

Analyser av familjebakgrundens betydelse för skolresultaten och skillnader mellan skolor

En kvantitativ studie av utvecklingen över tid i slutet av grundskolan



Analyser av familjebakgrundens betydelse för skolresultaten och skillnader mellan skolor

En kvantitativ studie av utvecklingen över tid i slutet av grundskolan

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

ISSN: 1103-2421
ISRN: SKOLV-R-467-SE

Grafisk produktion: AB Typoform
Omslagsbild: Maskot Bildbyrå AB

Skolverket, Stockholm 2018

Förord

Skolverket ska enligt sin instruktion (2015:1047) genom sin verksamhet främja att alla barn och elever får tillgång till en likvärdig utbildning och annan verksamhet av god kvalitet i en trygg miljö. Vidare ska myndigheten främja en stärkt likvärdighet mellan skolor. En del i detta arbete är att Skolverket ska bistå med kunskaper om hur likvärdig den svenska skolan är och hur likvärdigheten utvecklas över tid.

Begreppet likvärdighet kan delas upp i tre grundläggande aspekter: lika tillgång till utbildning, lika kvalitet på utbildningen och att utbildningen ska vara kompensierande. Denna rapport fokuserar på i vilken utsträckning utbildningen kompensierar för elevernas olika förutsättningar genom att undersöka resultaten i slutet av grundskolan i förhållande till elevernas socioekonomiska bakgrund. Dessutom analyserar rapporten hur resultat och elevsammansättning skiljer sig mellan skolor.

Rapporten är den tredje som Skolverket publicerar och innehåller en omfattande genomgång av hur likvärdigheten, sett till resultaten i slutet av grundskolan, förändras över tid. Genom den här rapporten kan utvecklingen studeras över en längre tidsperiod än tidigare, och den bygger dessutom på mer detaljerade uppgifter.

Vi vill tacka de forskare som har deltagit i diskussionen om våra likvärdighetsanalyser under de senaste åren. Särskilt vill vi tacka Björn Öckert, Institutet för arbetsmarknadspolitisk- och utbildningspolitisk utvärdering, Kajsa Yang Hansen, institutionen för pedagogik och specialpedagogik på Göteborgs universitet och Maria Brandén, Institutet för analytisk sociologi vid Linköpings universitet som alla har gett värdefulla synpunkter på analyserna i denna rapport. Alla analyser, tolkningar och slutsatser som görs i rapporten är helt och hållet Skolverkets egna.

De statistiska analyserna har i huvudsak genomförts av Anders Auer. Daniel Gustafsson och Oscar Oelrich har också bidragit med statistiska analyser för rapporten. Flera andra medarbetare på Skolverket har gjort insatser för att granska och förbättra rapporten.

Skolverket mars 2018

Peter Fredriksson
Generaldirektör

Jonas Sandqvist
Undervisningsråd

Innehåll

Förord	3
Resultatsammanställning	6
1. Inledning	8
Bakgrund	8
Likvärdighetsbegreppet, syfte och avgränsningar	8
Resultatmått, bakgrundsvariabler och metod	10
Rapportens innehåll och disposition	12
2. Familjebakgrundens betydelse för skolresultaten	13
Betydelsen av socioekonomisk bakgrund för betygen	13
Sammanfattning	17
3. Skolornas elevsammansättning – skolsegregation	18
Skolornas elevsammansättning utifrån socioekonomisk bakgrund	18
Skolornas elevsammansättning utifrån migrationsbakgrund	20
Sammanfattning	21
4. Resultatskillnader mellan skolor	22
Mellanskolvariation	22
Skolsegregation som orsak till ökade resultatskillnader mellan skolor	25
Skolnivåeffekter – betydelsen av skolans sammansättning för elevens förväntade resultat	28
Sammanfattning	32
5. Diskussion av resultaten	33
Bilaga B1	40
B1.1 Datakällor och exkluderingar	40
B1.2 Variabler som ingår i analyserna	41
B1.3 Anpassning till byte av betygssystem och betygsskala	54
B1.4 Takeffekter	61
B1.5 Förändring av antalet skolor och byte av skolkod	63
B1.6 Statistiska metoder som används i analyserna	68
Bilaga B2	70
B2.1 Kompletterande analyser för betydelsen av socioekonomisk bakgrund ..	70
B2.2 Syskonkorrelationer	73
B2.3 Betydelsen av socioekonomisk bakgrund enligt PISA	77
B2.4 Förändrade demografiska faktorer för utlandsfödda elever	82

Bilaga B3	85
B3.1 Skolsegregation utifrån nyanlända respektive elever som invandrat efter skolstart.....	85
Bilaga B4	86
B4.1 Variation i resultat mellan klasser och mellan kommuner	86
B4.2 Skolnivåeffekter.....	89
B4.3 Hur mycket minskar mellanskolvariationen när skolnivåeffekter utifrån socioekonomisk bakgrund läggs till modellen?.....	93
Bilaga B5	94
B5.1 Variation i resultat mellan elever	94
Referenser	99

Resultatsammanställning

I den här rapporten analyserar Skolverket i vilken utsträckning utbildningen kompenserar för elevernas olika förutsättningar genom att undersöka resultaten i slutet av grundskolan i förhållande till elevernas socioekonomiska bakgrund. Dessutom analyserar vi hur resultat och elevsammansättning skiljer sig mellan skolor. Det är viktiga faktorer att följa som en del av att analysera utbildningens likvärdighet.

