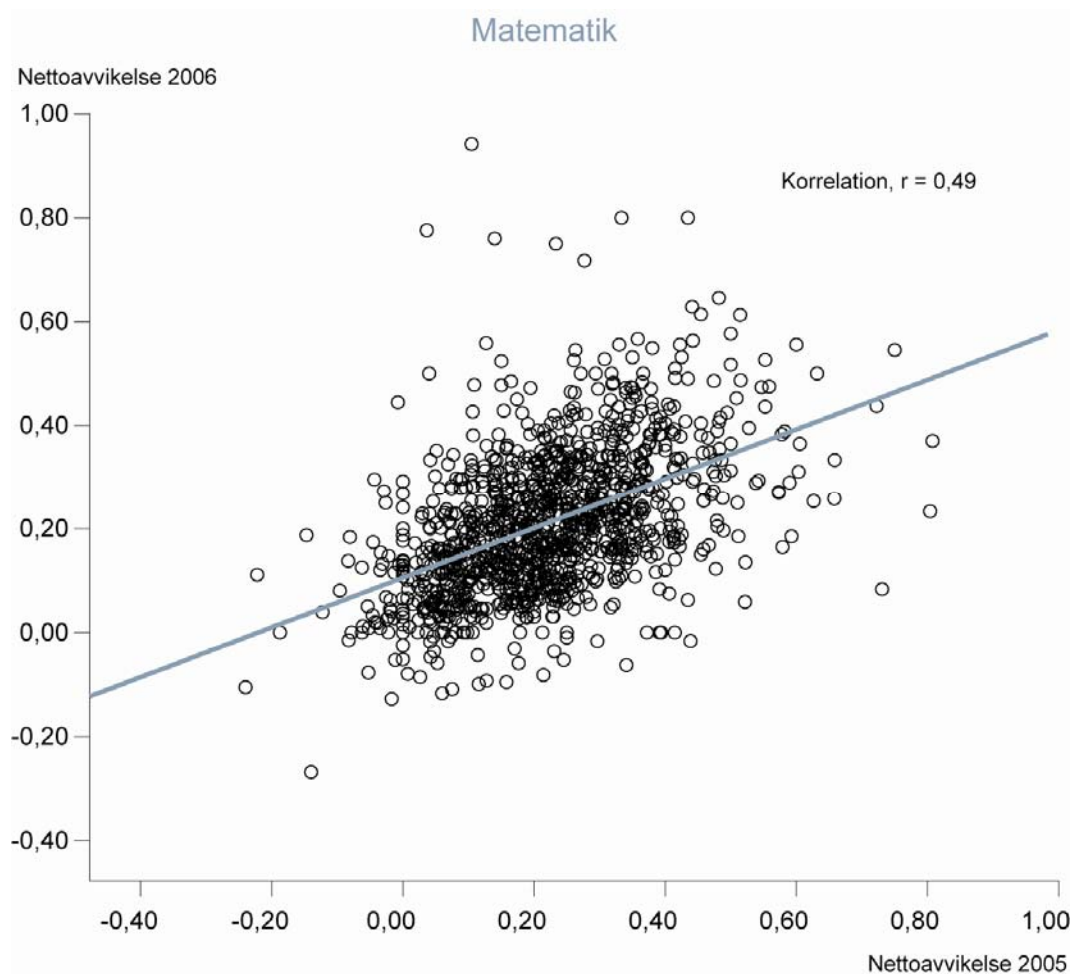
I föregående avsnitt diskuterades huruvida de observerade skillnaderna mellan skolor indikerade att det fanns skillnader i skolors sätt att sätta betyg. Även om en del av dessa skillnader beror på naturliga förklaringar så finns skäl att misstänka att en del av skillnaderna också beror på skillnader i hur betygen sätts. Men det är mycket svårt att säga vad som skall anses vara en acceptabel nivå på sådana skillnader. Det finns naturligtvis ingen magisk gräns, utifrån denna indikator (standardavvikelsen i skolors genomsnittliga avvikelse mellan slutbetyg och provbetyg) där skillnaderna mellan skolors betygsättning skall anses oacceptabelt stor. Men även om det är svårt att bedöma vad nivån på denna indikator säger om likvärdighet i betygsättning kan förändring över tid säga något om huruvida likvärdigheten har förbättrats eller försämrats över tid.

Därmed kan tidsserien som åskådliggörs i figur 4.2 tolkas som en indikator på att likvärdigheten i betygsättningen varken förbättrats eller försämrats under denna tidsperiod. Det är naturligtvis endast en av flera möjliga indikatorer på likvärdighet i betygsättning. Även om proven som indikator har brister så är den samtidigt en av få möjligheter att göra en sådan bedömning av likvärdigheten i betygsättningen över tid.

Är det samma skolor som har höga nettoavvikelser från år till år?

En annan viktig fråga i sammanhanget är om det är samma skolor som år efter år har relativt stora nettoavvikelser mellan provbetyg och slutbetyg? Om det är samma skolor som år efter år avviker mycket kanske de finns skäl att tro att det handlar om skillnader i betygskulturer skolorna emellan. Ett flertal analyser har gjorts för att undersöka detta eventuella samband för åren 2003 - 2006. Samtliga analyser tyder på att det finns ett tydligt samband mellan en skolas nettoavvikelse i ett givet ämne från ett år till ett annat.

Detta samband illustreras i figur 4.3 för matematik. Där syns skolors nettoavvikelse i matematik för 2005 på den horisontella axeln samt för 2006 på den vertikala axeln. Varje skola representeras således av en cirkel i diagrammet. Ett samband kan skönjas då skolor som tenderar att avvika mycket 2005 också tenderar att avvika mycket 2006 men det finns också många och stora avsteg från detta generella mönster, dvs. mönstret är inte deterministiskt.



Figur 4.3 Samband mellan skolors nettoavvikelse i matematik 2005 och 2006.

Korrelationskoefficienten, $r = 0,488$, statistiskt signifikant på 1 % nivå.

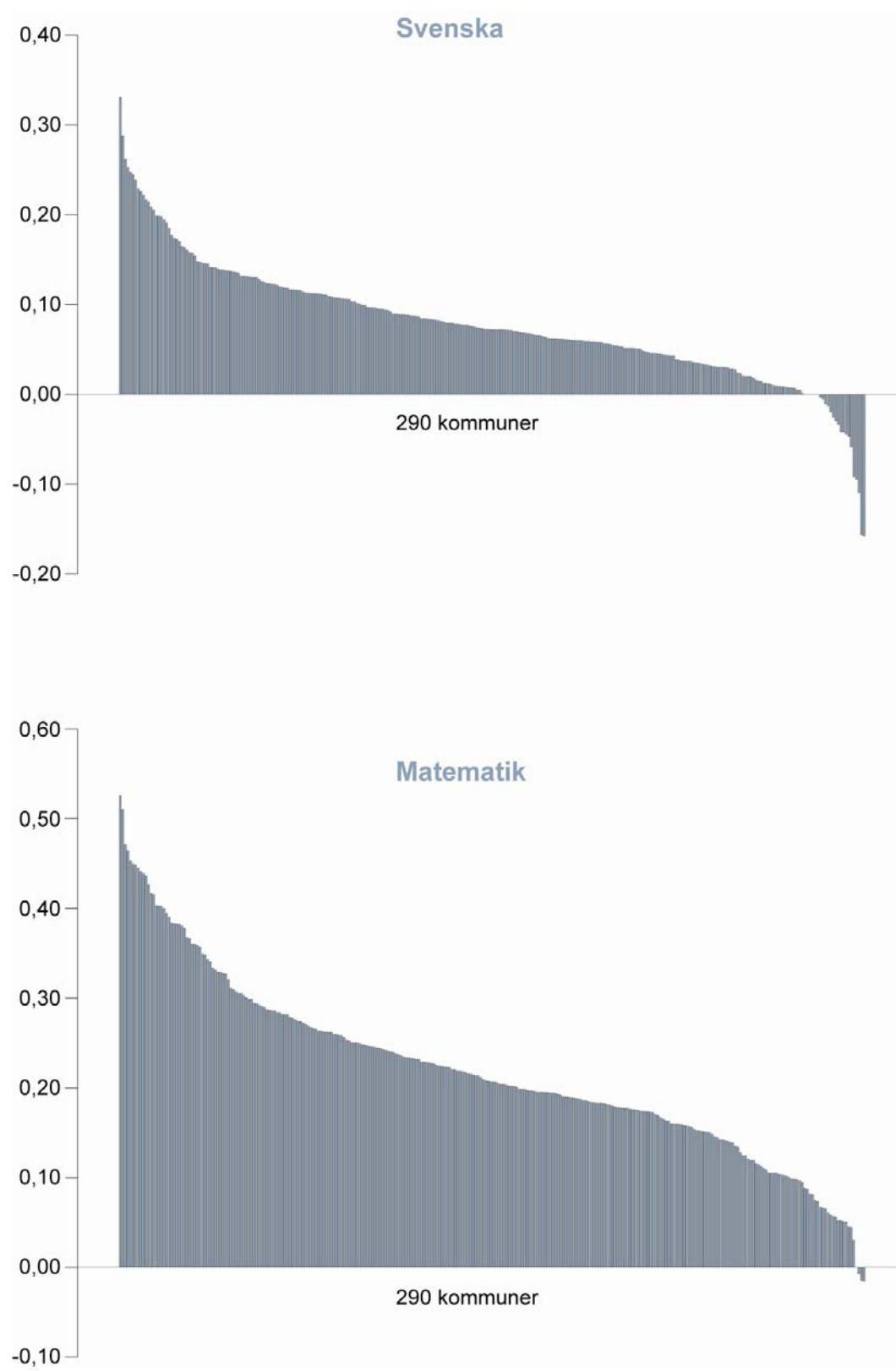
Analyserna visar även att det finns ett visst samband mellan ämnena på en skola. Detta innebär att det finns en generell tendens att skolor som avviker mycket i t ex svenska också tenderar att avvika mycket i engelska och matematik. Detta samband är dock starkare mellan engelska och svenska än mellan svenska och matematik eller engelska och matematik.

Att det finns ett statistiskt signifikant samband mellan skolors nettoavvikelse över tid och även mellan ämnen tyder på att de skillnader som finns i skolors nettoavvikelser knappast kan bero på slumpmässiga avvikelser. Däremot kan en del av de förklaringarna som togs upp i kapitel 1 fortfarande vara giltiga. Detta diskuteras vidare i kapitel 8.

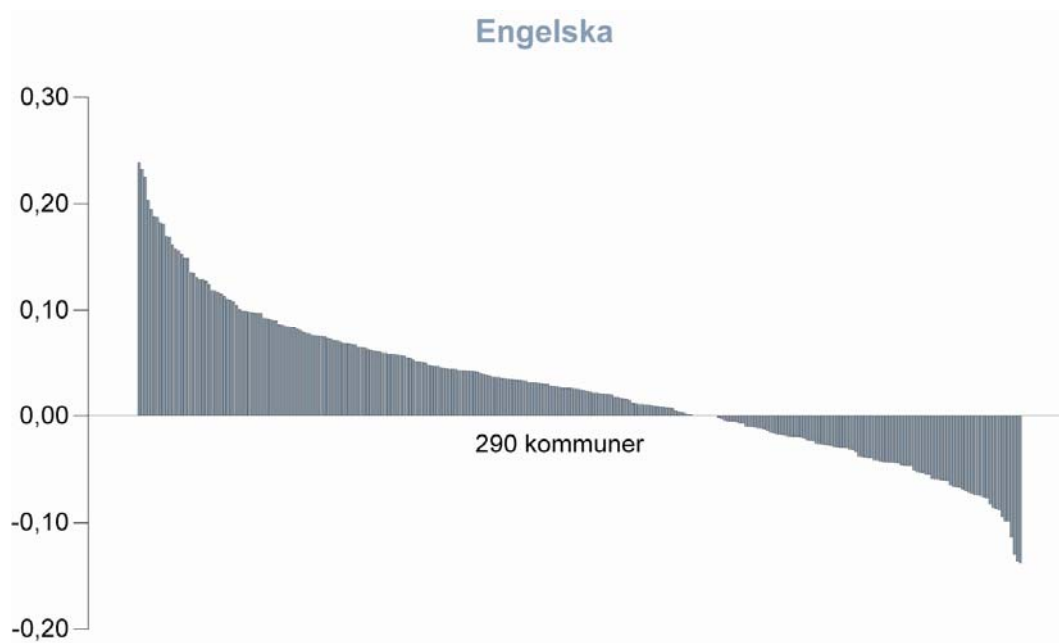
4.2 Kopplingen mellan provbetyg och slutbetyg på kommunnivå

Föregående avsnitt visade hur variationen ser ut mellan skolor vad gäller kopplingen mellan provbetyg och slutbetyg. Men hur ser bilden ut om man aggregerar provbetyg och slutbetyg till kommunnivå? Då de flesta kommuner har mer än en skola med år 9-elever så vore det rimligt att avvikelserna som syntes på skolnivå borde bli mindre, kanske till och med helt försvinna.

Figur 4.4 a-c visar kommunernas nettoavvikelse på samma sätt som skolornas nettoavvikelse åskådliggjordes i figur 4.1 a-c.



Figur 4.4a-b Nettoavvikelse mellan slutbetyg och provbetyg per kommun, svenska och matematik 2006.



Figur 4.4c Nettoavvikelse mellan slutbetyg och provbetyg per kommun, engelska 2006.

Figurerna 4.4a-c påminner mycket om motsvarande figurer för skolnivån med den skillnaden att avvikelserna inte är riktigt lika extrema. I matematik t ex har en kommun en nettoavvikelse på 53 procent, i och för sig en ganska liten kommun med rapporterade resultat för 97 elever fördelade på två skolor. Samtidigt finns det bara tre kommuner som uppvisar en negativ avvikelse, som mest -2 procent, en kommun med resultat från två skolor och 127 elever.

Variationen (mätt som standardavvikelsen) i kommuners nettoavvikelse är störst i matematik (9,9 procent) och mindre i svenska och engelska (6,7 procent vardera). Aggregeringen till kommunnivå har därmed minskat skillnaderna i nettoavvikelse något jämfört med skolnivån, från 14,1 procent till 9,9 procent i matematik, från 11,1 procent till 6,7 procent i svenska och från 10,8 till 6,7 procent i engelska. Men samtidigt är det uppenbart att det fortfarande finns relativt stora skillnader mellan olika kommuner med avseende på hur de sätter betyg i förhållande till resultaten på de nationella proven.

Även när skillnader mellan kommuners nettoavvikelse studeras för hela tidsperioden visar det sig att år 2006 inte är något extremt år. Och på samma sätt som för skolor uppvisar matematik en större spridning bland kommuner jämfört med svenska och engelska.

4.3 Speglar skillnader i skolors nettoavvikelser faktiska skillnader i kunskaper?

Analyserna i avsnitt 4.1 visar att det finns en ganska betydande skillnad i hur skolorna sätter slutbetyg i förhållande till provbetyg. Detta skulle kunna betyda att uttolkningen och tillämpningen av betygskriterierna varierar mellan olika skolor i

landet, att eleverna inte betygssätts efter samma måttstock, och att eleverna därmed får en generösare bedömning på skolor med hög nettoavvikelse och en strängare bedömning på skolor med låg eller negativ nettoavvikelse (förklaringar tillhörande kategori 2 och 3 i kapitel 1.10). Men det är också möjligt att den variation som syns i skolornas nettoavvikelse beror på att elever som går på skolor med höga nettoavvikelser faktiskt har kunskaper som motsvarar de relativt höga betygen (förklaringar tillhörande kategori 1 i kapitel 1.10). Det går inte att ge ett klart svar på den frågan med tillgängliga data men detta i avsnitt skall två analyser genomföras för att försöka få ytterligare klarhet i, i vilken utsträckning variationen i skolors nettoavvikelse verkligen indikerar att elever betygsätts efter olika kravnivåer beroende på vilka skolor de går i.

Hur klarar sig elever från grundskolor med höga nettoavvikelser senare på gymnasiet?

Om det verkligen är så att skolor som sätter relativt höga slutbetyg i förhållande till hur eleverna presterat på de nationella proven (dvs. hög nettoavvikelse), faktiskt också tolkar och tillämpar betygskriterierna på ett relativt generösare sätt för ett givet betyg, så borde dessa elever i genomsnitt också klara sig relativt sämre på gymnasiet i förhållande till elever från övriga skolor med samma slutbetyg från grundskolan.⁴⁹

Denna hypotes går att testa för den elevkohort som skrev de nationella proven år 1998 då det också finns uppgifter på elevnivå om deras slutbetyg från gymnasiet. Det som undersöks är förhållandet mellan elevernas meritvärde, genomsnittliga slutbetyg från gymnasiet samt skolans genomsnittliga nettoavvikelse utifrån de tre ämnena svenska, matematik och engelska.⁵⁰

Analysen visar på ett statistiskt signifikant negativt samband⁵¹ mellan hur elever klarar sig på gymnasiet (utifrån vad som kan förväntas utifrån deras grundskolebetyg) och hur mycket grundskolan de gick på avvek mellan slutbetyg och provbetyg. Ju högre genomsnittlig nettoavvikelse på elevers grundskola, desto sämre gick det för eleverna från den skolan på gymnasiet (relativt deras betyg). Detta stöder därmed hypotesen att skillnader i skolors nettoavvikelser också speglar skillnader i betygsättning.

Utbildningsförvaltningen i Stockholms kommun har även undersökt sambandet mellan grundskolebetyg och gymnasiebetyg⁵² och funnit systematiska skillnader mellan olika grundskolor. Utbildningsförvaltningen tolkar detta som att det finns skillnader i grundskolors betygsättning. Dessa data gäller slutbetyg från grundskolan år 2002 och analysen är inte baserad på resultat från de nationella proven utan bygger helt och hållet på ett skattat samband mellan grundskolebetyg och gymnasiebetyg.

⁴⁹ Betygen från grundskolan säger naturligtvis inte med någon 100-procentig säkerhet hur bra en enskild elev klarar sig på gymnasiet men det finns ett statistiskt signifikant positivt samband mellan elevers grundskolebetyg och gymnasiebetyg (Utbildningsförvaltningen 2006). Detta innebär att i genomsnitt, när man analyserar ett flertal elever, kan man förvänta sig att elever med relativt höga grundskolebetyg i genomsnitt också får relativt höga gymnasiebetyg.

⁵⁰ Analysen gäller således skolenivån och **inte** enskilda elevers relativa framgång på gymnasiet utifrån deras enskilda avvikelse mellan slutbetyg och provbetyg. Data tillåter inte sådana analyser på elevnivå.

⁵¹ Korrelationskoefficienter på -0,41 respektive -0,46 uppmättes beroende på metod. Båda korrelationerna är signifikanta på 1 % nivå.

⁵² Utbildningsförvaltningen 2006.

Men för ett mycket begränsat antal grundskolor (10 skolor) från år 2002, går det att koppla ihop data från utbildningsförvaltningens studie med resultaten från de nationella proven för ämnet matematik⁵³. När detta görs syns ett signifikant negativt samband (korrelation = -0,717 och signifikant på 5 % nivå) mellan grundskolors nettoavvikelse i matematik och hur elever från dessa skolor presterade i matematik A-kursen på gymnasiet. Elever från grundskolor med höga nettoavvikelser (i matematik) tenderar att prestera förhållandevis sämre på gymnasiet (betyg i matematik A) jämfört med elever från grundskolor med relativt låga nettoavvikelser. Således ytterligare ett bevis, om än baserat på ett mycket begränsat antal skolor, för att skillnader i skolors nettoavvikelser indikerar skillnader i skolornas betygsättning.

Dessa resultat, från 1998 respektive 2002 och som bygger på olika sätt följa hur elever från olika grundskolor klarar sig på gymnasiet, stöder därmed hypotesen att de variationer som syns i skolors nettoavvikelse, till viss del i alla fall, kan förklaras av att skolorna *inte* tolkar och tillämpar mål och betygsriterier på ett likartat sätt. Detta resultat stöder därmed argumentet att variationer i skolors nettoavvikelse mellan slutbetyg och provbetyg kan fungera som en indikator på hur likvärdigt betygen sätts på systemnivå.

Hur klarar sig elever från grundskolor med höga nettoavvikelser enligt PISA?

PISA är en internationell studie som mäter 15-åriga (= årskurs 9) elevers kompetens i läsförståelse, matematik och naturvetenskap.⁵⁴ Ett urval av cirka 150 grundskolor genomför ett kunskapsprov som i motsats till de nationella proven bedöms centralt och dessutom har betydligt striktare bedömningsanvisningar och är i den meningen mer "objektiva". En nackdel är att inte alla elever i skolan gör PISA-provet utan endast cirka 30 elever per skola. En annan nackdel är att PISA-provet inte nödvändigtvis mäter exakt samma kunskaper och kompetenser som de nationella proven eller de svenska kursplanerna prövar.⁵⁵

På samma sätt som i föregående avsnitt, där elevers gymnasiebetyg analyserades, undersöks här om det finns något samband mellan hur det går för elever på PISA-testet utifrån vad som kan förväntas utifrån deras slutbetyg, och hur hög nettoavvikelse elevernas skola har. Elever från skolor som sätter relativt höga betyg jämfört med vad eleverna presterar på de nationella proven (hög nettoavvikelse) kan då förväntas prestera relativt sämre på PISA-testet jämfört med elever med samma betyg men som kommer från skolor där nettoavvikelsen är lägre. Detta gäller utifrån antagandet att skolornas nettoavvikelse fungerar som en indikator på betygsättningen. Resultaten gäller för år 2003 års ämnesprov i matematik och svenska samt PISA 2003.

Resultaten visar på ett statistiskt signifikant samband för matematik.⁵⁶ Skolor med relativt höga nettoavvikelser i matematik tenderar att klara sig relativt sämre enligt PISA-testet, jämfört med skolor som har lägre nettoavvikelser men med samma betyg.

⁵³ För svenska blev antalet skolor för litet (7 skolor) och för engelska finns inget dataunderlag.

⁵⁴ För mer information om PISA på svenska, se (Skolverket 2004).

⁵⁵ Överensstämmelsen mellan PISA och den svenska läroplanen och kursplanerna är dock god visar analyser som nyligen gjorts på uppdrag av Skolverket. Se Skolverket 2006e och Skolverket 2006f.

⁵⁶ Korrelationen för matematik var -0,31 och signifikant på 1 % nivå. För svenska uppmättes en korrelation på -0,12, vilken inte var statistiskt signifikant.

För svenska finns inget statistiskt signifikant samband. Men det finns tre omständigheter som gör att analysen för svenska är tveksam. Eftersom svenska (läsförmåga i PISA) inte var ett huvudämne i PISA 2003 så är underlaget för resultat på skolnivå mycket osäkrare än för matematik. Det andra skälet är att PISA enbart prövar en relativt begränsad del av svenskämnet: läsförmågan. Det tredje skälet är att elever som läser svenska som andraspråk (många utlandsfödda elever) inte ingår i beräkningarna av slutbetyg och provbetyg men de ingår i PISA-populationen (under förutsättning att de befunnit sig i Sverige i minst 1 år). Detta innebär att en skola med många utlandsfödda elever kan komma att prestera sämre på PISA än vad slutbetygen antyder på grund av att det genomsnittliga slutbetyget är baserat bara på de elever som läst enligt kursplanen i svenska. Men skolan behöver inte ha en positiv nettoavvikelse mellan slutbetyg och provbetyg. Detta kan leda till att eventuella samband blir svåra att påvisa.

Även när det gäller matematik gäller att det svårt att påvisa statistiskt signifikanta samband mellan skolors nettoavvikelser och PISA-resultat eftersom det endast är ett urval av elever som genomför PISA i varje skola. Analysen som bygger på elevers gymnasiebetyg är mer stabil eftersom den bygger på uppgifter från så gott som samtliga elever.

Sammantaget måste sägas att den samlade empirin från avsnitt 4.3 ger ytterligare stöd åt förklaringen att skillnader i skolors nettoavvikelse kan ses som en indikator på skillnader i skolors betygsättning utifrån ett systemperspektiv. Detta innebär inte att det går att peka ut enskilda skolor med hög nettoavvikelse och hävda att de sätter för höga betyg. Eftersom det inte finns några föreskrifter för hur skolor skall förhålla sig till resultaten på de nationella proven går det inte att säga att vissa skolor sätter för höga betyg då det lika väl kan vara så att det är de andra skolorna som sätter för låga betyg. Men ur ett systemperspektiv går det att säga att de relativt stora variationer som går att observera i skolors nettoavvikelse indikerar att uttolkningen eller tillämpningen av betygsriterierna varierar mellan skolorna.

4.4 Hur ser relationen slutbetyg – provbetyg ut mellan klasser?

Hela detta kapitel har fokuserat på relationen slutbetyg – provbetyg på skolnivå och kommunnivå. Kapitel 2 tog framförallt upp motsvarande relation på nationell nivå. Men hur ser det ut när det gäller avvikelser mellan prov och slutbetyg mellan klasser/undervisningsgrupper/lärare inom en given skola?

Inom ramen för denna rapport har det inte funnits möjligheter att analysera denna variation i detalj, delvis på grund av det inte finns tillförlitliga data på klassnivå. Men det är viktigt att påpeka att det finns starka skäl att tro att det kan finnas en betydande variation i hur slutbetyg sätts i förhållande till provbetyg mellan olika klasser eller undervisningsgrupper inom en given skola.

Som ett enstaka och förvisso extremt exempel på variation mellan klasser inom skolor fanns år 2006 i ämnet matematik en skola med 189 elever fördelade på 10 undervisningsgrupper där en klass (20 elever) hade en nettoavvikelse på +70 procent medan en annan klass (15 elever) hade en nettoavvikelse på -20 procent.

Som redan nämnts i kapitel 1.8 kan avvikelser mellan olika klasser inom en skola ”skymma” resultaten när det gäller variationen mellan lärare. En skola som har nettoavvikelsen noll kan i själva verket bestå av en klass där alla elever fått ett höjt

slutbetyg i förhållande till proven och en klass där alla elever fått ett sänkt betyg. Nettoavvikelsen mellan skolor säger alltså långt ifrån allt om hur likvärdigheten mellan olika lärares betygsättning ser ut i förhållande till de nationella proven.

4.5 Sammanfattning

Det finns klara skillnader i hur både skolor och kommuner sätter betyg i förhållande till resultaten på de nationella proven. Det finns skäl att misstänka att åtminstone en del av de observerade skillnaderna mellan skolor indikerar att skolors sätt att tolka och tillämpa betygsriterierna skiljer sig. Denna förklaring får ytterligare stöd av de kompletterande analyser som görs med PISA-data och uppgifter om gymnasiebetyg. Andra tänkbara förklaringar till skillnaden mellan provbetygen och slutbetygen för skolorna diskuteras vidare i kapitel 8.

En viktig slutsats är därför att variationen mellan skolor i hur de sätter betyg i förhållande till resultaten på de nationella proven kan fungera som en indikator på hur likvärdig betygsättningen är ur ett systemperspektiv. Detta betyder dock inte att det går att säga att enskilda skolor sätter för höga eller för låga betyg. Däremot att de skillnader som observerats mellan skolors betygsättning i förhållande till provbetygen, indikerar att tolkning och tillämpning av betygsriterier inte är lika på alla skolor.

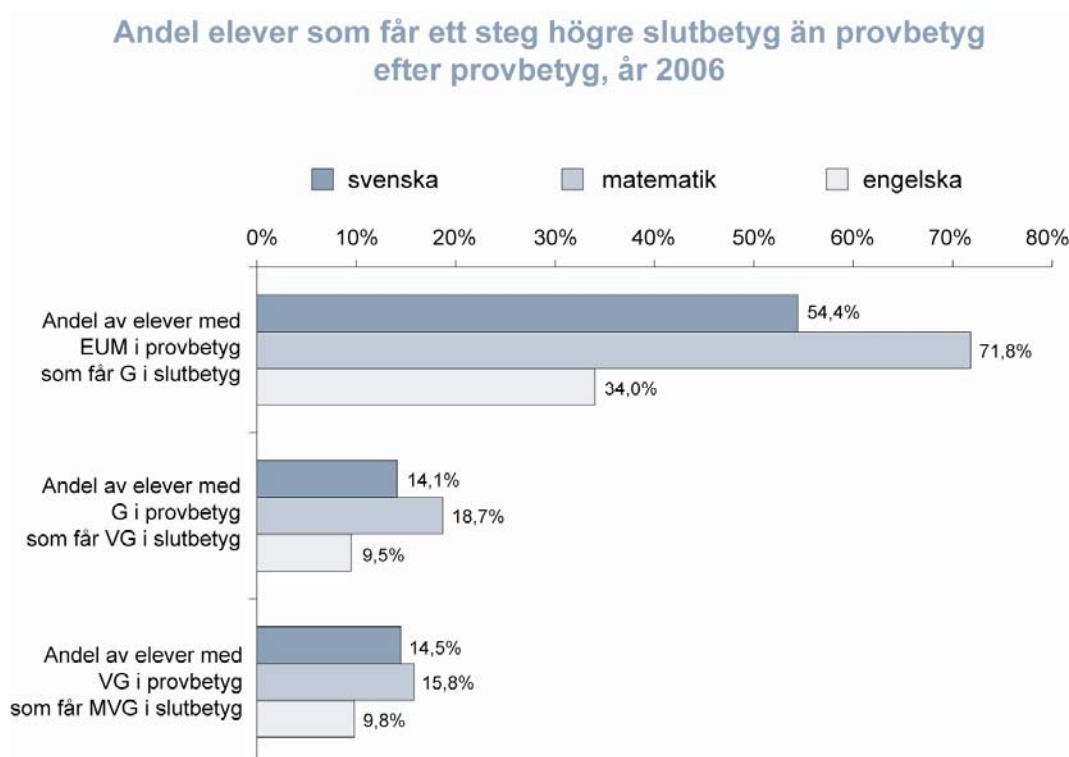
En annan viktig iakttagelse är att skillnaderna mellan hur skolor avviker mellan slutbetyg och provbetyg inte har förändrats över tidsperioden 1998-2006, något som skulle kunna tolkas som att de skillnader som verkar finnas i hur skolor sätter betyg inte verkar ha ökat eller minskat över tidsperioden. Utifrån denna indikator på likvärdig betygsättning har således likvärdigheten vare sig förbättrats eller försämrats sedan de nationella proven infördes i slutet av 1990-talet.

Kapitel 5. Vilka är det som får ett högre slutbetyg än provbetyg?

Kapitel 4 visade att betygsättningen i förhållande till provbetygen uppvisar ganska stora variationer mellan olika skolor och även mellan olika kommuner. En intressant fråga är då på vilket sätt betygsättningen skiljer sig mellan olika skolor. Sätter skolor med höga nettoavvikelser konsekvent höga betyg i förhållande till provbetyget eller är det bara för elever med vissa provbetyg, t ex G, som skolan tenderar att sätta ett högre slutbetyg för? Detta kapitel avser att försöka ge svar på dessa frågor genom att analysera skolors *betygsättningsprofiler*. Med *betygsättningsprofiler* menas att beskriva mer i detalj vilka elever som får ett högre slutbetyg än provbetyg. Är det framförallt elever som inte fått ett godkänt provbetyg som får ett högre slutbetyg eller är avvikelserna mellan slutbetyg och provbetyg relativt jämnt fördelade bland de olika provbetygen?

5.1 Betygsprofil – nationell bild

Figur 5.1 visar andelen elever som får ett högre slutbetyg än provbetyg av de elever som har ett givet provbetyg i ett ämne. Observera att då andelen elever som får samma eller lägre slutbetyg i förhållande till provbetyg inte finns representerade i figuren summerar inga andelar ihop till 100 procent.



Figur 5.1 Andel elever som får ett steg högre slutbetyg än provbetyg av de elever som har ett givet provbetyg, år 2006.

Siffrorna bygger på samtliga elever i riket (förutom bortfall). (svenska: 98 229 elever, matematik: 103 139 elever, engelska: 103 805 elever). Elever som läser svenska som andraspråk har exkluderats i analyserna.

Det som figur 5.1 framförallt bidrar med, jämfört med de resultat som presenterades i kapitel 2 (figur 2.3), är att det nu går att se exakt vilka betyg det är som höjs, inte bara den totala andelen elever som får ett steg högre betyg (jämför med figur 2.3). Till exempel kan ses att i svenska, av de elever som inte fick godkänt enligt provbetyget, har 54,4 procent ändå fått G i slutbetyg. Däremot har bara 14,1 procent av de elever som fick G i provbetyg i svenska fått VG i slutbetyg. Av dem som erhållit VG i provbetyg i svenska är det 14,5 procent som erhåller MVG i slutbetyg. (För de elever som har MVG i provbetyg går det förstås inte att höja slutbetyget och därför presenteras inga sådana siffror i figuren.)⁵⁷ För matematik får hela 71,8 procent av eleverna som inte uppnått målen enligt provbetyget ändå godkänt enligt slutbetyget. För engelska är motsvarande andel 34,0 procent.

Ovanstående andelar kan tolkas som sannolikheten att, givet ett visst provbetyg, erhålla ett högre slutbetyg (ett steg högre slutbetyg).⁵⁸ Sett över alla tre ämnena finns ett tydligt mönster att sannolikheten att få ett högre slutbetyg jämfört med provbetyget är störst för elever som inte fått godkänt enligt provbetyget. Det är alltså vanligast att lärare höjer upp elever som inte klarat godkänt på provet jämfört med att höja elever som fått G eller VG på provet.

5.2 Betygsprofiler för skolor med hög respektive låg nettoavvikelse

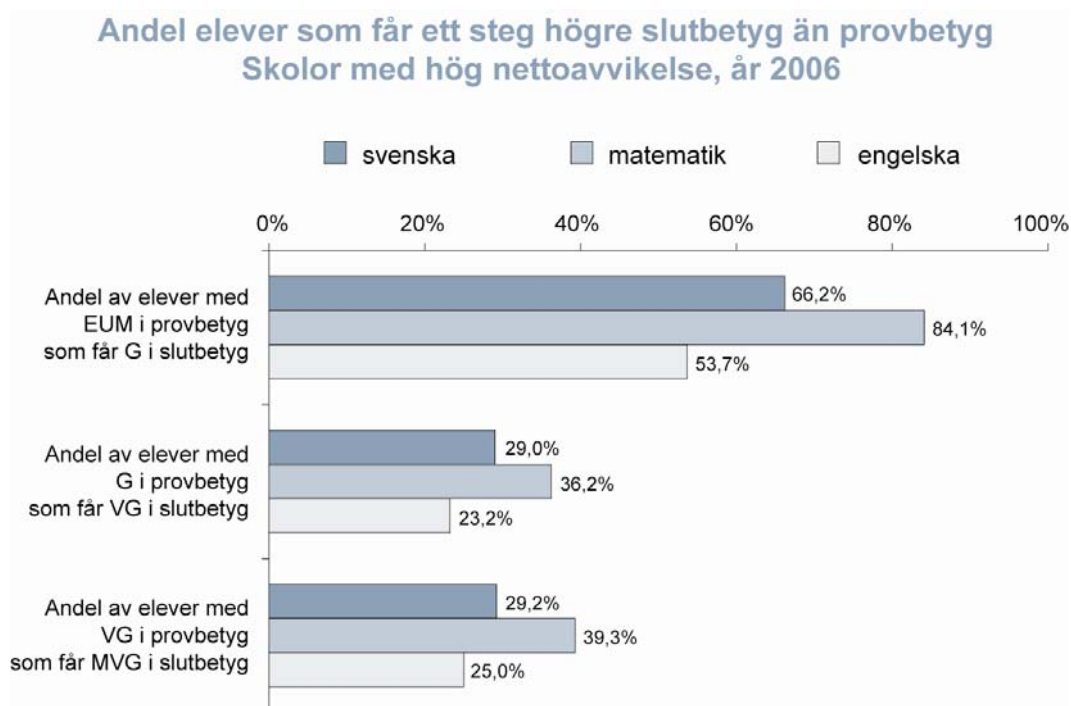
På vilket sätt skiljer sig betygsättningen mellan skolor med relativt hög nettoavvikelse jämfört med skolor med relativt låg nettoavvikelse? Som tidigare visats är det så att skolor med hög nettoavvikelse per definition sätter högre slutbetyg i genomsnitt jämfört med det genomsnittliga provbetyget. Motsatsen gäller för skolor med låg nettoavvikelse. Utifrån detta kan man förvänta sig att sannolikheten att få ett högre slutbetyg än provbetyg genomgående är större för varje provbetyg för skolor med relativt hög nettoavvikelse jämfört med skolor med relativt låg nettoavvikelse. Men frågan är om det bara är en nivåskillnad i andelen elever som får ett högre slutbetyg än provbetyg mellan dessa två grupper av skolor eller om *betygsprofilen* skiljer sig på något ytterligare sätt?

I figur 5.2 a-b har separata betygsprofiler tagits fram för dels de 10 procent skolor som har störst nettoavvikelse (figur 5.2a) samt dels för de 10 procent skolor med lägst nettoavvikelse (eller störst negativ nettoavvikelse) (figur 5.2b). Således jämförs bara de skolor som avviker mest respektive minst i fråga om nettoavvikelse.⁵⁹ På samma sätt som i figur 5.1 visas i figurerna andelen elever som får ett slutbetyg som är ett steg högre än provbetyget av de elever som har ett givet provbetyg. På samma sätt som i figur 5.1 kan dessa andelar (relativa frekvenser) tolkas som sannolikheten att få ett högre slutbetyg, givet ett visst provbetyg.

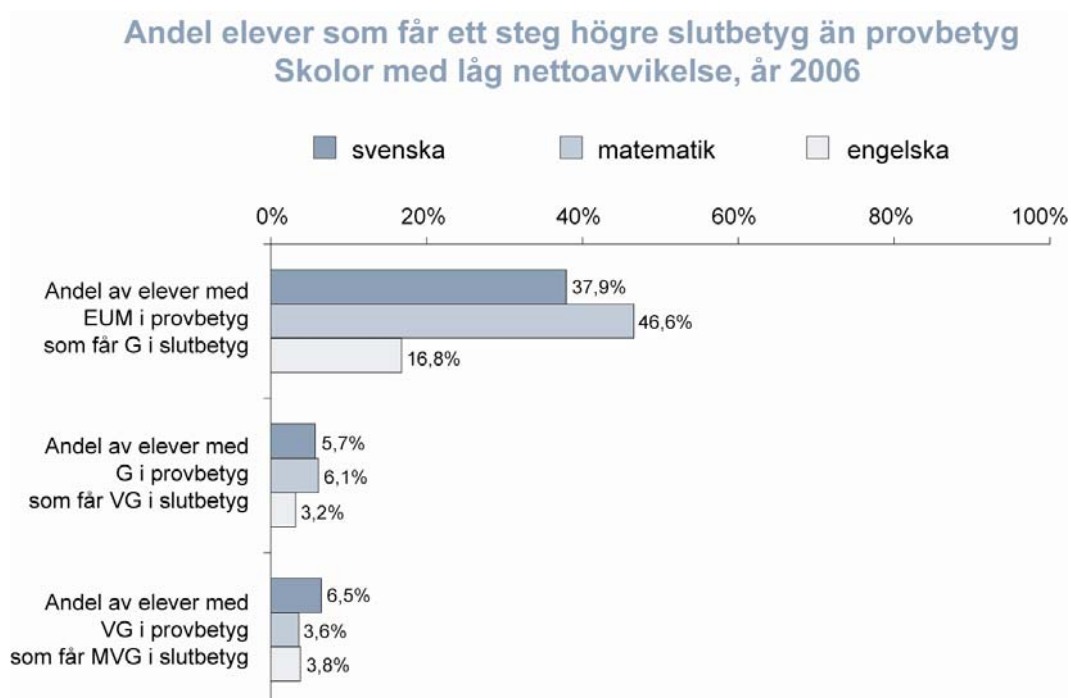
⁵⁷ Om istället andelen elever med MVG i provbetyg och som fick VG i slutbetyg beräknas är det 9 % i svenska, 6,3 % i matematik och 10,3 % i engelska.

⁵⁸ Andelen elever som får ett högre slutbetyg anger hur vanligt förekommande det är att erhålla ett högre slutbetyg givet ett visst provbetyg. Om en elev väljs slumpmässigt från en grupp elever med ett visst provbetyg, så är sannolikheten att just den eleven är en elev som fått ett steg högre slutbetyg lika med andelen elever i den gruppen som har ett högre slutbetyg. I grundläggande statistisk teori säger man att den relativa frekvensen (andelen) är en väntevärdesriktig skattning av sannolikheten.

⁵⁹ I figur 3.1a-c i kapitel 3 motsvaras dessa skolor av de skolor som ligger längst till vänster respektive längst till höger i fördelningen av skolornas nettoavvikelser.



Figur 5.2a Andelen elever som får ett steg högre slutbetyg än provbetyg, givet ett visst provbetyg. Data gäller de 10 % skolor med **högst** nettoavvikelse och baseras på 8060 elever i 129 skolor för svenska, 8882 elever i 130 skolor för matematik och 8035 elever i 131 skolor för engelska.



Figur 5.2b Andelen elever som får ett steg högre slutbetyg än provbetyg, givet ett visst provbetyg. Data avser de 10 % skolor med **lägst** nettoavvikelse och baseras på 8531 elever i 129 skolor för svenska, 9031 elever i 130 skolor för matematik och 9992 elever i 131 skolor för engelska.

Figur 5.2 a-b visar hur stora skillnaderna är i betygsättning mellan skolor med hög nettoavvikelse respektive skolor med låg nettoavvikelse. Vid en första anblick ser figur 5.2a och 5.2b mycket likartade ut. Både skolor med låg som hög nettoavvikelse uppvisar en betydande större sannolikhet att ge ett högre slutbetyg än provbetyg om eleverna inte har fått godkänt på det nationella provet. En observerbar skillnad mellan figurerna däremot, men som är förväntad, är att sannolikheten att få ett högre slutbetyg givet ett visst provbetyg är avsevärt högre för varje provbetygsteg i samtliga ämnen för skolorna med relativt höga nettoavvikelser jämfört med skolorna med relativt låga nettoavvikelser. Till exempel för svenska är sannolikheten att få G i slutbetyg givet EUM i provbetyg 66,2 procent om man går på en skola med hög nettoavvikelse (figur 5.2a) jämfört med endast 37,9 procent om man går på en skola med låg nettoavvikelse (figur 5.2b), dvs. nästan dubbelt så stor sannolikhet.

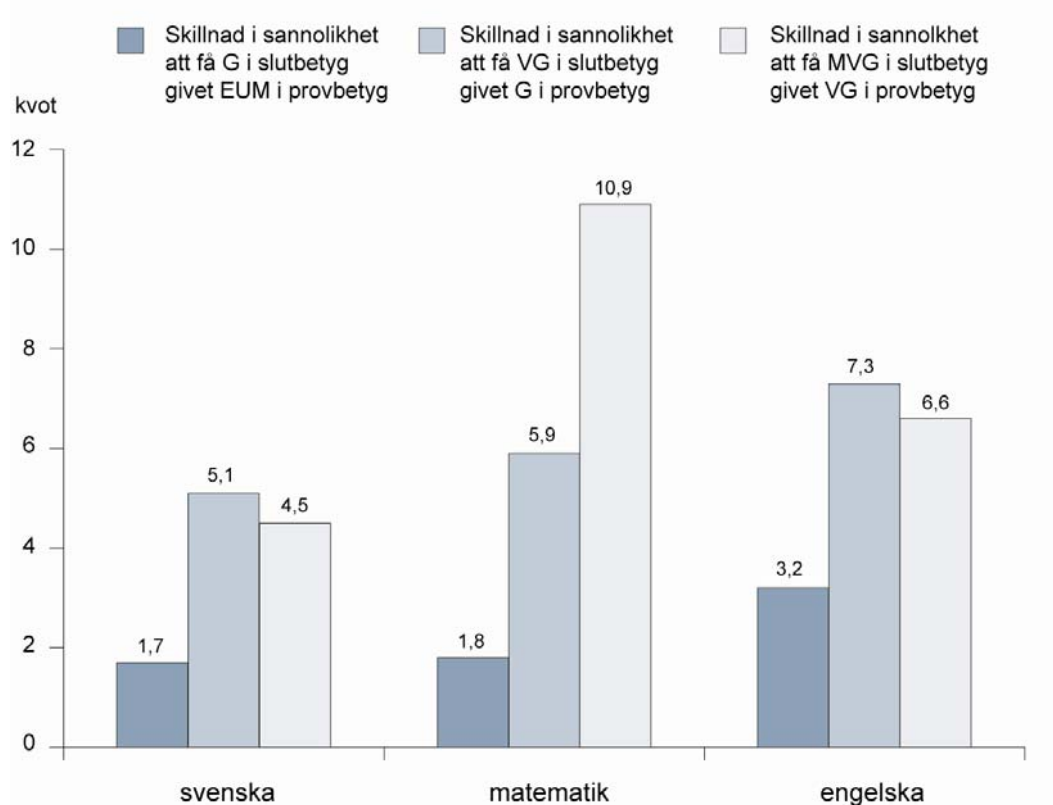
Eftersom det verkar vara ett generellt mönster att en förhållandevis större andel elever med EUM i provbetyg får ett högre slutbetyg är det tänkbart att skolor med en hög andel lågpresterande elever får en högre total nettoavvikelse, inte för att de nödvändigtvis skulle vara mer generösa i betygsättningen utan för att de har fler elever som skall "petas över ribban". Detta förefaller helt rimligt men vad figurerna 5.2 a-b visar är att denna orsak inte ensam kan förklara hela den skillnad i betygsättning som finns mellan skolor med höga respektive låga nettoavvikelser. I figur 5.2 a-b jämförs bara elever med samma provbetyg mellan skolor med hög respektive låg nettoavvikelse och därmed tas hänsyn till att skolorna kan ha olika sammansättning, dvs. att andelen elever med EUM i provbetyg är olika på olika skolor. Figur 5.2 a-b visar därmed att eftersom sannolikheten är genomgående högre för samtliga provbetyg på skolor med hög nettoavvikelse måste det finnas ytterligare förklaringar till de skillnader som finns i hur skolor sätter betyg i förhållande till provbetygen. Det kan naturligtvis finnas andra förklaringar, som till viss del i alla fall, kan hjälpa till att förklara dessa skillnader och som inte nödvändigtvis implicerar att det skulle finnas skillnader i tolkningen av betygskriterier. Men en annan tänkbar förklaring är att betygsättningskulturen faktiskt skiljer sig mellan dessa skolor, dvs. att det finns skillnader i hur skolorna tolkar och tillämpar mål och betygskriterier. Olika förklaringar diskuteras vidare i kapitel 8.

Men det mest intressanta med figur 5.2 a-b är att vid en närmare analys visar det sig att det finns ytterligare skillnader i hur betygen sätts. Visserligen har det redan konstaterats att sannolikheten att få ett högre slutbetyg än provbetyg är högre i skolor med hög nettoavvikelse för samtliga provbetygsteg. Men en jämförelse av figurerna a och b visar att skillnaderna inte är lika stora för varje provbetyg.

För svenska så är sannolikheten att få G i slutbetyg givet EUM i provbetyg 66,2 procent respektive 37,9 procent i skolor med hög respektive låg nettoavvikelse. Därmed är sannolikheten att få ett G i slutbetyg givet ett EUM i provbetyg 1,7 gånger större ($66,2/37,9 = 1,7$) om man går på en skola med hög nettoavvikelse jämfört med om man går på en skola med låg nettoavvikelse. Motsvarande sannolikhet för att få ett VG i slutbetyg givet ett G i provbetyg är däremot hela 5,1 gånger större ($29,0/5,7 = 5,1$) om man går på en skola med hög nettoavvikelse jämfört med om man går på en skola med låg nettoavvikelse. Och på samma sätt är sannolikheten för att få MVG i slutbetyg givet VG i provbetyg $29,2/6,5 = 4,5$ gånger större om man går på en skola med hög nettoavvikelse. Följaktligen är skillnaderna i betygsättning mellan skolorna i ämnet svenska betydligt större för de högre betygen än vad de är för betyget G.

I figur 5.3 redovisas dessa skillnader i sannolikheter (sannolikhetskvoter) för samtliga ämnen. En sannolikhetskvot på 1,0 skulle innebära att det inte fanns någon skillnad i sannolikhet att få ett högre slutbetyg än provbetyg mellan de två skolkategorierna som jämförs. En kvot större än ett innebär att sannolikheten att få ett högre slutbetyg än provbetyg är större i skolor med relativt hög nettoavvikelse. Staplarnas höjd är således ett mått på hur stora skillnaderna är i betygsättning mellan skolor med hög respektive låg nettoavvikelse. Ju högre stapeln är desto större skillnad. Figur 5.3 tillför egentligen ingen ny information i jämförelse med figur 5.2 a-b men tydliggör och sammanställer endast skillnaderna i sannolikhet mellan figur 5.2a och figur 5.2b.

Skillnad i sannolikhet (uttryckt som kvoter) att få ett högre slutbetyg än provbetyg mellan skolor med hög respektive låg nettoavvikelse



Figur 5.3 Skillnad i sannolikhet (uttryckt som sannolikhetskvoter) att erhålla ett slutbetyg som är ett steg högre än provbetyget, givet ett visst provbetyg, mellan de 10 % skolor med högst nettoavvikelse och de 10 % skolor med lägst nettoavvikelse.

Figur 5.3 visar att det finns en klar skillnad i betygsättningsprofil mellan skolor med relativt hög respektive relativt låg nettoavvikelse då sannolikhetskvoterna (höjden på staplarna) skiljer sig väsentligt mellan provbetygstegen inom varje ämne. Detta

innebär att även om det finns tydliga skillnader mellan skolorna gällande sannolikhet att få G givet EUM i provbetyg så är skillnaderna ännu större för de högre betygstegen. I matematik t ex är sannolikheten nästan 11 gånger (10,9) större att få MVG i slutbetyg givet VG i provbetyg om man går på en skola med hög nettoavvikelse jämfört med om man går på en skola med låg nettoavvikelse medan motsvarande skillnad i sannolikhet för elever med EUM i provbetyg endast är knappt två gånger större (1,8).

Ovanstående resultat visar att skolorna inte skiljer sig lika mycket åt när det gäller betygsättning av elever som inte fått godkänt på det nationella provet som de skiljer sig åt vad gäller betygsättning av elever som fått G eller VG i provbetyg. Vad beror detta mönster på? En förklaring är att lärare oavsett skola kanske inte gärna vill underkänna elever (förklaring (j) i kapitel 1). Eller annorlunda uttryckt, lärarnas betygsättning i förhållande till provresultaten kanske skiljer sig mellan skolorna men för elever som inte uppnått målen enligt provbetyget suddas dessa skillnader till viss del ut på grund av att lärare generellt, oavsett skola, inte gärna underkänner elever. Denna förklaring och andra förklaringar diskuteras vidare i kapitel 8.

5.3 Sammanfattning

Ur ett nationellt perspektiv är det betydligt vanligare att elever som inte fått ett godkänt provbetyg får ett högre slutbetyg än provbetyg jämfört med elever som erhållit de högre provbetygen (G och VG). När skolor med hög respektive låg nettoavvikelse jämförs visar det sig att skillnaderna i betygsättning mellan dessa skolor är betydligt större för de högre betygen än vad de är för det lägre betyget.

Kapitel 6. Konkurrens, fristående skolor och relationen provbetyg – slutbetyg

I skoldebatten om konsekvenserna av det fria skolvalet och den kraftiga ökningen av antalet fristående skolor har det framförts att en av konsekvenserna skulle vara att skolor känner sig pressade att sätta höga betyg som inte motsvarar elevernas faktiska kunskaper.⁶⁰ Argumentet är att skolor med höga genomsnittliga betyg kan framstå som bättre skolor, dvs. man ser det genomsnittliga slutbetyget som en kvalitetsindikator.

En tidigare studie av Christina Wikström har undersökt effekten av konkurrens på betygsättning när det gäller gymnasieskolan.⁶¹ Studien finner att skolor i kommuner där det finns andra konkurrerande skolor tenderar att sätta något högre betyg (vad hon kallar betygsinflation) men att skillnaderna är små. Däremot visar samma studie att fristående gymnasieskolor tenderar att sätta högre betyg än andra kommunala skolor. Wikström har utgått från jämförelser mellan genomsnittligt gymnasiebetyg och elevers resultat på högskoleprovet. I denna rapport ligger som bekant fokus på betygsättning i grundskolan och det kunskapsmått som betygen jämförs med är resultaten från de nationella proven.⁶² Ett grundläggande antagande, liksom i tidigare kapitel i rapporten, är att resultaten från de nationella proven kan ses som ett mer eller mindre ”oberoende” kunskapsmått, vilket naturligtvis kan ifrågasättas (se diskussion i kapitel 1).

6.1 Konkurrensutsatta skolor

Uti från relationen provbetyg – slutbetyg är det rimligt att, om konkurrenshypotesen är riktig, förvänta sig att skolor i områden som är relativt konkurrensutsatta i genomsnitt skall uppvisa större avvikelser mellan slutbetyg och provbetyg (större nettoavvikelser) jämfört med skolor i områden som är relativt mindre konkurrensutsatta.

I denna analys har tre olika definitioner på konkurrensutsatt skola använts. Den första innebär att alla skolor som ligger i de tre storstäderna Stockholm, Göteborg och Malmö (cirka 180 skolor) klassas som konkurrensutsatta medan övriga skolor (cirka 1120 skolor) klassas som icke konkurrensutsatta. Den andra definitionen innebär att förutom storstadsskolorna även skolor i förortskommuner klassas som konkurrensutsatta, dvs. alla skolor i de tre storstadsregionerna klassas som konkurrensutsatta (totalt cirka 390 skolor) och övriga skolor (ej storstadsregion) som icke konkurrensutsatta (cirka 910 skolor). Den tredje definitionen innebär att alla skolor som ligger i en kommun med minst tre skolor eller med minst två skolor, varav minst en fristående skola, klassas som konkurrensutsatta. Med denna definition klassas cirka 1100 skolor som konkurrensutsatta medan endast cirka 200 skolor definieras som icke konkurrensutsatta.

Dessa definitioner är inte perfekta indikationer på hur konkurrensutsatta skolorna verkligen är, för det skulle ingående analyser av varje enskild skola behöva göras,

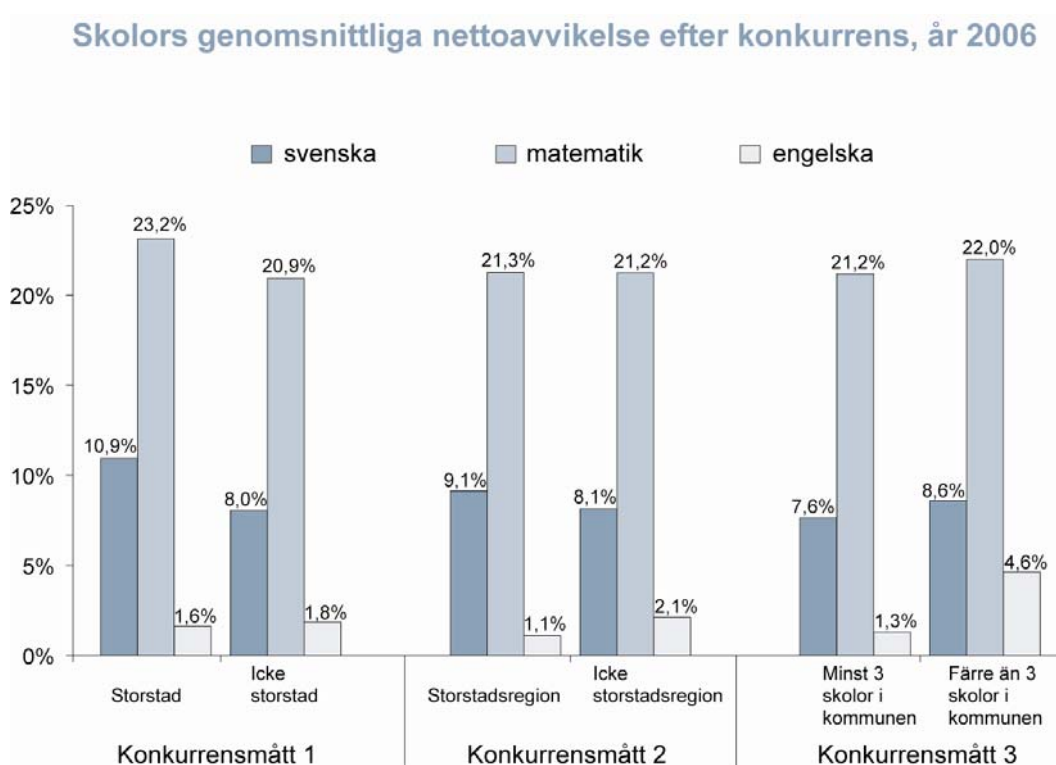
⁶⁰ Wikström 2005.

⁶¹ Wikström 2005. Se även Cliffordsson 2004 och Söderström 2006.

⁶² I Wikströms studie fungerar således resultaten från högskoleprovet som ett ”oberoende” kunskapsmått medan i denna rapport används resultaten från det nationella provet som det ”oberoende” kunskapsmåttet.

vilket inte har varit möjligt inom ramen för denna rapport. Men trots dessa brister kan det vara intressant att se om det finns några systematiska skillnader utifrån dessa definitioner.

I figur 6.1 visas genomsnittlig nettoavvikelse för skolor enligt de tre kategoriseringarna, storstad – icke storstad, storstadsområde – icke storstadsområde samt konkurrensutsatt – icke konkurrensutsatt. För meningsfulla jämförelser skall således skolor i storstad jämföras med skolor i icke storstad, skolor i storstadsregion med skolor i icke storstadsregion samt konkurrensutsatta skolor med icke konkurrensutsatta skolor.



Figur 6.1 Skolors genomsnittliga nettoavvikelse efter kategori. Tre olika definitioner på konkurrens: 1. Skola i storstad. 2. Skola i storstadsregion. 3. Skola i kommun med minst tre skolor eller minst två skolor varav minst en fristående skola. Data avser år 2006.

Till att börja med jämförs skolornas genomsnittliga nettoavvikelser i storstäderna (Stockholm, Göteborg och Malmö) med motsvarande nettoavvikelser för skolor i övriga landet (de två stapelgrupperna till vänster i figur 6.1.) För svenska är den genomsnittliga nettoavvikelsen 10,9 procent för skolor i storstad jämfört med 8,0 procent för övriga skolor (icke storstad), en skillnad på knappt 3 procentenheter. I matematik är den genomsnittliga nettoavvikelsen 23,2 procent för skolor i storstad och 20,9 procent för skolor som inte ligger i en storstad, en skillnad på drygt 2 procentenheter. För engelska är den genomsnittliga nettoavvikelsen ungefär lika stor för skolor i storstäder och icke storstäder, 1,6 procent respektive 1,8 procent.

Om istället kategoriseringen av konkurrensutsatta skolor vidgas till att även omfatta skolor som ligger i förortskommuner till storstäderna (jämförelse av de två stapelgrupperna i mitten i figur 6.1), så gäller för svenska att de genomsnittliga nettoavvikelse för skolor i storstadsregioner är 9,1 procent jämfört med 8,1 procent för skolor utanför storstadsregionerna, således en skillnad på 1 procentenhet. För matematik är den genomsnittliga nettoavvikelsen nästan lika stor för båda kategorierna, 21,3 respektive 21,2 procent. För engelska är den genomsnittliga nettoavvikelsen 1,1 procent för skolor i storstadsregion jämfört med 2,1 procent för övriga skolor, således en något högre nettoavvikelse för skolor utanför storstadsregionerna.

Med det tredje konkurrensmåttet är den genomsnittliga nettoavvikelsen i svenska för konkurrensutsatta skolor 7,6 procent respektive 8,6 procent för skolor som inte klassats som konkurrensutsatta. Motsvarande nettoavvikelse är 21,2 respektive 22,0 procent i matematik samt 1,3 respektive 4,6 procent i engelska. Således är den genomsnittliga nettoavvikelsen snarare större i de skolor som inte klassas som konkurrensutsatta för samtliga tre ämnen.

Ovanstående analys tar dock inte hänsyn till en del andra faktorer som kan påverka resultaten. Till exempel framkom i kapitel 4 att sannolikheten att få ett högre slutbetyg än provbetyg är betydligt större om man inte fått godkänt (EUM) på det nationella provet. Detta innebär att skolor med många relativt svaga elever kan framstå som att de sätter relativt höga betyg i förhållande till provbetygen men egentligen beror det bara på att de har en hög andel lågpresterande elever. Om det finns systematiska skillnader i denna sammansättning med avseende på skolor i storstäder/storstadsregioner eller det tredje konkurrensmåttet jämfört med övriga skolor kan detta påverka resultaten. På samma sätt finns det ytterligare faktorer som kan ha en systematisk inverkan på resultaten.

Den mest lämpliga metoden att undersöka om det finns några systematiska skillnader i relationen provbetyg – slutbetyg mellan skolor som klassats som konkurrensutsatta (enligt något av de tre kriterierna) och övriga skolor är i en regressionsmodell där det kontrolleras för andra relevanta faktorer. I den modell som använts har det kontrollerats för elevens kön och meritvärde på elevnivå samt skolans genomsnittliga provbetyg och om skolan är fristående eller kommunal (på skolnivå). Dessutom har kontroll gjorts för vilket provbetyg eleven har.⁶³ På så sätt kan man förenklat säga att man endast jämför elever med exakt samma värden på de bakgrundsvARIABLES som ingår i modellen, t ex innebär kontroll för faktorn kön att man endast jämför pojkar med pojkar och flickor med flickor.

Enligt resultaten från flernivåmodellen syns inte några tecken på att skolor som klassats som förmodat mer konkurrensutsatta skulle tendera att sätta högre slutbetyg i förhållande till provbetygen jämfört med övriga skolor, oavsett om man definierar konkurrensutsatta skolor som skolor i storstäder, skolor i storstadsregioner eller som skolor i kommuner med ett visst antal skolor. Detta gäller för samtliga tre undersökta ämnen (svenska, matematik och engelska).

En närliggande hypotes om effekter på betygsättning av konkurrens från andra skolor är att det uppstår en större variation i skolors nettoavvikelse i områden som är

⁶³ Kontroll för elevens provbetyg har gjorts genom att dels genomföra separata analyser för elever med varje provbetyg (EUM, G, VG), dels genom att använda så kallade dummyvariabler för elevens provbetyg. Resultaten blev i praktiken identiska oavsett metod.

relativt konkurrensutsatta. Anledningen enligt denna hypotes skulle vara att skolors betygsättning kan påverkas i olika utsträckning även om konkurrenstrycket är lika stort, beroende på lärarnas integritet och skolledningen. En analys av spridningen (variationen) i skolors nettoavvikelse visar inte på några betydande skillnader mellan konkurrensutsatta respektive icke konkurrensutsatta skolor.⁶⁴

Slutsatsen är att utifrån de data som rapporten använt och med de definitioner på konkurrensutsatt skola som använts, syns inga indikationer på att konkurrensutsatta skolor skulle sätta högre slutbetyg i förhållande till provbetyg. Men analysen skall inte ses som ett slutgiltigt bevis för att så verkligen är fallet. Förutom den trubbiga definitionen på vad som är en ”konkurrensutsatt skola” är en möjlighet att konkurrensutsatta skolor också kanske rättar de nationella proven lite ”snällare”, något som i så fall döljer eventuella mönster i dessa analyser. Detta går inte att pröva med tillgängliga data.

6.2 Fristående skolor

Ett par vetenskapliga undersökningar gällande gymnasieskolan och en journalistisk granskning gällande grundskolan har kommit fram till att det skulle vara lättare att få högre betyg om man går på en fristående skola givet samma egentliga prestationsnivå.⁶⁵ I detta avsnitt testas denna hypotes genom att jämföra nettoavvikelsen mellan slutbetyg och provbetyg för elever i fristående skolor och elever i kommunala skolor. Denna analys kan också ses som ett alternativt test av konkurrensypotesen då det kan tänkas att många fristående skolor startas där det finns utrymme för konkurrens och att just de fristående skolorna i högre utsträckning kanske skulle känna pressen att attrahera elever och därmed kanske frestas att sätta högre betyg.⁶⁶

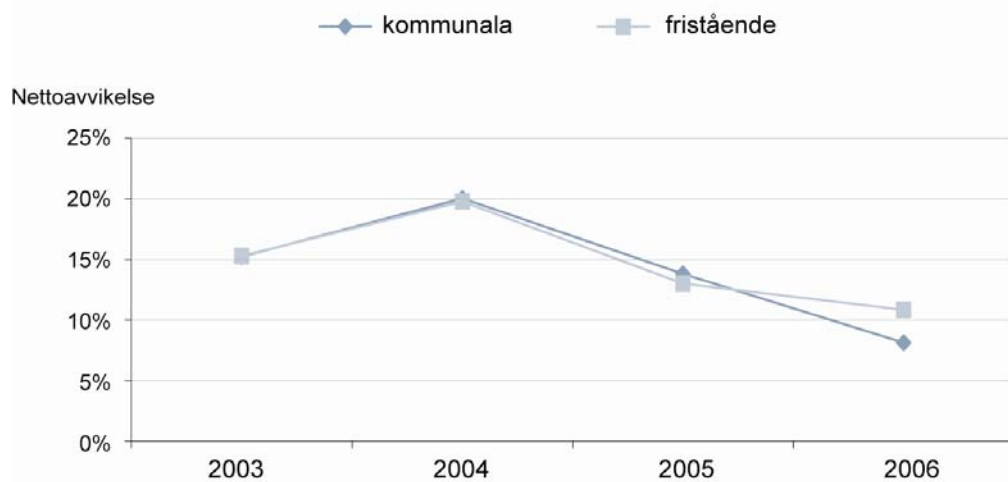
Figur 6.2a-c visar den genomsnittliga nettoavvikelsen mellan slutbetyg och provbetyg för kommunala respektive fristående skolor för tidsperioden 2003-2006 och ger en första indikation över sambandet mellan nettoavvikelse och huvudmannatyp. Åren före 2003, då insamlingarna av provresultat baserades på skolurval innehåller inte tillräckligt många fristående skolor för att göra meningsfulla jämförelser. Därför presenteras här endast resultat för åren 2003-2006.

⁶⁴ I svenska och matematik är spridningen marginellt större bland konkurrensutsatta skolor (oavsett definition) medan i engelska är förhållandet snarare omvänt. Men även i engelska är skillnaderna endast marginella.

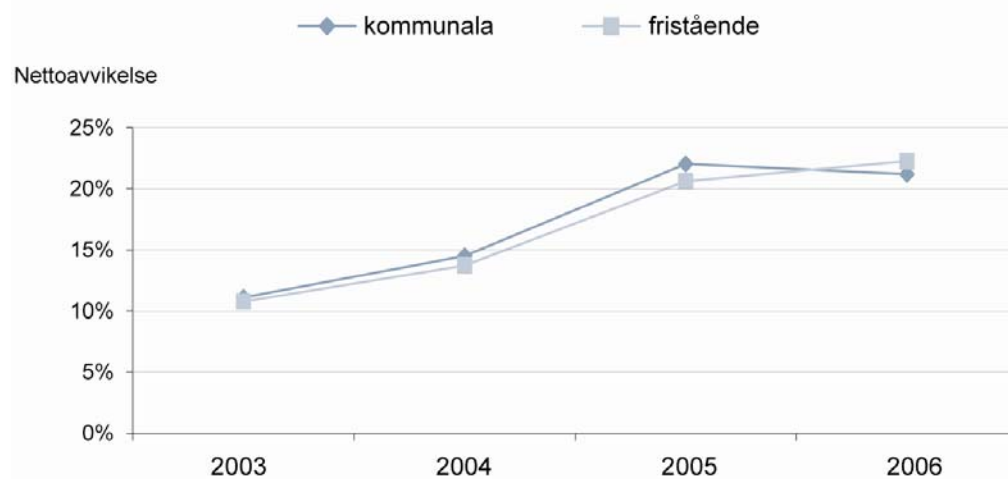
⁶⁵ Wikström 2005. Cliffordsson 2004. Sveriges Radio 2007.

⁶⁶ Wikström 2005.

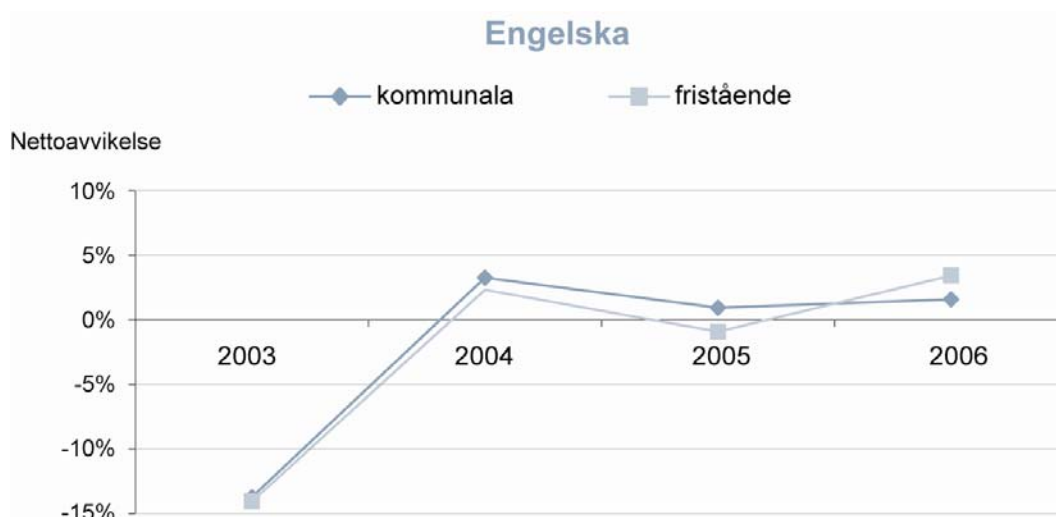
Nettoavvikelse mellan slutbetyg och provbetyg efter typ av huvudman Svenska



Nettoavvikelse mellan slutbetyg och provbetyg efter typ av huvudman Matematik



Figur 6.2a-b Nettoavvikelse mellan slutbetyg och provbetyg efter typ av huvudman, svenska och matematik.



Figur 6.2c Nettoavvikelse mellan slutbetyg och provbetyg efter typ av huvudman, engelska.

Figur 6.2a-c visar att skillnaderna i nettoavvikelse mellan kommunala och fristående skolor är mycket små. Figur 6.2a-c avser dock endast det enkla sambandet mellan nettoavvikelse och kommunal eller fristående skola, dvs. ingen kontroll har gjorts för andra faktorer/variabler som kan tänkas påverka sambandet.

När effekten av att gå på en fristående skola för sannolikheten att få ett högre slutbetyg än provbetyg analyseras i en regressionsmodell, där hänsyn tas till ett antal andra variabler som kan tänkas påverka sambandet, syns inte heller några signifikanta skillnader.⁶⁷ Detta innebär, enligt resultaten som gäller år 2006, att sannolikheten att få ett högre slutbetyg än provbetyg inte är högre på en fristående skola jämfört med en kommunal skola när hänsyn tas till övriga relevanta bakgrundsfaktorer.⁶⁸

Totalt sett ger inte ovanstående analys något stöd för hypotesen att det skulle vara lättare att få höga betyg i fristående skolor. Inte heller när motsvarande analys genomförs på data från år 2005 syns några samband som stöder denna hypotes. Inte heller i en tidigare av Skolverket publicerad studie av relationen provbetyg – slutbetyg i fristående skolor som är baserad på data från 2004 syns några skillnader mellan fristående skolor och kommunala skolor.⁶⁹

För att sammanfatta avsnittet om fristående skolor så finns det, utifrån dessa analyser, ingenting som tyder på att fristående skolor skulle sätta för höga betyg i förhållande till provbetyg i större utsträckning än kommunala skolor. Detta innebär naturligtvis inte att det inte är möjligt att det finns fristående skolor som sätter höga slutbetyg i förhållande till provbetygen. Det finns en stor variation i skolors

⁶⁷ Endast i engelska och för elever med VG i provbetyg kunde ett signifikant positivt samband påvisas. För övriga provbetyg i engelska och för alla provbetyg i svenska och matematik påvisades inga signifikanta samband.

⁶⁸ Regressionsmodellen som använts är identisk med den regressionsmodell som användes i föregående avsnitt med den skillnaden att nu är det skolans huvudmannaskap som står i fokus för analysen. På elevnivå ingår kön och meritvärde samt elevens provbetyg och på skolnivå ingår skolans genomsnittliga provbetyg samt om skolan är konkurrensutsatt eller ej. Analysen gäller år 2006.

⁶⁹ Skolverket 2005d.

nettoavvikelse bland fristående skolor precis som det finns en stor variation i skolors nettoavvikelse bland kommunala skolor, men det finns inget generellt mönster att dessa skolor i genomsnitt skulle sätta högre slutbetyg i förhållande till provbetygen när hänsyn tas till andra relevanta bakgrundsfaktorer.

Utifrån dessa data och denna analys går det således inte att finna några belägg för att den ökade konkurrensen skulle ha drivit på skolor att sätta högre betyg, i alla fall inte på ett olikformigt sätt, dvs. att fristående skolor eller skolor som kategoriserats som konkurrensutsatta i högre utsträckning skulle utmärka sig vad gäller detta påstådda beteende.

Men det finns ett par svagheter i analysen som gör att det trots allt inte går att se detta som slutgiltiga bevis för att så inte är fallet. För det första används meritvärde som kontrollvariabel och det skulle rent teoretiskt kunna vara så att om fristående skolor sätter för höga betyg, och gör så inom samtliga ämnen, så blir analysen till viss del självuppfyllande då den så att säga ”kontrollerar bort” skillnader i betygsättning genom att kontrollera för meritvärde. För det andra skulle det kunna vara så att lärare på fristående skolor, om de sätter ”för höga” betyg, även rättar de nationella proven lite ”snällare”. I så fall skulle också eventuella skillnader döljas. Båda dessa argument är även giltiga när det gäller analysen av konkurrensutsatta skolor.

I relation till Wikströms resultat, som visade på tydliga skillnader i hur fristående skolor sätter betyg så skall påminnas att Wikströms studie gäller gymnasieskolan. Det kan också vara så att det är lättare att skilja ut konkurrensutsatta skolor på gymnasienivå då det som regel finns färre gymnasieskolor i en kommun än grundskolor. En möjlig tolkning är också att konkurrensstrycket helt enkelt är större på gymnasiet jämfört med grundskolan. Debatten om huruvida fristående skolor sätter högre betyg i förhållande till kommunala skolor skall utifrån ovanstående resultat inte anses avslutad och hypotesen slutligt avfärdad utan det kan bara konstateras att utifrån resultaten på de nationella proven finns ingen skillnad mellan kommunala och fristående skolor, när hänsyn tas till andra relevanta faktorer. På samma sätt, och med samma förbehåll, finns inte heller några skillnader mellan skolor som klassificeras som konkurrensutsatta och övriga skolor.

6.3 Sammanfattning

Hypotesen om att skolor som är konkurrensutsatta skulle tendera att sätta högre betyg får inget stöd utifrån analyserna av relationen provbetyg – slutbetyg när hänsyn tas till andra faktorer. Detta gäller även när olika varianter på definitionen av en konkurrensutsatt skola testas. När fristående skolor jämförs med kommunala skolor framträder inte heller några generella skillnader vad gäller den genomsnittliga skillnaden mellan slutbetyg och provbetyg.

En svaghet med samtliga analyser i detta kapitel (liksom delvis även i andra kapitel) är att om det finns systematiska skillnader i betygsättningen och om det dessutom finns systematiska skillnader i bedömningen av proven (som går åt samma håll), så försvagas de systematiska mönstren som analyserna undersöker. Genom att kontrollera för meritvärde finns också en teoretisk risk att man kontrollerar bort eventuella systematiska skillnader i betygsättning om konkurrensutsatta och/eller fristående skolor genomgående sätter för höga betyg i alla ämnen. Ovanstående resultat skall därför inte ses som slutliga bevis för att det inte kan finnas systematiska skillnader mellan skolor som är konkurrensutsatta och övriga skolor eller mellan fristående skolor och kommunala skolor. Om däremot analyserna trots nämnda

svagheter hade visat på signifikanta samband hade det varit ett ganska tungt bevis för att det verkligen finns systematiska skillnader.

Kapitel 7. Relationen provbetyg-slutbetyg och skolans genomsnittliga prestationsnivå

I kapitel 4 visades att relationen mellan slutbetyg och provbetyg varierar mellan olika skolor. I detta avsnitt undersöks om nettoavvikelsen på en skola, dvs. skolans genomsnittliga skillnad mellan slutbetyg och provbetyg, är kopplad till hur väl elevernas allmänna prestationsnivå är på just den skolan. Syftet med en sådan analys är att eventuellt kunna få ökad klarhet i huruvida hypotesen att det skulle vara svårare att få höga betyg på en skola där den genomsnittliga prestationsnivån är relativt hög, har någon grund.⁷⁰

Dagens betygssystem är kriterierelaterat som redan beskrivits i kapitel 1, vilket innebär att lärarna skall sätta betyg utifrån fasta och förutbestämda mål och betygskriterier som är lika för alla elever i hela landet, oavsett skola. Detta innebär att en elevs prestation skall bedömas oberoende av hur andra elever på samma skola presterar. Som redan nämnts så finns det en potentiell risk att tolkningen av betygskriterierna kan variera mellan lärare och mellan skolor. En inte helt orimlig hypotes är då att en lärares tolkning av betygskriterierna kan påverkas av hur eleverna i skolan och i klassen brukar prestera i de olika skolämnena (se förklaring (g) i kapitel 1). Till exempel kan en lärare som länge arbetat på en skola med en genomsnittligt låg prestationsnivå kanske tolka mål och betygskriterier något ”snällare” jämfört med en lärare som varit verksam på en skola med relativt hög genomsnittlig prestationsnivå. I ett av uppnåendemålen i matematik står t ex att eleven skall kunna ”lösa enkla ekvationer”. Vad menas med ”enkla”? Det finns risk att tolkningen av vad som är en ”enkel” ekvation ges olika innebörd beroende på lärarens erfarenhet av vad som kan anses vara rimligt, vilket i sin tur kan vara beroende av hur duktiga eleverna i lärarens omgivning i allmänhet är i matematik.

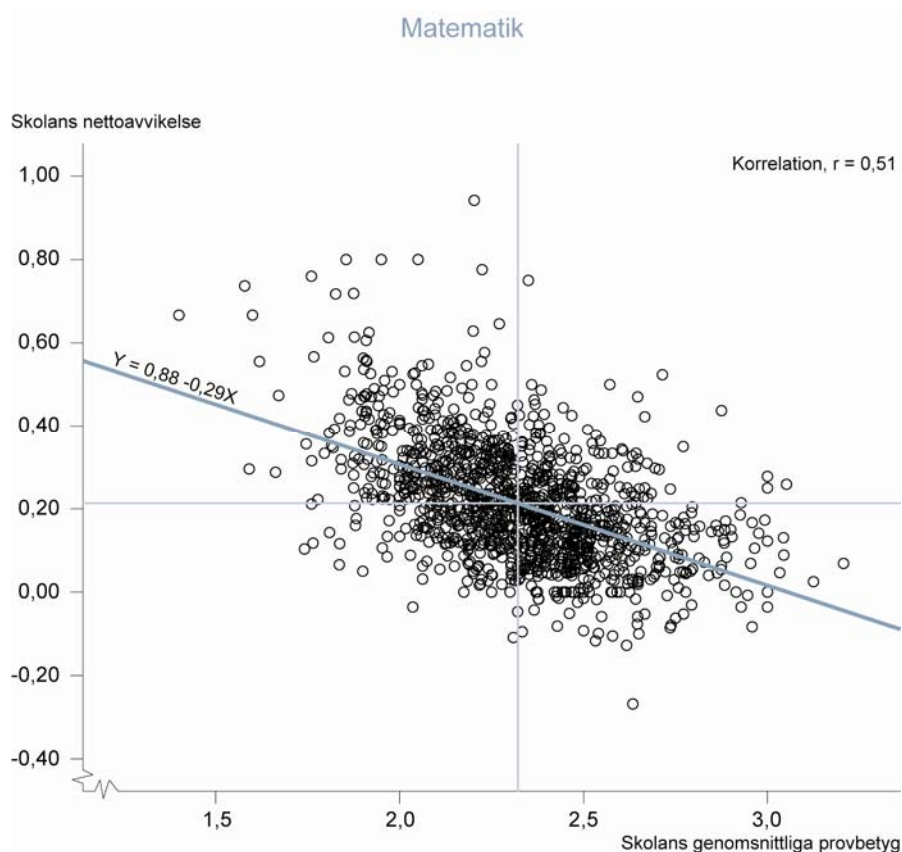
Om denna hypotes är riktig så är lärarens betygsättning till viss del *relativ* i den meningen att den är beroende av hur duktiga eleverna är på skolan. Detta är vad som i detta kapitel menas med *relativ* betygsättning. Enligt hypotesen utgår lärarna visserligen från samma mål och betygskriterier, men eftersom det finns ett visst tolkningsutrymme använder olika lärare till viss del även elevernas allmänna prestationsnivå som referenspunkt för tolkningen av mål och kriterier.

Huvudsyftet med detta avsnitt är således att undersöka om det finns någon koppling mellan skolors nettoavvikelse (skolans genomsnittliga skillnad mellan slutbetyg och provbetyg) och skolors genomsnittliga prestationsnivå. Som mått på skolans genomsnittliga prestationsnivå används i dessa analyser skolans genomsnittliga provbetyg för det aktuella ämnet.

Nettoavvikelse och skolors genomsnittliga provbetyg – en första inblick

För att få en första inblick i hur sambandet mellan skolors genomsnittliga prestationsnivå och nettoavvikelse ser ut, visas i figur 6.3 det enkla sambandet mellan skolors genomsnittliga provbetyg och skolors nettoavvikelse för matematik år 2006. Varje ring representerar en skola och skolans genomsnittliga provbetyg i matematik anges på den horisontella axeln och skolans nettoavvikelse i matematik på den vertikala axeln.

⁷⁰ En intervjustudie som jämför studieförberedande och yrkesförberedande gymnasieprogram på detta tema finns redovisad i Skolverket 2005f och Korp 2006.



Figur 7.1 Samband mellan skolors genomsnittliga provbetyg och nettoavvikelse. Matematik år 2006. 1304 skolor. Den gröna vertikala linjen anger medelvärdet i provbetyg för samtliga skolor medan den gröna horisontella linjen anger medelvärdet för samtliga skolors nettoavvikelse. Den röda linjen anger det skattade linjära sambandet mellan skolors genomsnittliga provbetyg och skolors nettoavvikelse.

Figur 7.1 visar att det finns ett tydligt negativt samband mellan skolors genomsnittliga provbetyg och nettoavvikelse. Det betyder att skolor som har höga resultat på det nationella provet tenderar att sätta slutbetyg som ligger nära eller under provbetygen, dvs. relativt låga nettoavvikelser (skolor i den nedre högra rutan av diagrammet). Skolor som har låga resultat på det nationella provet tenderar i högre grad att sätta slutbetyg som ligger högre än provresultatet, dvs. relativt höga nettoavvikelser (skolor i den övre vänstra rutan av diagrammet). Men figuren visar också att det finns många skolor som inte följer detta mönster. Det finns många skolor med höga genomsnittliga provbetyg och samtidigt relativt höga nettoavvikelser (skolor i den övre högra rutan). Och det finns också gott om exempel på skolor med låga genomsnittliga provbetyg men med relativt låga nettoavvikelser (skolor i den nedre vänstra rutan). Sambandet mellan prestationsnivå och

nettoavvikelser är alltså långt ifrån 100-procentigt men tendensen är ändå tydlig och sambandet är statistiskt signifikant på 1 % nivå.⁷¹

Nettoavvikelse och skolors genomsnittliga provbetyg – efter kontroll för andra bakgrundsfaktorer

Enligt samma princip som i avsnitt 6.1 är det mest korrekta sättet att analysera sambandet mellan nettoavvikelse och skolors genomsnittliga prestationsnivå att skatta sambandet i en regressionsmodell (i två nivåer) där hänsyn tas till andra relevanta bakgrundsfaktorer. I denna modell ingår, på samma sätt som tidigare, elevens kön och meritvärde på elevnivån samt vilket provbetyg eleven har. På skolnivån ingår variabeln som analysen fokuserar på, skolans genomsnittliga provbetyg, samt om skolan är en kommunal eller en fristående skola samt även om skolan är konkurrensutsatt eller inte.

Resultaten presenteras i tabell 7.1 nedan. Frågan var alltså om det finns en effekt av skolans genomsnittliga provbetyg när det gäller sannolikheten att en elev skall få ett slutbetyg som är högre än provbetyget. Utifrån hypotesen om att det skulle vara svårare att få högre betyg om man går på en skola med många högpresterande elever, så bör regressionsanalysen uppvisa negativa samband mellan sannolikheten att få högre slutbetyg än provbetyg och skolans genomsnittliga provbetyg. Det vill säga att ju högre skolans genomsnittliga provbetyg är, desto lägre bör sannolikheten vara för en elev att få ett högre slutbetyg än provbetyg, allt annat lika.

Tabell 7.1 Effekter av *skolans genomsnittliga provbetyg* när det gäller sannolikheten att få ett högre slutbetyg än provbetyg.⁷²

	Elever med EUM i provbetyg	Elever med G i provbetyg	Elever med VG i provbetyg
<i>Svenska</i>	Negativt samband	Inget samband	Inget samband
<i>Matematik</i>	Inget samband	Negativt samband	Negativt samband
<i>Engelska</i>	Inget samband	Inget samband	Negativt samband

5 % signifikansnivå har använts för att bestämma om det finns ett samband eller ej.

För svenska syns ett negativt samband för elever med EUM i provbetyg. För matematik syns negativa samband för provbetygen G och VG och för engelska syns ett negativt samband för elever med VG i provbetyg.

Tabell 7.1 ger ett delvis splittrat intryck och det är svårt att utifrån dessa analyser och data dra några definitiva slutsatser om huruvida denna form av relativ betygsättning förekommer generellt eller inte. En möjlig tolkning av den samlade resultatbilden är att det i alla fall inte finns några klara belägg för att det förekommer relativ betygsättning i någon större utsträckning i svenska och engelska. Däremot kan det,

⁷¹ Att det finns ett samband kan bekräftas dels av en negativ korrelationskoefficient som är statistiskt signifikant på 1 % nivå eller (vilket i detta fall är samma sak) att lutningskoefficienten i en skattad regression mellan genomsnittligt provbetyg och nettoavvikelse är signifikant negativ på 1 % nivå.

⁷² Anledningen till att inga resultat finns för elever med MVG i provbetyg är att det inte går att få ett högre slutbetyg än provbetyg för dessa elever, eller annorlunda uttryckt, sannolikheten är noll oavsett vilken skola eleverna går på.

utifrån resultaten, finnas vissa skäl att misstänka att det kan förekomma relativ betygsättning inom ämnet matematik. Även när motsvarande analyser görs för 2005 data är mönstret mycket likartat, med matematik som det ämne som uppvisar vissa indikationer på att betygsättningen kan variera beroende på skolans genomsnittliga prestationsnivå i ämnet.

7.1 Sammanfattning

Hypotesen om relativ betygsättning, som innebär att det skulle vara svårare att få höga betyg på skolor där den genomsnittliga prestationsnivån är hög får ett visst stöd för ämnet matematik men inget stöd för svenska och engelska.

Samma reservationer som i sammanfattningen av föregående kapitel är giltiga även när det gäller slutsatserna om ”relativ” betygsättning.

Kapitel 8. Resultat och diskussion

8.1 Inledning

Syftet med denna rapport har varit att utifrån en statistisk analys av relationen provbetyg – slutbetyg, analysera vissa aspekter av likvärdig betygsättning på systemnivå. En viktig förutsättning eller restriktion för analyserna är att det inte finns några juridiskt bindande förordningar eller rekommendationer för hur lärare/skolor/kommuner skall förhålla sig till resultaten på de nationella proven. En central utgångspunkt för denna rapport är därmed att det utifrån nuvarande bestämmelser vad gäller betyg och nationella prov inte går att uttala sig om huruvida enskilda skolor eller huvudmän sätter för höga eller för låga betyg, oavsett hur mycket de avviker från resultaten på de nationella proven.

Hur bra är proven som ett oberoende mått på kunskaperna i ett ämne? Det är omöjligt att ge ett klart svar på den frågan. I den här rapporten, när provbetygen används som en indikator på likvärdig betygsättning mellan skolor, är det viktigt att proven inte varierar särskilt mycket från år till år. Så är inte heller fallet. I figur 2.2 kan man se att provresultaten varierar mellan olika år men skillnaden är inte särskilt stor. Slutbetygen varierar dock betydligt mindre från år till år och det tycks inte vara så att ett ovanligt lätt eller svårt ämnesprov påverkar betygsättningen för en viss årskull i någon större utsträckning. Det innebär att de nationella proven knappast påverkar likvärdigheten negativt mellan olika årskullar.

Genom att jämföra betygen på de nationella proven med slutbetygen går det att få ett underlag för en diskussion om på vilka grunder lärare sätter betyg. Skillnader mellan lärare och skolor när det gäller i vilken utsträckning man avviker från provresultaten kan fungera som en *indikation* på att betygen sätts på olikartade grunder. Stora avvikelser på en skola och små på en annan tyder på att det kan finnas olika synsätt när det gäller grunderna för betygsättningen.

8.2 Rapportens viktigaste resultat

Nedan följer sammanställning av rapportens viktigaste resultat. En mer utförlig diskussion om tänkbara förklaringar till resultaten följer i nästa avsnitt.

- Ungefär 8 av 10 elever får samma slutbetyg som provbetyg. Det gäller oavsett ämne och det är inte någon stor förändring mellan 1998-2006.⁷³ Dock kan konstateras att korrelationen på individnivå var något lägre för de år (2002-2004) när det inte gjordes en sammanvägning till ett samlat provbetyg i svenska och engelska. Sammantaget kan dock konstateras att proven varken blivit mer eller mindre styrande för betygsättningen av enskilda elever över åren. Enligt provsystemets grundförutsättningar är det, utifrån denna empiri, heller inte nödvändigtvis något eftersträvänsvärt att vare sig öka eller sänka andelen som får samma slutbetyg som provbetyg. Proven ska inte vara examensprov och det finns flera legitima skäl för en lärare att sätta ett annat betyg på en elev än vad denne fått på det nationella provet.
- Det är betydligt vanligare att en elev får ett höjt än ett sänkt slutbetyg i förhållande till provbetyget i matematik och svenska. I engelska är det ungefär lika många som får ett höjt som ett sänkt slutbetyg i förhållande till

⁷³ Även i det gamla relativa betygssystemet låg korrelationen på ungefär 0,80. Skolverket 1995.

provbetyget. Mönstret är mycket stabilt över åren 1998-2006. Detta innebär att de genomsnittliga slutbetygen i matematik och svenska överstiger de genomsnittliga provbetygen på nationell nivå varje år sedan provens tillkomst 1998. I engelska ligger prov- och slutbetygens genomsnitt mycket nära varandra i stort sett varje år. Detta indikerar att det kan finnas skillnader mellan hur betygsättningen görs i de olika ämnena.

- Att det finns avvikelser mellan provbetyg och slutbetyg på nationell nivå behöver inte vara ett problem ur likvärdighetssynpunkt. Om lärare ger eleverna ett högre betyg i förhållande till provbetygen i samma utsträckning för samtliga betygssteg i samtliga skolor så påverkar det inte konkurrensen mellan eleverna. Det skulle därmed inte heller påverka likvärdigheten mellan eleverna. Det enda problemet skulle då vara att de nationella proven inte fungerar som ett riktigt bra instrument för att pröva på vilken betygsnivå eleverna ligger. Men så ser inte verkligheten ut (se nedanstående punkt).
- Ett huvudresultat är att det är stora skillnader mellan hur skolor (och kommuner) sätter betyg i förhållande till de nationella proven. Även om det finns ett stort antal skolor som har liknande nettoavvikelser är det sammantaget ändå så pass stora skillnader att det är svårt att hävda att det finns en nationellt likvärdig betygsättning. Det är svårt att ge *ett* enda mått på *hur* lika eller olika skolorna sätter betyg i förhållande till provresultaten. Man måste använda flera mått. Genom att se på intervallet -10 till +10 procentenheter i nettoavvikelser från noll. (nettoavvikelser = genomsnittlig andel som får ett steg högre eller lägre betyg än prov). Då ligger i engelska drygt 70 procent, i svenska cirka 60 procent och i matematik drygt 20 procent inom det intervallet. Skillnaden mellan skolor kan också beskrivas i form av spännvidden i nettoavvikelser. Här finns för samtliga tre ämnen exempel på skolor som har en mycket stor positiv nettoavvikelse (över 70 procent) och samtidigt finns det skolor som har en kraftig negativ nettoavvikelse. Spännvidden mellan de mest "generösa" respektive "stränga" skolorna är ungefär lika stora i de tre ämnena. Den genomsnittliga spridningen (standardavvikelsen) från det nationella genomsnittet är för 2006 14,1 procentenheter i matematik, 11,1 procentenheter i svenska och 10,8 procentenheter i engelska.
- Det går inte att utifrån detta resultat dra slutsatsen att skolor med hög positiv nettoavvikelse nödvändigtvis sätter för höga betyg och skolor med negativ nettoavvikelse sätter för låga betyg. Men den stora variationen i nettoavvikelser indikerar att en del skolor troligen tolkar betygskriterierna annorlunda i förhållande till en del andra skolor, vilket i sin tur är en indikation på att det finns bristande likvärdighet i betygsättningen.
- Det finns en klar tendens att det är samma skolor som har höga respektive låga nettoavvikelser i ett ämne från år till år. Det finns även en viss tendens att höga respektive låga nettoavvikelser samvarierar mellan olika ämnen, särskilt mellan engelska och svenska. Delvis kan det bero på att det i vissa fall är samma lärare som sätter betyg i svenska och engelska. Att avvikelserna i viss mån samvarierar från år till år och även mellan ämnen tyder på att det i viss utsträckning finns en "skolkultur" för hur man betygsätter eleverna i förhållande till proven, att lärarna på något vis får en samsyn på hur man ska

sätta betyg i förhållande till de nationella proven. Det kan alltså gälla både att en skola i betygsättningen ligger mycket nära provbetygen och långt ifrån provbetygen. Samtidigt finns det många avsteg från dessa mönster (figur 3.3). Det är långt ifrån ovanligt att en skola har en stor nettoavvikelse för ett ämne ett år för att nästa år ha en låg nettoavvikelse. Och det är inte heller ovanligt att det är stora avvikelser i ett ämne på en skola och låga avvikelser i ett annat ämne. Samtidigt som det går att skönja en ”skolkultur” där man år från år har en liknande avvikelse mellan prov och betyg så finns det alltså samtidigt många skolor där en sådan ”kultur” inte alls tycks finnas.

- En indikation på att skolors olika avvikelser mellan nationella prov och betyg också står för olika ”faktiska” kunskapsnivåer hos eleverna är resultaten från analyserna av gymnasiebetygen och en jämförelse med den internationella PISA-undersökningen. Elever som gått på skolor med höga nettoavvikelser tenderar att få lägre gymnasiebetyg jämfört med elever (med samma grundskolebetyg) som gått på skolor med låga nettoavvikelser mellan prov och betyg. Skolor med höga nettoavvikelser i matematik tenderar också att få sämre resultat i PISA-undersökningen i matematik jämfört med skolor med låga nettoavvikelser. Återigen går det inte att med denna analys avgöra vilka skolor som sätter mest ”rättvisa” betyg men analysen styrker hypotesen att skolorna utifrån ett systemperspektiv sätter betyg på olika grunder.
- Variationen mellan hur skolor sätter betyg i förhållande till provbetygen har genomsnittligt sett inte minskat men heller inte ökat under tidsperioden från det att de första slutbetygen i det nya betygssystemet började sättas fram till idag (1998-2006). Då flera åtgärder vidtagits för att förbättra samstämmigheten i lärarnas betygsättning är detta ett negativt resultat ur likvärdighetssynpunkt. Antingen har åtgärderna inte fått tillräckligt genomslag eller så har de motverkats av andra faktorer under perioden.
- En analys av vilka elever det är som får ett högre slutbetyg än provbetyg visar att det framförallt är elever som inte fått ett godkänt provbetyg som får en höjning till G i slutbetyg. I matematik år 2006 fick t ex drygt 70 procent av eleverna med ett icke godkänt provbetyg ändå ett godkänt slutbetyg. Av de elever som hade G i provbetyg var det ungefär 20 procent som fick VG i slutbetyg och av elever med VG i provbetyg fick cirka 16 procent MVG i slutbetyg. Mönstret är likartat i svenska och engelska men respektive andelar är lägre. Dessa siffror är genomsnitt för hela riket. En mer detaljerad analys visar på en stor variation mellan olika skolor. Det finns ett antal skolor där en mycket stor andel (mer än 90 procent) av eleverna som inte nått godkänt på provet ändå får ett godkänt slutbetyg. Det finns samtidigt ett antal skolor där nästan inga elever som inte nått godkänt på provet sedan får ett godkänt slutbetyg.
- Skolor med mycket höga nettoavvikelser sätter generellt högre betyg givet ett visst provbetyg för samtliga provbetygsteg jämfört med skolor med mycket låga nettoavvikelser. Men det som framförallt skiljer sig i betygsättning mellan skolor med mycket höga respektive mycket låga nettoavvikelser är hur de högre betygen VG och MVG sätts. Exempelvis har en matematikelev (se figur 4.3) som går på en skola som tillhör de tio procenten med högst nettoavvikelser ungefär dubbelt så stor sannolikhet att få en höjning från ett

icke godkänt provbetyg till G i slutbetyg jämfört med en elev som går på en skola tillhörande de tio procent med lägst nettoavvikelse. Men sannolikheten för att få VG i slutbetyg, givet G i provbetyg, är 6 gånger större och sannolikheten att få MVG i slutbetyg, givet VG i provbetyg, är hela 11 gånger större i matematik om en elev går i en skola tillhörande de tio procenten med högst nettoavvikelser jämfört med en elev som går i en skola som tillhör de tio procent med lägst nettoavvikelser.

- Hypotesen om att skolor som är mer konkurrensutsatta skulle tendera att sätta högre betyg får inget stöd utifrån analyserna av relationen provbetyg. Detta gäller även när olika varianter på definitionen av en konkurrensutsatt skola testas. Denna hypotes bör dock undersökas mer i detalj innan några definitiva slutsatser dras.
- Utifrån rapportens analyser finns inga belägg för att elever på fristående skolor lättare skulle få högre betyg än elever på kommunala skolor. Det finns fristående skolor som sätter höga slutbetyg i förhållande till provbetygen men det gäller i lika hög utsträckning för kommunala skolor.
- Det finns vissa tendenser till att det kan vara svårare att få bra betyg för elever som går på en skola där den genomsnittliga prestationsnivån är relativt hög. Detta gäller framförallt matematik. I svenska och engelska syns däremot inga tydliga mönster. En möjlig förklaring till mönstret i matematik är att lärarens tolkning av betygsriterierna till viss del skulle vara beroende av skolans elevers allmänna prestationsnivå.

8.3 Förklaringar av resultaten

Vilka förklaringar finns till ovanstående resultat? Rapporten argumenterar för att även om det inte är möjligt att identifiera enskilda skolor och bedöma huruvida de gör en "korrekt" tolkning av betygsriterierna utifrån hur de förhåller sig till provbetygen, så är själva variationen i hur olika skolor sätter betyg i förhållande till provbetygen en *indikator* på hur likvärdigt betygssystemet är. Denna indikator visar på en betydande variation mellan olika skolor. Utifrån denna observation är det mycket tveksamt om det går att hävda att det råder likvärdig betygsättning mellan skolor och även till viss del mellan kommuner.

Varför ingen förändring över tid?

Ett grundläggande resultat är att variationen i nettoavvikelser mellan skolor inte minskat, men heller inte ökat, sedan de första nationella ämnesproven infördes 1998. Rapporten har konsekvent försökt skilja på individnivå och skolnivå. På individnivå har det inte heller skett några förändringar – ungefär 8 av 10 elever får samma slutbetyg som provbetyg sett över hela tidsperioden 1998-2006. Detta är emellertid inte problematiskt sett ur hur proven är tänkta att fungera. Proven ska inte vara examensprov. Men variationen mellan skolor är något helt annat. Avvikelse mellan skolor indikerar att betygsättningen inte sker på ett likvärdigt sätt. Borde man inte förvänta sig att lärarnas samstämmighet i betygsättningen förbättrats med tiden?

När det målrelaterade betygssystemet infördes är det rimligt att tro att lärarna kände sig osäkra i sina tolkningar av de nya målen och betygsriterierna. Med de nationella proven fick lärarna ett stöd genom att mål och betygsriterier konkretiserades. Med tiden kan man förvänta sig att lärarna blivit mer och mer förtrogna med det nya

betygssystemet och att deras tolkningar blivit mer i överensstämmelse med betygskriterierna och därmed även sinsemellan. När det nya betygssystemet infördes var tanken att MVG-kriterier skulle utformas lokalt av lärarna själva. Efter mycket kritik fick Skolverket i uppdrag att utforma nationella MVG-kriterier som började tillämpas år 2000. Redan 1998 utkom dock Skolverkets allmänna råd om MVG-kriterier. Efter att Skolverket genomförde en nationell kvalitetsgranskning av betygsättningen 2000-2001, där stora brister i likvärdigheten påtalades, har ytterligare insatser gjorts. Skolverket gav 2001 ut en av sina hittills mest spridda skrifter: *Bedömning och betygsättning - kommentarer med frågor och svar*.⁷⁴ 2003 utkom allmänna råd och kommentarer för en Likvärdig betygsättning.⁷⁵ Sedan hösten 2004 finns en särskild betygssida på Skolverkets hemsida innehållande bland annat bedömningsexempel för matematik, svenska och engelska i årskurs 9 och andra rådgivande material om betygsättning. Sedan 2003 gör skolverket systematiska inspektioner av alla kommuner och skolor vart sjätte år. Inspektionen uppmärksammar och ställer frågor om den ser stora avvikelser mellan prov och betyg. Sedan 2003, när de nationella proven började samlas in från samtliga skolor, publiceras provresultaten för varje skola på Skolverkets hemsida. Där kan var och en jämföra prov- och betygsresultat på skolnivå.⁷⁶ Detta bidrar till att öka insynen och borde därmed sätta ljuset på skolor som har stora avvikelser. Myndigheten för skolutveckling har också under flera år haft ett särskilt uppdrag och genomfört en rad aktiviteter för att utveckla en likvärdig betygsättning.⁷⁷ Till detta kommer att många kommuner under senare år gjort omfattande insatser för att få till stånd en mer likvärdig betygsättning.⁷⁸

Allt detta sammantaget borde rimligtvis ha haft en positiv effekt på likvärdigheten i betygsättningen och därmed minskat nettoavvikelsena mellan olika skolor. Men så har alltså inte skett. Om man går tillbaka till de tre kategorierna av förklaringar till avvikelser mellan prov och betyg som nämndes i kapitel 1.10 så har staten vidtagit åtgärder inom både kategori 2 och 3. Som nämndes ovan har flera åtgärder vidtagits för att stödja lärarnas gemensamma tolkning av mål och kriterier (kategori 2). Införandet av allmänna råd och utbildningsinspektion har angripit det som nämndes inom kategori 3, att lärare sätter betyg på grunder utanför ramen för betygsbestämmelserna.

Finns det andra händelser som kan tänkas ha motverkat en utveckling mot en mer likvärdig betygsättning? Det har ibland framförts att en ökande konkurrensen mellan skolor, som en följd av det fria skolvalet i kombination med ett ökande antal fristående skolor, skulle ha lett till större avvikelser mellan slutbetyg och provbetyg. Här skulle även publiceringen av skolornas betygsresultat på nätet ha kunnat bidra genom att öka pressen på skolorna att sätta höga betyg som ett sätt att locka elever. Inom ramen för denna rapport har det inte gått att göra några djupgående analyser av

⁷⁴ Skolverket 2001.

⁷⁵ Skolverket 2004e.

⁷⁶ <http://siris.skolverket.se>

⁷⁷ http://www.skolutveckling.se/kunskap_bedomning/

⁷⁸ En del kommunala insatser finns samlade på Myndighetens för skolutvecklings hemsida http://www.skolutveckling.se/kunskap_bedomning/bedomning_betyg/vardkommuner/. Det finns många andra exempel på fleråriga kommunala projekt kring likvärdig betygsättning. Se t ex beskrivning på Stockholms och Göteborgs hemsidor: <http://www.stockholm.se/Extern/Templates/Page.aspx?id=77600> <http://www.goteborg.se/skolutveckling>

eventuella konkurrenseffekter. De analyser som gjorts visar dock inte på några tendenser att skolor i mer konkurrensutsatta områden skulle sätta högre slutbetyg i förhållande till provbetygen än andra skolor. Jämförelsen mellan fristående och kommunala skolor visar inte heller på skillnader i betygsättning relativt de nationella proven.

Det är anmärkningsvärt att det inte finns några indikationer på att likvärdigheten i betygsättningen mellan skolor har förbättrats de senaste åren om man utgår från relationen till provbetygen. Det kan vara så att de insatser som gjorts inte har slagit igenom ordentligt eller så finns det andra faktorer än de som undersökts som motverkat en ökning av lärarnas samstämmighet i betygsättningen.

Vilka förklaringar till skillnader mellan skolor är mest trovärdiga?

Vad finns det då för förklaringar till de stora avvikelserna mellan provresultat och betyg i de olika ämnena och vad är det framförallt som *orsakar* den ganska stora variationen mellan olika skolor? Inom ramen för den här rapporten har det inte varit möjligt att empiriskt pröva alla tänkbara förklaringar som räknades upp i kapitel 1. Det är ändå viktigt att med hjälp av data från denna studie och från andra källor försöka föra en diskussion om rimligheten i de olika förklaringarna. Det är också viktigt att försöka sortera förklaringarna utifrån om de ligger inom ramen för prov- och betygssystemet eller inte.

Förklaringarna kan grovt sett delas upp i tre kategorier (se kap 1.10). Det finns förklaringar till avvikelser som ligger inom ramen för prov- och betygssystemet, varav vissa påverkar och andra inte påverkar likvärdigheten i betygsättningen. Ett exempel på den första kategorin är stödinsatser gjorda efter proven. Om en skola varit mer lyckosam än en annan med de stödinsatser som satts in efter de nationella proven vilket sedan lett till faktiska förbättringar och högre betyg är det inget som påverkar likvärdigheten i betygsättningen⁷⁹. Ett exempel på den andra kategorin är om en skola gör en annorlunda tolkning av betygskriterierna jämfört med en annan skola. Att lärare på olika skolor gör olika tolkningar ligger också inom ramen för betygssystemet men leder ändå till att likvärdigheten i betygsättningen påverkas negativt. Den tredje kategorin handlar om förklaringar som *inte* ligger inom ramen för betygssystemet och som påverkar likvärdigheten i betygsättningen. Ett sådant exempel är om en skola tillämpar andra betygskriterier än de som står i kursplanerna som grund för betygsättningen, medan en annan skola håller sig till de nationella kriterierna.

Förklaringar som bedöms som mindre troliga

Nedan följer först de förklaringar som bedöms som mindre troliga.

Förklaring (d): Eleverna har arbetat med ett annat stoff än det provet behandlar

Detta är mer av en teoretisk invändning och har redan avfärdats som mindre trolig i en tidigare rapport från Skolverket⁸⁰. Enligt ansvarsfördelningen i skolan är det lärare och elever som ska välja stoff och arbetsmetoder. Om valet av stoff är ett annat än det som ingår i nationella provet skulle följderna kunna bli att den på provet uppvisade måluppfyllelsen blir sämre än den måluppfyllelse eleverna uppvisat i klassrumsarbetet. Men frågan är hur vanligt det är att lärare och elever väljer stoff som avviker från det gängse förekommande. Mycket tyder på att läromedlens

⁷⁹ Däremot kan man diskutera om det påverkar likvärdighet i tillgång till möjligheter.

⁸⁰ Skolverket 2004a s.73.

ställning är stark när det gäller valet av innehåll i undervisningen. I en nyligen publicerad skolverksrapport dras följande slutsats:

Flertalet lärare ser också läroboken som en garanti för att alla kursplanens mål och moment ingår i undervisningen och att eleverna är förberedda inför de nationella ämnesproven. Tidigare forskning drar samma slutsatser kring läroböckers funktioner.⁸¹

Läromedelsförfattare brukar av tradition studera de nationella proven och deras innehåll. Detta får rimligen genomslag i läromedlen och stoffvalet torde därmed i praktiken inte skilja sig särskilt mycket mellan olika skolor. Dessutom är i synnerhet svenska och engelska ämnen med påtaglig progression som inte är särskilt stoffberoende. Matematik är antagligen mer stoffberoende och därför skulle man kunna tänka sig att val av ett annat stoff skulle kunna förklara att just matematiken har störst nettoavvikelser. Men å andra sidan är matematik det ämne av de tre där kursplanerna är mest konkreta när det gäller att ange innehåll i kurserna. Dessutom förefaller det osannolikt att lärare på olika skolor skulle välja ett radikalt annorlunda stoff än sådant som man tror kommer att behandlas på de nationella proven. Även om lärarna inte kan veta exakt vilket stoff som de nationella proven tar upp förefaller det rimligt att tro att man ändå på något sätt försöker anpassa sig till proven. Förklaringen är därmed inte särskilt trovärdig för att förklara variationen mellan skolor.

Förklaring (e) Lärare planerar undervisningsmomenten olika

Förklaringen går ut på att det kan finnas skolor som väntar med att gå igenom eller repetera sådana moment som behandlades på provet till efter att det nationella provet hade gjorts. Eftersom de nationella proven görs på vårterminen i årskurs 9 förefaller det inte vara en rimlig förklaring. Mer rimligt är i så fall att moment som är centrala i ett nationellt prov behandlas vid olika tidpunkter på olika skolor. Var det länge sedan eleverna läste ett moment som är centralt på proven så kan resultaten bli sämre. Även om lärare och elever inte vet vilket stoff som kommer på de nationella proven finns det klara bevis för att en stor majoritet lärare och elever ändå försöker förbereda sig inför de nationella proven genom att repetera sådana uppgifter som brukar komma på proven. Man går igenom tidigare prov och tar del av det förberedelsematerial (endast svenska) som sänds ut innan proven.⁸²

Förklaring (a) Läraren har att ta hänsyn till fler mål än de som ingår i nationella proven.

De nationella proven har en begränsning i och med att de endast kan omfatta ett visst antal uppgifter på de sammanlagt 4-5 timmar de olika delproven pågår. De som utformar proven måste av det skälet prioritera mellan vilka mål man prövar och *hur mycket* man prövar ett visst mål. Dessutom finns den begränsningen att proven till största delen är skriftliga⁸³. Det gör att det kan vara svårt att pröva vissa mål. Vilka kursplanemål som täcks in av de nationella proven och *hur mycket* olika mål täcks av proven redovisas för lärarna i lärarhandledningar och bedömningsanvisningar till de nationella proven. I Bedömningsanvisningarna för ämnesprovet i svenska 2007 står exempelvis:

Provresultaten är en del av lärarens underlag för betygsättning. Allt har inte kunnat prövas i ämnesprovet (till exempel annat litterärt kunnande än vad provet exemplifierat). Vissa förmågor (till exempel de muntliga) har en relativt liten del i hela provet.

⁸¹ Skolverket 2006d s.92

⁸² Skolverket 2004a s.38ff, Skolverket 2004b s.33ff, Skolverket 2006a, Skolverket 2006c.

⁸³ I alla tre provämnerna ingår dock en muntlig del.

I engelska framgår i lärarinformationen till proven precis vilka mål i kursplanen som provet prövar. Ett par uppnåendemål prövas ”i viss mån” och tre mål inte alls: Kunna använda hjälpmedel, Kunna reflektera över sitt lärande och Kunna planera och genomföra arbetsuppgifter. Ungefär samma förhållande gäller matematikprovet. Slutsatsen är alltså att de nationella proven prövar de flesta målen i kursplanen. Men vissa mål prövas bara i begränsad utsträckning och några inte alls.

Till de mål som prövas i nationella prov ska alltså ytterligare läggas de mål som proven inte prövar. Om man då till de mål som de nationella proven prövar lägger de ytterligare mål som läraren har att pröva så bör - om man antar att några elever som klarat alla mål på provet inte klarar något eller några av de mål läraren har att pröva - andelen elever som totalt sett *inte* nått samtliga mål vara *högre* i slutbetyget än på det nationella provet. Man kan alltså knappast förklara de högre slutbetygen med att läraren har att ta hänsyn till fler mål än vad som ingår i det nationella provet eftersom det snarare borde höja andelen elever som inte nått samtliga mål. Förklaringen till att slutbetygen är högre än provbetygen måste alltså vara en annan. Det är inte heller en trolig förklaring till skillnader mellan skolor. All erfarenhet tyder på att lärarna prioriterar de mål som prövas i proven.⁸⁴ Dessutom är det inte troligt att eleverna på en hel skola skulle vara bättre eller sämre på de mål som i mindre utsträckning prövas i de nationella proven, t ex muntlig förmåga.

Förklaring (b) Lärarna bedömer proven olika.

Det går inte att säga så mycket om hur väl lärarnas bedömningar av provresultaten stämmer överens med det centrala bedömningsunderlaget och inte heller hur väl bedömningen stämmer överens lärare emellan. Några mer omfattande systematiska studier av hur lärarna bedömer proven har inte gjorts. De begränsade undersökningar som finns räcker inte för en generell bedömning av likvärdigheten i bedömningen av årskurs 9-proven.⁸⁵ Undersökningarna för ämnesproven årskurs 9 (från år 2001 och 2002) visar inte på några dramatiska avvikelser i bedömningen. Och sannolikt har läget förbättrats med tiden.⁸⁶ Det finns tecken som tyder på att även sambandet mellan prov mellan lärare på en skola och mellan skolor har ökat med åren. Samtidigt går det att se att det bildas tydliga ”toppar” runt betygsgränserna, särskilt vid G-gränsen, vilket tyder på att elever tilldelas en välvillig poängsättning för att nå över ett visst betygssteg.⁸⁷ På det hela taget är det dock inte sannolikt att bristande överensstämmelse när det gäller bedömningen skulle kunna ha någon avgörande betydelse för att förklara avvikelser till slutbetygen mellan skolor och klasser. Om det finns ett visst tolkningsutrymme i proven så är det troligast att lärarens tolkning av prov och betyg samvarierar, dvs. om läraren gör en relativt generös tolkning vid bedömning av provet så är det troligt att läraren även gör en relativt generös tolkning vid betygsättningen. Om detta är sant så är variationen i lärarens betygsättning *underskattad* utifrån en analys av relationen provbetyg – slutbetyg.

⁸⁴ Se förklaring (c) ovan.

⁸⁵ För årskurs 9 finns bara två undersökningar: Kjellström och Olofsson 2004 (matematik) och Ohlsson-Wahlsten 2002 (engelska). För gymnasiet finns (i detta avseende) fyra mindre studier. En studie på matematik B-provet 2002: Boesen 2004. En studie på matematik A-provet 2001: Olofsson 2006. Två studier på prov från 1997: Lindström 1998 (matematik E), Östlund-Stjärnegårdh 2002 (svenska B). Se även Skolverket 2005f s.93ff och Korp 2006 s.200ff för en intervjustudie om bedömningen av gymnasieproven.

⁸⁶ Olofsson 2006.

⁸⁷ Skolverket 2004a s.76, 83. Se även Korp 2006 s.205.

Förklaring (b) Läraren har ett bredare underlag än provresultatet för att bedöma elevernas måluppfyllelse

En inte ovanlig situation vid betygsättningen borde vara att läraren konstaterar att en elev inte klarade tillräckligt många mål (i vissa fall poäng) på provet för att bli godkänd, men att läraren i andra sammanhang sett att eleven uppnått målen och att eleven därför i slutändan kan godkännas om han eller hon dessutom uppfyller de mål läraren har att pröva. Det är då inte det förhållandet att läraren ska pröva fler mål som är förklaringen utan förhållandet att läraren har ett mycket *bredare* underlag för att bedöma de mål som proven behandlar. Denna förklaring förefaller rimlig men hur stark kan den vara för att förklara skillnader mellan skolor? Det är rimligt att betrakta det nationella provet som ett i raden av olika bedömningsunderlag även om det nationella provet väger tungt. En elev kan naturligtvis ha en mindre lyckad dag vid de nationella provtillfällena. Om eleven i andra sammanhang visat på kunskapskvaliteter som motsvarar ett visst betyg kan detta motivera läraren att sätta ett högre betyg än vad eleven fått på det nationella provet. Omvänt skulle man kunna tänka sig att läraren ansett att en elev tidigare inte nått de mål som det nationella provet prövar men att han eller hon blir godkänd på det nationella provet och därigenom visar att målen nu är nådda. Det skulle kunna förklara den vanliga kommentaren att nationella proven endast kan höja, inte sänka betyget. De nationella proven kan ses som ytterligare en chans att visa att man har uppnått vissa kunskapskvaliteter. Sammantaget kan alltså förklaringen vara relevant för en enskild elevs avvikelse mellan prov och betyg.

För att förklara skillnader mellan skolor är dock förklaringen mindre relevant. Det förutsätter att lärare på olika skolor har olika synsätt på vilken vikt de lägger vid provresultatet i förhållande till andra bedömningsunderlag. Enkätundersökningar tyder på att lärare har en ganska samstämmig bild av vilken vikt man lägger vid de nationella proven vid betygsättningen.⁸⁸

Förklaring (k) Konkurrensen ger betygsinflation och variation mellan skolor

Denna förklaring har delvis testats empirisk i rapporten (kapitel 6) och visade sig inte ha någon förklaringskraft utifrån de analyser som genomfördes. Det är kanske för tidigt att hävda att det inte finns några som helst konkurrenseffekter på betygsättningen utifrån de relativt begränsade analyser som gjorts i denna rapport, men även om konkurrenseffekter skulle existera är det mindre troligt att de skulle kunna förklara någon större del av den variation som finns, utifrån den empiri som presenterats.

Förklaringar som bedöms som mer troliga

Det återstår nu fem förklaringar som alla kan ha betydelse för att förklara skillnaderna i betygsättningen i förhållande till provresultaten mellan olika skolor och kanske även mellan olika ämnen. Alla dessa förklaringar har säkerligen en betydelse men det går inte att på säkra grunder avgöra *hur* starka förklaringarna är. Det är ändå viktigt att försöka resonera kring hur betydelsefull respektive förklaring kan vara. Inte minst eftersom vissa förklaringar ligger inom ramen för bestämmelserna om betygsättning medan andra inte gör det. Det är också viktigt att diskussionen om

⁸⁸ Skolverket 2004a s. 35f.

betygsättningen i förhållande till de nationella proven förs utifrån de förklaringar som är mest relevanta och som man då kanske kan åtgärdas.

Förklaring (c) Särskilda insatser sätts in för elever som inte klarat proven.

Denna förklaring ligger inom ramen för prov- och betygssystemet och betyder inte att likvärdigheten mellan elever riskeras. Det är rimligt att tänka sig att elever som inte nått målen får mer stöd än de som redan fått ett godkänt provbetyg. Att inte bli godkänd i ett behörighetsämne innebär ju att man inte kommer in på ett nationellt program på gymnasiet. Men hur stark denna förklaring kan tänkas vara är svårt att uttala sig om. Det finns dock förhållanden som talar emot att denna förklaring skulle vara stark. Tidsfaktorn är viktig. Alla prov genomförs på våren i årskurs 9. Det nationella provet i svenska görs före sportlovet, engelskprovet före påsklovet och matematikprovet efter påsklovet.⁸⁹ Det finns inte många veckor för stödinsatser efter proven. Att analysen visat att det är i matematik som de klart största höjningarna görs upp till ett godkänt slutbetyg gör att förklaringen blir ännu mer tveksam. Tiden för att göra stödinsatser efter det nationella provet är betydligt kortare i matematik i jämförelse med engelska och framförallt svenska.⁹⁰ Det är knappast rimligt att tänka sig att det sker någon mer avgörande kunskapsutveckling dessa sista veckor i grundskolan som inte skett tidigare. Man måste också ta i beaktande att det går ytterligare tid innan lärarna har rättat elevernas prov.

Men det finns andra omständigheter som talar för denna förklaring. Det kan vara så att matematikämnet karaktär gör det lättare att göra stödinsatser som går att bedöma för lärarna. Matematikkursplanens uppnåendemål (som ligger till grund för betyget godkänt) består mer av ”moment” än i svenska och engelska. Om det exempelvis visade sig att en elev i det nationella provet klarade godkänt på allt utom målet att ”kunna använda begreppet sannolikhet i enkla slumpvariationer” kanske en särskild insats efter provet kan leda till förbättringar. Uppnåendemålen i svenska och engelska är inte så avgränsade och specificerade och det kan därför vara svårare att bedöma om en elev förbättrat sig på kort tid. Giltigheten i denna typ av förklaring går inte att avgöra genom denna studie. För detta krävs andra mer kvalitativa undersökningar. Det förefaller ändå osannolikt att stödinsatserna skulle kunna få sådana effekter att de cirka 12 procent som inte nått målen i matematik på provet reduceras till 4 procent i slutbetygen. Att man på olika skolor skulle vara olika framgångsrik när det gäller stödinsatser går förstås också att föreställa sig. Olika lyckosamma stödinsatser kan vara en förklaring till variationen mellan skolor när det gäller nettoavvikelser. Men en sådan förklaring kan ändå inte vara en särskilt stark faktor för att förklara de stora skillnaderna mellan skolor. Det är inte enbart i höjningarna upp till godkänt som variationen finns och det förefaller inte troligt att stödinsatserna skulle vara så olika lyckosamma för eleverna vid olika skolor.

Förklaring (j) Behörighetskraven gör att lärare inte gärna vill underkänna elever.

Detta är en av två rimliga förklaringar som inte kan anses ligga inom ramen för bestämmelserna om betygsättning. Elever får godkänt betyg trots att läraren egentligen bedömer att eleven inte når samtliga uppnåendemål i kursplanerna. Om det förhåller sig på det viset rubbas likvärdigheten mellan elever. Elever som egentligen inte har kunskaper för att komma in på gymnasiet konkurrerar då om

⁸⁹ Provdatum <http://www.skolverket.se/sb/d/310>

⁹⁰ Som nämnts i kap 1.7 ingår *inte* i databasen betyg satta efter sommarkurser och efterföljande provningar.

platser med andra som godkänts på ”legitima” grunder. I denna studie finns ingen ny empiri som kan belysa denna förklaring. Men i den intervjustudie som Skolverket gjorde med lärare om grundskolans nationella prov finns en diskussion som anknyter till denna förklaring även om den knappast ger den ett starkt stöd. Intervjuerna antyder i vart fall att lärare hellre ”friar än faller” elever.⁹¹ En studie av lokala terminsbetyg indikerar också att det kan finnas skäl att misstänka att skolorna kan sätta lägre krav för godkänt än vad uppnåendemålen och därmed provbetygen anger.⁹² I en tidigare rapport från Skolverket⁹³ står följande om denna förklaring:

Även information som på annat sätt kommer till Skolverket tyder på att detta är en förklaring som det inte går att bortse ifrån. Det händer att lärare ringer och berättar om sina våndor inför betygssättningen. Man känner sig ensam och utlämnad i en svår process. Även om man egentligen tycker att eleven inte nått alla mål så trycker eleven på om att bli godkänd, och elevens föräldrar gör samma sak. Ofta har inte heller skolledningen något intresse av att elever inte blir godkända utan ser helst att andelen elever som inte når målen är så liten som möjligt. Det ser bättre ut i både den kommunala och nationella statistiken. Även studievägledare uppges ibland uttrycka dubier mot att elever blir underkända eftersom de vet att det individuella programmet i många fall inte fungerar särskilt bra. Det är i sådana fall förståeligt, om än inte försvarbart, att läraren ger vika och sätter det godkända betyg alla vill ha.

Även om forskningen inte är övertygande och övrig information bygger på mer anekdotisk information går det inte att avfärda denna förklaring. Den här typen av beteende vid betygssättningen är självklart ingenting lärare skyltar med öppet och är därför svår att bedöma.⁹⁴ Med tanke på de stora skillnader som finns mellan olika skolor när det gäller hur många elever som ”höjs” till ett godkänt betyg och med tanke på den kraftiga höjningen i matematik så är det troligt att denna förklaring kan ha en relevant betydelse.

Förklaring (i) Betygen sätts på andra grunder än provbetygen.

Detta är den andra mer rimliga förklaringen som inte ligger inom ramen för prov-, och betygssystemet. Provbetygen sätts utifrån mål och betygskriterier i kursplanerna. Det samma ska ju också gälla för slutbetygen. Beror avvikelser på att betygen sätts på andra grunder än de som anges i förordningar och kursplaner är risken stor att betygen inte blir likvärdiga. Särskilt om dessa grunder varierar mellan olika lärare och skolor. Med hjälp av dataunderlaget från denna rapport har det inte varit möjligt att pröva denna förklaring. Men på denna punkt finns det en del annan empiri. De få och empiriskt begränsade försök som gjorts att analysera lärares betygssättning visar att det tycks förekomma att även andra faktorer än betygskriterierna ligger till grund för lärares betygssättning. Utifrån enkät- och intervjuundersökningar går det att dra slutsatsen att lärarna har en mycket hög tilltro till att de nationella proven täcker mål och kriterier i kursplanerna och proven utgår från läroplanens och kursplanernas kunskapssyn. Det är också något som lärarna uppskattar.⁹⁵ Men det innebär inte att det är givet att alla lärare faktiskt utgår från enbart kursplanens mål och kriterier vid betygssättningen.

⁹¹ Skolverket 2004b s.47.

⁹² Tholin 2006.

⁹³ Skolverket 2004a

⁹⁴ Ett exempel finns i en insändare av en anonym lärare i tidningen Skolvärlden 2004.

⁹⁵ Skolverket 2004a kap 2. Skolverket 2004b s.141ff. Skolverket 2006a, 2006b och 2006c.

Det finns två rapporter som jämför de nationella proven med lärarnas egna sätt att bedöma eleverna.⁹⁶ Lars Naeslund har dragit slutsatsen att:

*Kommentarer från lärare som är positiva till de nationella provens sammansättning och egenskaper ger stöd åt tolkningen att undervisningen och den lokala kunskapsbedömningen ingalunda behöver vara mer konservativ än de nationella proven.*⁹⁷

Naeslunds relativt positiva slutsats om överensstämmelsen mellan kunskapssyner i de nationella proven och i skolpraktiken är baserad på ett fåtal intervjuer med lärare och elever och inte på faktiska analyser av lärarnas egna prov. Det har däremot Jesper Boesen gjort, dock enbart för gymnasiet.⁹⁸ Genom att jämföra de nationella proven i matematik med 52 lärarkonstruerade prov framkommer en mer problematisk bild ur likvärdighetssynpunkt. En del lärarprov visade sig t ex utelämna mål som är formulerade i kursplanerna. Ett av bedömningsområdena som är angivna i kursplanerna: elevernas kommunikativa förmåga i matematik, återfanns i mycket få uppgifter i lärarnas egna prov. Det förekom också sällsynt med uppgifter som prövade andra kompetenser som återfinns i kursplanerna och även i de nationella proven som t ex begreppsförståelse, modellering, problemlösning mm. Skolverket har i sina nationella kvalitetsgranskningar av betygsättningen i början av 2000-talet kommit fram till liknande slutsatser som Boesen - att det inte sällan finns en svag koppling mellan det som provas i lärargjorda prov och kursplanernas mål och kriterier.⁹⁹

Även dessa undersökningar har dock sina begränsningar. Den centrala slutsatsen i detta sammanhang är att det ändå är rimligt att tro att det finns vissa skillnader, i alla fall på en del skolor, mellan vad som provas i de nationella proven och vad som sedan ligger till grund för lärarnas egenkonstruerade prov. *Hur* stora dessa skillnader är kan man inte ha någon säker uppfattning om.

Det finns också andra studier som pekar på att andra "icke legitima" betygsgrunder kan ha betydelse vid betygsättningen. Några studier pekar på att det förekommer att elevernas allmänna *beteende och attityder* vägs in i betygsättningen.¹⁰⁰ T ex är närvaro, flit och ambition ganska vanligt förekommande i skolorna lokala terminskriterier.¹⁰¹ I Jörgen Tholins avhandling om lokala terminskriterier fanns han att det ämne som också har nationella prov, engelska, också är det ämne där de lokala kriterierna i högst utsträckning överensstämmer med de nationella målen och kriterierna (de andra två ämnena var idrott och hälsa och kemi). Men samtidigt fann Tholin en stark fokusering på grammatik i de lokala engelskakriterierna som inte har sin motsvarighet i den nationella kursplanen och han fann också att variationen mellan olika skolors lokala kriterier är stor även i engelska.¹⁰²

I en sammanställning av alla Skolverkets inspektioner under 2004¹⁰³ dras följande slutsats:

⁹⁶ Skolverket 2004b. Boesen 2006.

⁹⁷ Skolverket 2004b. s.114

⁹⁸ Boesen 2006.

⁹⁹ Skolverket 2000 s.26ff, Skolverkets 2002 s.20ff.

¹⁰⁰ Skolverket 2000, 2005h och 2005i. Tholin 2006. Selghed 2004. Korp 2006.

¹⁰¹ Tholin 2006 s. 169. Skolverket 2000 s.26.

¹⁰² Tholin 2006 s.156ff.

¹⁰³ Skolverket 2005e s.20.

Inspektörerna har bedömt att närmare vart tredje rektorsområde som har elever som får betyg i grundskolan behöver utveckla sitt arbete med bedömning och betygsättning. Beslut tas inte alltid med stöd av underlag som är allsidigt och tar hänsyn till kunskapens olika former. Betygsättningen bedöms inte alltid följa nationella bestämmelser.

Det förefaller rimligt att tro att denna förklaring kan ha betydelse för att nettoavvikelserna skiljer sig åt mellan olika skolor, *hur* stor betydelse är däremot omöjligt att uttala sig om.

Förklaring (f) Lärare på olika skolor gör olika tolkningar av mål och betygskriterier, dels sinsemellan och dels i förhållande till de nationella proven.

Det här är den kanske mest centrala förklaringen som kan påverka likvärdigheten i betygsättningen även om den ligger inom ramen för betygsbestämmelserna. Om det skulle vara så att olika tolkningar av mål och kriterier helt förklarade de varierade nettoavvikelserna mellan skolorna så ligger detta i så fall inom ramen för de bestämmelser som finns för betygsättningen. Lärarna är suveräna i sin betygsättning och det finns ingen som kan överpröva lärarnas tolkning av nationella mål och kriterier så länge som de håller sig till dessa och inget annat.

Att sätta betyg innebär att tolka de nationella kursmålen och betygskriterierna och att ställa dem i relation till olika elevprestationer. Enligt förordningarna ska uppnåendemålen och kriterierna finnas "som stöd för betygsättningen".¹⁰⁴ Betygen ska sättas "med hjälp av" uppnåendemål och kriterier.¹⁰⁵ Till hjälp för bedömningen finns också kursplaneavsnittet *Bedömningens inriktning* där ämnets centrala kunskapskvaliteter klargörs. Kursplanernas uppnåendemål och betygskriterier är dock inte utformade så att de är självklara och entydiga för alla läsare. Uppnåendemål och kriterier är utformade på ett sådant sätt att de förutsätter en professionell tolkning av lärarna. Ett exempel: Ett uppnåendemål i svenska för årskurs 9 är att: "kunna läsa, reflektera över och sätta in i ett sammanhang några skönlitterära verk och författarskap med betydelse för människors sätt att leva och tänka".¹⁰⁶ En person som inte är svensklärare kan ställa sig frågan vad som här menas med "kunna", "reflektera", "betydelse", "några" och "sammanhang"? Hur ska man veta vilka verk och författarskap som har denna betydelse? Men nu är betygskriterierna skrivna till en profession: till lärarna, och inte till allmänheten eller till eleverna. Systemet förutsätter att landets svensklärare utifrån sin profession har en gemensam uppfattning om de begrepp som används i betygskriterierna och att de därför också kan bedöma eleverna på liknande grunder.¹⁰⁷ Bedömning av kunskaper förutsätter en hög grad av professionell kompetens vilket också uppmärksamats av svenska forskare och lärarutbildare i stigande utsträckning.¹⁰⁸

Det finns ingen garanti för att det bedömningsunderlag som ligger till grund för provbetygen överensstämmer med alla lärares tolkning vid betygsättningen. Det är tänkbart att den *samlade* lärarkårens tolkning av uppnåendemål och betygskriterier skiljer sig från den tolkning som ligger till grund för proven och att det är i

¹⁰⁴ Läroplan för det obligatoriska skolväsendet (Lpo 94) punkt 2.7.

¹⁰⁵ Grundskoleförordningen 7 kap 8 §.

¹⁰⁶ Se Kroksmark 2004 och Måhl 2005a för en utförlig diskussion om möjligheten och rimligheten för lärare att göra gemensamma tolkningar av mål och kriterier.

¹⁰⁷ Måhl 2005b s.63ff.

¹⁰⁸ Två tydliga exempel på det ökade intresset för kunskapsbedömning är antologin *Pedagogisk bedömning*, Lindström och Lindberg 2005, och att en tidskrift som Grundskoletidningen Skolbarn har starkt fokus på kunskapsbedömning under senare år. www.fortbild.se.

matematik den tolkningen skiljer sig åt mest. Det är i matematik, där skillnaden är som störst på nationell nivå (se figur 2.2), som måluppfyllelsen är som lägst, vilket möjligen kan leda till att lärarna uppfattar kravnivåerna som för höga i matematik. Denna förklaring får dock inget stöd i de lärarenkäter som skickas in av lärare i samband med de nationella proven i matematik.¹⁰⁹ På frågorna om kravgränserna för de olika provbetygen svarade nästan 90 procent av matematiklärarna att de var ”lagom” både vad det gäller Godkänt, Väl godkänt och Mycket väl godkänt.¹¹⁰ Av dem som inte ansåg att G-gränsen låg på en lagom nivå var det en något *större* andel ansåg att gränsen var för låg än som ansåg att den var för hög. Av dem som inte ansåg att VG- och MVG-gränserna var lagom var det däremot fler som ansåg att kraven var för höga än för låga. Lärarenkäten i engelska har enbart frågor om betygsgränserna för varje delprov men sammantaget visar svaren på snarlika resultat som i matematik.¹¹¹ I lärarenkäten i svenska finns inte direkta frågor om betygsgränserna i proven men lärarsvaren visar ändå, på samma sätt som i matematik och engelska, att det inte finns något starkt missnöje med att proven skulle ligga på en för hög nivå i förhållande till de angivna betygsgränserna.¹¹² Den intervjustudie med lärare om de nationella proven i årskurs 9 som Skolverket låtit genomföra pekar i samma riktning. De lärare som intervjuats tycker inte att kraven för godkänd på proven är för höga, snarare tvärtom.¹¹³ Allt tyder alltså på att i samtliga ämnen så är en mycket stor majoritet av lärarna nöjda med kravgränserna för de olika betygsstegen i de nationella proven. Om det finns någon tendens är det snarast att en del lärare uppfattar att vissa delprov ibland har lite för lågt ställda gränser för vissa betyg, även om motsatsen också förekommer för vissa delprov vissa år.

Hur är det då med lärarnas syn på betygsgränserna i slutbetygen? Det finns inte någon empiri för alla tre provämnena. Men i en enkätundersökning från 2004 ställdes en fråga till lärare om de olika betygsstegen i ett av provämnena, svenska, samt i ämnena fysik och idrott och hälsa.¹¹⁴ Resultatet är snarlikt lärarnas uppfattning om betygskraven i de nationella proven. En stor majoritet, cirka 80-90 procent, ansåg att nivån för de olika betygen låg på en ”lagom” nivå. Tendensen är (även om dataunderlaget här är litet) att av den andel som inte tyckte att betygsnivåerna var lagom så ville fler lärare höja snarare än sänka kraven för ett godkänt betyg och sänka kraven för VG och MVG. Det gäller även lärare i svenska. Sammantaget kan konstateras att en stor majoritet av lärarna själva uppger att de är nöjda med de anvisade kravnivåerna för betygsstegen både för de nationella proven och för slutbetygen. Går det därmed att dra slutsatsen att en annorlunda tolkning av uppnåendemålen och betygskriterierna inte skulle ligga bakom avvikelserna mellan prov och betyg i de olika ämnena? Det är inte helt självklart.

Även om lärarna säger sig vara nöjda med betygsgränserna i prov och kursplaner innebär det inte att lärarnas *faktiska* tolkning är samstämmig. I lärarenkäten från 2004 framkommer att lärarna i stor utsträckning utgår från de nationella betygskriterierna när de sätter betyg och de förefaller i hög utsträckning ha uppfattningen att de själva

¹⁰⁹ Skolverket 2006a

¹¹⁰ Från och med 1 april 2007 är betygsbeteckningen Godkän(t) inte Godkän(d). Grundskoleförordningen 7 kap 6 §.

¹¹¹ Skolverket 2006c.

¹¹² Skolverket 2006b.

¹¹³ Skolverket 2004b s.52 och 142.

¹¹⁴ Bearbetning av data från Skolverket 2005c – enbart lärare som angett att de satt betyg i ämnet.

sätter rättvisa betyg.¹¹⁵ Lärarna enkätsvar visar samtidigt att de upplever att tolkningen av kursplanerna är långt ifrån oproblematisk. Nästan hälften av svensklärarna svarar i enkäten att målen är för otydliga, cirka en fjärdedel menar att flera begrepp som används i kursplanen är svåra att förstå och ungefär fyra av fem svensklärare tycker att ett införande av mer preciserade nationella betygskriterier är ”mycket önskvärt” eller ”värt att överväga”.¹¹⁶ Dessutom svarar över hälften av svensklärarna att det är stora eller vissa problem att få till stånd en likvärdig betygsättning i landet. 3 procent svarar ”nej” på frågan och resten svarar ”vet ej”. Det skulle kunna betyda att lärarna inte tror att man gör samma tolkningar av kriterierna på alla skolor. Inte heller inom kommunen anser svensklärarna att förutsättningarna för en likvärdig betygsättning är god. Svarsmönstret är ungefär likadant som när det gäller den nationella nivån. På skolnivå anser dock svensklärarna att förutsättningarna för en likvärdig betygsättning är bättre. Bara någon enstaka procent av svensklärarna anser det vara ”stora problem” att få till stånd en likvärdig betygsättning inom den egna skolan. Men däremot är det lika många som svarar: ”vissa problem” som ”nej” när det gäller likvärdig betygsättning på skolnivån. Man skulle kunna förvänta sig att svensklärarna, som ju har nationella prov, skulle vara mer tillfreds med likvärdigheten i betygsättningen jämfört med fysik- och idrottslärarna. Men så är inte fallet. Svarsmönstren på ovanstående frågor är i stort sett likartade i de andra två ämnena.

En olikartad tolkning av mål och kriterier skulle därmed ändå kunna ha en betydelse för att förklara avvikelserna mellan prov och betyg mellan olika skolor. Även om lärarna uttrycker att kravnivåerna för betygsstegen ligger på en lagom nivå både för de nationella proven och för slutbetygen innebär det inte att lärarna i praktiken gör samma tolkning när de sätter slutbetygen. Det finns flera studier som visar att lärarnas tolkningar av mål och betygskriterier sällan prövas mot andra lärares tolkningar utöver dem på den egna skolan.¹¹⁷ Hur lika eller olika tolkningarna är finns det dock inte mycket kunskap om ännu.¹¹⁸ Det går inte att bedöma exakt hur stark denna förklaring är för nettoavvikelsena mellan olika skolor men att den sannolikt har betydelse råder knappast någon tvekan om. Tolkningsförklaringen ges dessutom empiriskt stöd av de jämförelser som gjorts med andra ”oberoende” kunskapsmätt som elevernas gymnasiebetyg och skolors resultat enligt PISA (se avsnitt 4.3). Analyserna av gymnasiebetygen visade att elever från skolor med höga nettoavvikelser klarar sig relativt sämre på gymnasiet jämfört med elever från skolor med låga nettoavvikelser, givet samma slutbetyg från årskurs 9. Det tyder på att skolor med höga respektive låga nettoavvikelser inte bedömer eleverna efter samma måttstock.

¹¹⁵ Skolverket 2005c.

¹¹⁶ Skolverket 2005c.

¹¹⁷ Skolverket 2004f s.116. Skolverket 2004a s.97, Skolverket 2005c s.16ff och Skolverket 2000.

¹¹⁸ Forskningen har inte ägnat detta problem något större intresse. Egentligen finns ingen empirisk svensk studie av lärares tolkning av mål och kriterier. Se forskningsöversikten i Tholin 2006 kapitel 2. Tholin har själv studerat lokala terminskriterier i grundskolan och i dessa funnit en betydande variation i uttolkningen mellan olika skolor. Terminskriterier som mått på lärarnas uttolkning av mål och kriterier är dock problematiska eftersom relationen till de nationella målen och kriterierna i årskurs 9 inte är givna. Inte heller Skolverket har gjort några systematiska studier av hur lärare tolkar mål och kriterier. I Skolverket 2000 berörs dock frågan på ett allmänt plan och i Skolverkets Nationella utvärdering 2003 finns ett par enkätfrågor till svensklärare som berör frågan om tolkning av betyget godkänt, Skolverket 2005g s.86ff.

En variant och fördjupning av ovanstående förklaring är (g) *Lärare tolkar kursplanerna relativt elevgruppens förutsättningar*. Denna förklaring går ett steg längre än förklaringen om att lärare tolkar betygskriterier olika då den förklarar *på vilket sätt* eller varför lärare gör olika tolkningar. För att illustrera så kan man åter använda exemplet (se ovan) från uppnåendemålet i svenska för årskurs 9 att: ”kunna läsa, reflektera över och sätta in i ett sammanhang några skönlitterära verk och författarskap med betydelse för människors sätt att leva och tänka”. Det går att tänka sig att det finns lärare som gör tolkningar av det uppnåendemålet i förhållande till den elevgrupp man är van att arbeta med genom att kraven på t ex *vilka* skönlitterära verk man ska kunna läsa och reflektera över skiljer sig åt mellan olika skolor. På en skola där eleverna har en vana hemifrån att läsa mycket böcker kanske lärarna menar att eleverna för att få godkänt betyg ska klara att läsa tjocka vuxenböcker och reflektera över verk av August Strindberg och Kerstin Ekman. På en skola med elever som har mindre utvecklade läsvanor hemifrån kanske lärarna tolkar samma mål som att eleverna ska klara att läsa och reflektera över kortare utdrag ur böcker skrivna för ungdomar.¹¹⁹ En orsak till de skillnader som finns i lärares tolkningar kan alltså ha att göra med att lärares tolkningar påverkas av elevgruppens generella prestationsförmåga. Förklaringen ligger (precis som (f)) inom ramen för bestämmelserna om betygsättning eftersom det handlar om en tolkning av kriterierna.

Rapportens analyser (kapitel 7) visade att det finns tendenser i matematik (men inte i svenska och engelska) till att det kan vara svårare att få bra betyg för elever som går på en skola där den genomsnittliga prestationsnivån är relativt hög. Det kan betyda att lärares tolkning av betygskriterierna till viss del är beroende av elevomgivningens allmänna prestationsnivå. Denna tolkning får ett visst stöd i en intervjustudie med gymnasielärare som undervisar i kärnämnen matematik, svenska och engelska. Där sammanfattas resultaten på följande vis:

*Vissa grupper av elever på yrkesförberedande program tenderar att gynnas i betygsättningen, genom att lägre krav ställs på dem för att de ska få betyget Godkänd (möjligen också VG och MVG, i den mån som lärarna tenderar att tolka kursplaner och betygskriterier i relation till elevgruppens genomsnittliga kunskapsnivå).*¹²⁰

Om det betyder att lärare gör relativa tolkningar av mål och kriterier eller att de medvetet bortser från vissa mål och kriterier, förklaring (i), framgår inte tydligt av intervjustudien. Det är naturligtvis praktiskt taget omöjligt att genom empiriska undersökningar helt skilja de olika förklaringarna åt.

Sammanfattningsvis finns således endast en sannolik eller rimlig förklaring till att förklara skillnader i skolors nettoavvikelser som dels ligger inom ramen för bestämmelserna om betygsättning och dessutom inte har negativa konsekvenser för likvärdigheten, nämligen *stödinsatser efter provet (c)*. Av övriga rimliga förklaringar ryms bara en inom ramen för betygsbestämmelserna, den om att *lärare på olika skolor gör olika tolkningar av betygskriterierna (f)*, men den har negativa konsekvenser för likvärdigheten. De sista två förklaringarna som kan anses rimliga ryms inte inom ramen för betygsbestämmelserna, *betygsättning på andra grunder (i)* samt att *lärare har svårt att underkänna elever (j)* och har också negativa konsekvenser för likvärdigheten. En rimlig slutsats är därför att även om en del av de skillnader som kan observeras i skolors nettoavvikelser kan förklaras med orsaker som ryms inom ramen för

¹¹⁹ Jmf Måhl 2005b s.63ff.

¹²⁰ Skolverket 2005f s.132f. Jmf Korp 2006 s.259.

betygssystemet och inte påverkar likvärdigheten så finns starka skäl att misstänka att en betydande del också speglar skillnader i hur betygen sätts.

8.4 Slutdiskussion

Att det finns problem när det gäller en likvärdig betygsättning är ingen nyhet. Vad denna rapport tillför diskussionen är en empirisk bekräftelse på några av de mönster på variationer mellan skolors avvikelser i förhållande till de nationella proven som redan visades i en tidigare rapport från Skolverket 2004.¹²¹ Resultaten är likartade för hela tidsperioden 1998-2006. Dessutom har ytterligare fördjupade analyser genomförts som styrker hypotesen om att de skillnader som syns mellan skolor i relationen provbetyg-slutbetyg speglar faktiska skillnader i betygsättning mellan skolor. Ett viktigt bidrag är att utifrån provbetygen som indikator så har likvärdigheten mellan skolor vare sig förbättrats eller försämrats under tidsperioden 1998-2006.

En begränsning med rapporten är att den inte på något sätt ger en heltäckande bild av likvärdighet i betygsättning utifrån de nationella proven. Huvudfokus i rapporten ligger på att analysera variation i hur *skolor* sätter betyg i relation till de nationella proven. Men utöver variation mellan skolor finns det säkerligen också en betydande variation mellan olika *klasser* inom skolor. Denna aspekt av variation i betygsättning syns inte när man analyserar skolors genomsnittliga nettoavvikelse eftersom en skola med nettoavvikelsen noll mycket väl skulle kunna bestå av en klass där samtliga elever får ett höjt slutbetyg i förhållande till provbetyget och en annan klass där samtliga elever får ett sänkt slutbetyg i förhållande till provbetyget. Dessutom finns det antagligen en viss variation i hur proven bedöms och som troligtvis samvarierar med hur slutbetygen sätts. Detta innebär att variationen i skolors nettoavvikelse snarare är underskattad än överskattad. En ytterligare aspekt av likvärdighet som rapporten inte analyserar är hur olika elevgrupper betygsätts i förhållande till provresultaten, t ex utifrån kön och migrationsbakgrund.

En ytterligare begränsning med rapporten är att den är kvantitativ och utgår från ett statistiskt betraktelsesätt där syftet är att finna eventuella systematiska skillnader. Med tillgängliga statistiska data går det inte att ta fram säkra förklaringar till varför vissa skolor avviker mer än andra i förhållande till provbetygen. I föregående avsnitt har diskuterats ett antal tänkbara förklaringar där till viss del empiri från andra typer av undersökningar har utnyttjats. Det skulle vara intressant att följa upp resultaten i denna rapport med mer fördjupade och mer kvalitativt inriktade analyser.

Men utifrån de statistiska analyserna och med proven som indikator på en likvärdig betygsättning är ändå en viktig slutsats att det finns uppenbara brister i likvärdigheten mellan skolor och även mellan kommuner. Och ett av rapportens viktigaste resultat är att likvärdigheten utifrån denna indikator inte har förbättrats över tid. Därmed följer frågan på vilket sätt likvärdigheten i betygsättning kan förbättras?

För att besvara den frågan är det viktigt att förstå att de ramar och förutsättningar som finns för rådande skolsystem och bygger på en hög grad av friutrymme på det lokala planet. Detta gäller inte bara i bemärkelsen att läraren är ensam ansvarig för att betygsätta varje elev. Dessutom är varje lärare/skola fri att själv bestämma hur den skall förhålla sig till resultaten på de nationella proven. Varje lärare/skola är också fri

¹²¹ Skolverket 2004a

att välja stoff och arbetssätt i syfte att få eleverna att uppnå de mål som finns i kursplanerna. Målen och betygskriterier i kursplanerna är skrivna på ett relativt öppet vis. I skriften *Grundskola för bildning* om den nya läroplanen och kursplanerna beskrivs skälen till de öppna målen på följande vis:

*Målstyrningen på skolans område kännetecknas av att målen för verksamheten är öppna, dvs. de förutsätter en lokal uttolkning. Detta kommer till uttryck i begreppet deltagande målstyrning. I en deltagande målstyrning med öppna mål deltar såväl den centrala som den lokala nivån i formuleringen av målen.*¹²²

*...de kvaliteter i kunskaper som eleverna förväntas utvecklas i skolan är inte av en sådan art att de låter sig exakt beskrivas eller formuleras. Alltför exakta beskrivningar hindrar snarare måluppfyllelsen eftersom de tenderar att trivialisera kunskapen.*¹²³

Det sågs således som inbyggd ”spänning” i systemet mellan ”öppna mål” och nationell likvärdighet. Det fria tolkningsutrymmet i kursplaner och betygskriterier samt frånvaron av en preciserad relation mellan nationella prov och slutbetyg ger en klar begränsning när det gäller möjligheten att uppnå en likvärdig i betygsättning.

Men vissa möjligheter finns för att förbättra stödet för en mer likvärdig betygsättning. Redan kvalitetsgranskningen av betygsättningen som genomfördes år 2000¹²⁴ pekade på stora brister i kommunikationen mellan lärare inom skolor och inte minst mellan skolor angående betygsättningsfrågor. Med ett stort tolkningsutrymme i betygskriterierna är en utbredd dialog mellan lärare, skolor och huvudmän ett nödvändigt villkor för en likvärdig betygsättning. Här finns uppenbarligen ett utrymme för förbättringar och som redovisats tidigare har staten vidtagit flera åtgärder för att försöka öka lärarnas samstämmiga tolkning av mål och kriterier.

Det finns naturligtvis möjligheter att göra större ingrepp i prov- och betygssystemen om man vill åstadkomma en mer likvärdig betygsättning. I *Utredningen om mål och uppföljning i grundskolan* framställs ett antal förslag när det gäller de nationella proven och betygsättningen.¹²⁵ Utredningen föreslår att de nationella provens olika syften ska begränsas. Det syfte som främst föreslås ska tonas ner, om än inte helt försvinna, är provens förebildliga roll som ett ”kommentarmaterial” till kursplanerna. I och med att utredningen föreslår att kursplanerna ska bli mer tydliga med ett angivet ”huvudsakligt innehåll” (stoff) och att det dessutom ska utarbetas kommentarmaterial till varje kursplan så menar utredaren att proven inte längre behövs för denna funktion. Denna slutsats överensstämmer helt med Skolverkets ställningstagande i en rapport till regeringen 2003¹²⁶ där verket ställer sig kritiskt till att provsystemet ska ha rollen att konkretisera kursmål och betygskriterier:

Kursmål och betygskriterier bör konkretiseras i god tid innan prov skall genomföras. Det är inte rimligt att man skall invänta ett antal nationella prov innan man kan få någon uppfattning om vad staten avser med sina formuleringar av kursmål och betygskriterier. Sekretessen medför vidare att det endast är de lärare som deltar i proven som får ta del av

¹²² Skolverket 1998 s.12.

¹²³ Skolverket 1998 s.38.

¹²⁴ Skolverket 2000.

¹²⁵ SOU 2007:28 kap 3.

¹²⁶ Skolverket 2003 s.16.

denna konkretisering. Man kan dessutom fråga sig hur de ämnen som inte har nationella prov skall få tillgång till någon konkretisering av sina kursmål och betygskriterier. Utredningsgruppen bedömer att en konkretisering av kursplaner och betygskriterier är nödvändig för att ge stöd för en likvärdig och rättvis betygssättning. En sådan konkretisering är dock inte viktig bara i de tre provämnena utan i alla ämnen - meritvärden baseras inte bara på provämnena - och bör därför utformas som särskilda kommentarmaterial.

Det innebär samtidigt att det blir möjligt att stärka provens styrande roll för betygssättningen och provens roll som mått på hur kunskapsmålen nås på skolnivå, huvudmannanivå och nationell nivå. Någon fastslagen relation mellan slutbetygen och proven föreslår dock utredningen inte. Den förebildliga funktionen hos dagens prov gör att man försöker att släppa fritt provuppgifterna efter en viss tid för att provuppgifterna därmed ska kunna diskuteras öppet och fungera som ett förebildligt kommentarmaterial. Utredaren menar att om det förebildliga syftet tonas ner (till förmån för kommentarmaterial till kursplanerna) så kan sekretessen när det gäller provuppgifterna förlängas. Det skulle möjliggöra att proven innehåller samma provuppgifter under längre tid och att därför stabiliteten mellan olika år skulle öka. På nationell nivå skulle då möjligheterna öka för att använda provresultaten som mått på elevernas kunskaper över tid.

Förslagen om tydligare kursplaner med innehållsangivelser och kommentarmaterial ökar också möjligheterna att de nationella proven blir mer stabila som utvärderingsinstrument av elevernas kunskaper. De som har till uppgift att utforma de nationella proven får ett tydligare underlag för sitt arbete eftersom stoffet är mer givet och att tolkningen av mål- och kriterier är förklarade i statliga kommentarmaterial.

Utredningens förslag innehåller emellertid en inbyggd motsägelse. I utredningen konstateras inledningsvis att

Det faktum att det finns betygskriterier borde kanske egentligen ha haft som konsekvens att det inte skulle behövas nationella prov. Väl formulerade och tydliga betygskriterier borde i ett perfekt system vara ett tillräckligt underlag för lärarnas bedömning och betygssättning.¹²⁷

När utredningen nu föreslår betydligt tydligare betygskriterier med ett angivet stoff och dessutom kommentarmaterial kopplat till kursplanerna förefaller det inte helt självklart att man *även* ska föreslå att de nationella proven får en mer styrande roll för betygssättningen. Men slutsatsen är ändå att det förefaller möjligt att med mer omfattande ändringar i kursplaner och bestämmelser om betygssättning och prov åstadkomma en större likvärdighet, givet att de mer utförliga och tydligare ”instruktionerna” får ett faktiskt genomslag i undervisning och bedömning, vilket inte är en självklarhet.¹²⁸ Men även om det inte skulle ske ett *direkt* genomslag genom att lärarna läser och följer kursplanerna så är det mycket sannolikt att det sker en ökad *indirekt* styrning av undervisning och betygssättning. Innehållsangivelser, målpreciseringar och kommentarer till kursplanerna får ett omedelbart genomslag i

¹²⁷ SOU 2007:28 s.256

¹²⁸ För en diskussion om statens svårigheter att styra skolan med tydliga instruktioner se t ex Lindensjö Lundgren 2002.

de nationella proven och på sikt genom läromedlen.¹²⁹ Värdet av förslagen måste i så fall vägas mot de effekter på lärarnas friutrymme, en trivialisering av kunskaperna och andra eventuella effekter som sådana åtgärder kan skapa. Det finns risker med att proven blir för styrande för undervisningen och ersätter läro- och kursplanen.¹³⁰

Om inte de föreslagna förändringarna genomförs återstår att fortsätta med den typ av åtgärder som staten redan börjat vidta och som räknades upp i inledningen av denna rapport: Information, stödjande material och insatser när det gäller betyg och prov, klargöranden i form av allmänna råd och kommentarer, satsning på utbildning och lärarutbildning i dessa frågor, skolinspektioner och fortsatta utvärderingsstudier samt inte minst åtgärder för att åstadkomma en ökad kommunikation och samsyn om betygsättningen mellan lärare och skolor. *Utredningen om mål och uppföljning i grundskolan* diskuterar också några förslag som syftar till att öka diskussionen och dialogen om skillnaderna mellan slutbetygen och proven.

- En skrivning förs in i läroplanen för grundskolan om att läraren ska kunna redovisa för eleven på vilka grunder betygen sätts.
- Rektor skulle kunna begära in skriftliga motiv från lärarna om de avser att sätta slutbetyg som systematiskt avviker från provbetygen.
- Skolan bör ange orsakerna till avvikelser mellan slutbetyg och prov i sina kvalitetsredovisningar.
- För att öka kvaliteten i bedömningarna av proven är det önskvärt om en ytterligare lärare sambedömer proven med den undervisande läraren.

Förhoppningen är att denna rapport har gett ett underlag för diskussion om betygsättningen i förhållande till de nationella proven och även bidragit till idéer om fortsatta utvärderingar i ämnet.

¹²⁹ Bachmann 2004.

¹³⁰ Lundahl 2006 s.392ff.

Referenser

- Bachmann, Kari, Kirsten Sivesind, Azita Afsar och Stefan Hopmann 2004: *Hvordan formidles læreplanen? En komparativ evaluering av læreplanbaserte virkemidler – deres utforming, konsistens og betydning for læreres praksis*. Høyskoleforlaget AS – Norwegian Academic Press.
- Boesen, Jesper 2004: "Bedömarreliabilitet. Med fokus på aspektbedömningen i det nationella B-kursprovet i matematik våren 2002." (PM nr 195) Enheten för pedagogiska mätningar, Umeå universitet. <http://www.umu.se/edmeas/publikationer/2004.html>
- Boesen, Jesper 2006: *Assessing mathematical creativity: comparing national and teacher-made tests, explaining differences and examining impact*. Umeå universitet.
- Cliffordsson, Christina 2004: "Betygsinflation i de målrelaterade gymnasiebetygen." Pedagogisk Forskning i Sverige 2004 årg 9 nr 1 s 1–14.
- Kjellström, Katarina och Olofsson, Gunilla 2004: "Pålitligheten i lärares bedömning av nationella prov – presentation av en undersökning." Matematikbienalen 2004: Malmö högskola.
- Kroksmark, Tomas 2004: "Då den målrationella styrningen möter skolan" i Grundskoletidningen Skolbarn 5/2004.
- Lindensjö, Bo och Ulf P Lundgren 2002: *Utbildningsreformer och politisk styrning*. HLS förlag.
- Lindström, Jan-Olof 1998: "Rättvis rättning i nationella prov." (PM nr 144) Enheten för pedagogiska mätningar, Umeå universitet. <http://www.umu.se/edmeas/publikationer/pdf/pmn144sec.pdf>
- Lindström, Lars och Viveca Lindberg 2005: *Pedagogisk bedömning. Om att dokumentera, bedöma och utveckla kunskap*. HLS förslag.
- Lundahl, Christian 2006: *Viljan att veta vad andra vet. Kunskapsbedömning i tidigmodern, modern och senmodern skola*. Arbetslivsinstitutet: Arbetsliv i omvandling 2006:8.
- Måhl, Per 2005a: "Vilka grundantaganden är i förväg uppställda krav baserade på" i Grundskoletidningen Skolbarn 1/2005.
- Måhl, Per 2005b: "I förväg uppställda krav" i Lindström, Lars och Viveca Lindberg: *Pedagogisk bedömning. Om att dokumentera, bedöma och utveckla kunskap*. HLS förslag 2005.
- Nycander, Maria 2006: "Pojkars och flickors betyg", Lärarexamensarbete 10p, Uppsala Universitet.
- Ohlsson-Wahlsten, Carina 2002: "Öppna svar – hur funkar det? Elever svarar och lärare bedömer i en läsförståelseuppgift i ett nationellt prov." D-uppsats i pedagogik med didaktisk inriktning. Göteborgs universitet: Institutionen för pedagogik och didaktik.
- Olofsson, Gunilla 2006: *Likvärdig bedömning? En studie av lärares bedömning av elevarbeten på ett nationellt prov i matematik A*. Rapport från PRIM-gruppen nr 23, Lärarhögskolan i Stockholm.
- Riksrevisionen 2004: *Betyg med lika värde? – en granskning av statens insatser*. RiR 2004:11.
- Selghed, Bengt 2004: *Ännu icke godkänt. Lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*. Lärarutbildningen. Malmö högskola
- Skolverket 1995: *Att använda standardprovresultat i den kommunala uppföljningen*, En arbetsrapport från Skolverkets kommunala arbetsgrupp för skolekonomiska frågor. Dnr 94:1908.

Skolverket 1998: *Grundskola för bildning – kommentarer till läroplan, kursplaner och betygskriterier*. www.skolverket.se

Skolverket 2000: *Betygsättningen* - Särtryck ur Nationella kvalitetsgranskningar 2000. www.skolverket.se

Skolverket 2001: *Bedömning och betygsättning. Kommentarer med frågor och svar*. www.skolverket.se

Skolverket 2002: *Betygsättningen i fristående skolor - Nationella kvalitetsgranskningar 2001-2002*. www.skolverket.se

Skolverket 2003: "Det nationella provsystemet – vad, varför och varthän?" i *Skolverkets bedömning av dagens system med nationella prov med avseende på kvalitet och kostnadseffektivitet*. Regeringsuppdrag. Redovisning av uppdrag avseende resultatinformation. Del E. www.skolverket.se

Skolverket 2004a: *Det nationella provsystemet i den målstyrda skolan. Omfattning, användning och dilemman*. Fritzes, Stockholm. www.skolverket.se

Skolverket 2004b: *Prövostena i praktiken. Grundskolans nationella provsystem i ljuset av användares synpunkter*. Fritzes. www.skolverket.se

Skolverket 2004c: *Mellan myt och verklighet*. Fritzes. www.skolverket.se

Skolverket 2004d: "Handlingsplan för en rättssäker och likvärdig betygsättning", dnr. 2004:556. www.skolverket.se

Skolverket 2004e: *Allmänna råd och kommentarer. Likvärdig bedömning och betygsättning*. www.skolverket.se

Skolverket 2004f: *Attityder till skolan 2003*. Rapport 243. www.skolverket.se

Skolverket 2004g: *PISA 2003 – Svenska femtonåringars kunskaper och attityder i ett internationellt perspektiv – Resultaten i koncentration*. www.skolverket.se

Skolverket 2005a: *Lärare och elever om gymnasieskolans nationella prov – en enkätstudie*. www.skolverket.se

Skolverket 2005b: *Inspektera för att förbättra - ett informationshäfte om Skolverkets utbildningsinspektion*. www.skolverket.se

Skolverket 2005c: "Hur används kursplanerna i skolan?" PM www.skolverket.se

Skolverket 2005d: "Jämförelse av ämnesbetyg och betyg på nationella prov för kommunala och fristående skolor, äp 9 2004." PM www.skolverket.se

Skolverket 2005e: *Utbildningsinspektionen 2004 Sammanställningar och analyser av inspektionsresultaten*. Rapport 266. www.skolverket.se

Skolverket 2005f: *Nationella prov i gymnasieskolan – ett stöd för likvärdig betygsättning?* www.skolverket.se

Skolverket 2005g: *Svenska och svenska som andraspråk årskurs 9. Ämnesrapport till rapport 251, Nationella utvärderingen av grundskolan 2003*.

Skolverket 2006a: *Ämnesprovet 2006 i matematik i grundskolans åk 9 och specialskolans åk 10*. www.skolverket.se

Skolverket 2006b: *Ämnesprovet 2006 i svenska och svenska som andraspråk i grundskolans åk 9 och specialskolans åk 10*. www.skolverket.se

Skolverket 2006c: *Ämnesprovet 2006 i engelska i grundskolans åk 9 och specialskolans åk 10*. www.skolverket.se

Skolverket 2006d: *Läromedlens roll i undervisningen*. Rapport 284. www.skolverket.se

Skolverket 2006e: *Med fokus på matematik och naturvetenskap*. www.skolverket.se

Skolverket 2006f: *Med fokus på läsförståelse*. www.skolverket.se

Skolverket 2006h: *Rullande stickprovsbaserat system för kunskapsutvärdering av grundskolans ämnen*. www.skolverket.se

Skolvärlden 2004: "Sur och grinig av nationella prov", insändare nr 13 2004.

SOU 2007:28. *Tydliga mål och kunskapskrav i grundskolan. Förslag till nytt mål- och uppföljningssystem*. Betänkande av Utredningen om mål och uppföljning i grundskolan.

Stockholms Utbildningsförvaltning 2006: "*Sambandet mellan grundskolebetyg och gymnasiebetyg läsåret 2004/05*" USK, Stockholm.

Sveriges radio 2007: "Fria skolor Fria betyg.", Radioprogrammet Kaliber, P1, 27 maj 2007.

Söderström, Martin: *Betygsintagning och elevers studieframgång i Stockholms gymnasieskolor*. Institutet för arbetsmarknadspolitisk utvärdering, Rapport 2006:15

Tholin, Jörgen 2006: *Att kunna klara sig i ökänd natur. En studie av betyg och betygs-kriterier – historiska betingelser och implementering av ett nytt system*. Högskolan i Borås. www.hb.se/till/tholin

Utbildningsdepartementet 1994: "Ett uppdrag till Statens skolverk om att utarbeta och tillhandahålla nationella prov". Dnr U94/1031/Gru

Utbildningsdepartementet 1999: "Ett uppdrag till Statens skolverk om ett nationellt provsystem". U1999/3290/S

Utbildningsdepartementet 2004: "Uppdrag till Statens skolverk avseende det nationella provsystemet". U2004/5293/S

Wikström, Christina 2005: *Criterion-Referenced Measurement for Educational Evaluation and Selection*. Enheten för pedagogiska mätningar. Umeå universitet.

Östlund-Stjärnegårdh, Eva: *Godkänd i svenska? Bedömning och analys av gymnasieelevers texter*. Uppsala 2002.