Betydelsen av familjebakgrund för betygsresultaten åren 2000–2015

- Analysen av samtliga elever visar att elevernas *socioekonomiska bakgrund* har fått en större betydelse för elevernas betygsresultat. Ökningen sker framför allt från och med slutet av 00-talet.
 - För utlandsfödda elever har den socioekonomiska bakgrunden fått en kraftigt ökad betydelse, samtidigt som den gruppen har vuxit genom ökad invandring. Detta är viktiga förklaringar till den ökade betydelsen för samtliga elever. Utlandsfödda elever har under senare år fått svårare förutsättningar att nå goda skolresultat genom att invandringsåldern stigit och föräldrarnas socioekonomiska status försämrats.
 - För elever födda i Sverige, både elever med svensk bakgrund och elever med utländsk bakgrund, har den socioekonomiska bakgrunden också fått en ökad betydelse. Men ökningen är betydligt mindre än för utlandsfödda elever.
 - Av de socioekonomiska bakgrundsfaktorerna är det föräldrarnas *utbildningsnivå* som har den största betydelsen för betygsresultaten, men utbildningsnivåns betydelse har inte ökat över tid. Däremot har föräldrarnas *inkomst* blivit allt viktigare, och framstår som den centrala förklaringen till den ökade betydelsen av socioekonomisk bakgrund för betygsresultaten. För utlandsfödda elever har även föräldrarnas utbildningsnivå och grad av bidragstagande bidragit till den ökande betydelsen av socioekonomisk bakgrund för skolresultaten.

Skillnader i skolors elevsammansättning (skolsegregation) 1998–2015

- Eleverna har blivit alltmer uppdelade mellan skolor utifrån socioekonomisk bakgrund, det vill säga skillnaderna i skolornas socioekonomiska sammansättning har ökat. Ökningen i socioekonomisk skolsegregation skedde framför allt från andra halvan av 00-talet och har avstannat de allra senaste åren.
- Sett över hela tidsperioden har eleverna blivit alltmer uppdelade mellan skolor utifrån om eleverna har någon form av utländsk bakgrund, men trenden har avstannat de senaste åren. Skolsegregationen utifrån utlandsfödda elever har däremot inte ökat över tidsperioden.

Skillnader i skolors resultat (mellanskolvariation och skolnivåeffekt) 2000–2016

- Resultatskillnaderna mellan skolor, den så kallade mellanskolvariationen, har nästan fördubblats 2000–2016. Ökningen är i stort sett kontinuerlig, men ökningstakten är högre från och med slutet av 00-talet.
- Större skillnader i skolors elevsammansättning, det vill säga en ökad skolsegregation, kan förklara delar men inte hela den ökade mellanskolvariationen under hela tidsperioden. Ökad skolsegregation kan förklara det mesta av den ökande mellanskolvariationen 2000–2010, men bara drygt hälften 2010–2015.
- Skolans socioekonomiska sammansättning har fått en större betydelse för elevers resultat (ökade *skolnivåeffekter*). Elever med samma socioekonomiska bakgrund får i ökande grad högre betyg om de går i en skola med gynnsam socioekonomisk sammansättning jämfört med om de gått i en skola med ogynnsam socioekonomisk sammansättning. Det gäller särskilt för elever med låg socioekonomisk bakgrund.

Inledning

I det här kapitlet beskriver vi bakgrunden till rapporten, vad det är som analyseras och hur rapporten är upplagd.

Bakgrund

I den här rapporten analyserar Skolverket med hjälp av kvantitativa analyser olika aspekter av likvärdighet i grundskolan och hur dessa har utvecklats över tid. Den första rapporten om likvärdighet¹ publicerades 2006 och gällde utvecklingen 1998–2004 och en andra rapport², publicerad 2012, förlängde tidsserien till och med 2011. Den här tredje rapporten beskriver utvecklingen 1998–2015 och fram till 2016 för vissa tidsserier.

Likvärdighetsbegreppet, syfte och avgränsningar

Syftet med rapporten är att beskriva och analysera likvärdigheten i slutet av svensk grundskola över tid. Skollagen (2010:800) lyfter fram tre aspekter på likvärdighet:³

- lika tillgång till utbildning
- lika kvalitet på utbildningen
- att utbildningen ska vara kompensatorisk.

De bestämmelser som visar på dessa olika aspekter finns alla i skollagens inledande kapitel.⁴ Aspekten *lika tillgång till utbildning* beskrivs så här:

Alla ska, oberoende av geografisk hemvist och sociala och ekonomiska förhållanden, ha lika tillgång till utbildning i skolväsendet om inte annat följer av särskilda bestämmelser i denna lag.⁵

Aspekten om *lika kvalitet på utbildningen* beskrivs så här:

Utbildningen inom skolväsendet ska vara likvärdig inom varje skolform och inom fritidshemmet oavsett var i landet den anordnas.⁶

Ordet kvalitet nämns inte i bestämmelsen, men av propositionen framgår:

Begreppet ”likvärdig” innebär inte att utbildningen ska vara likformig i betydelsen likadan utan att kvaliteten i verksamheten ska vara så hög att de fastställda målen kan uppnås oavsett var i landet verksamheten bedrivs.⁷

1 Skolverket (2006).

2 Skolverket (2012).

3 Skolverket (2012) kap. 1.2.

4 Se även Enkvist m.fl. (2017) för en aktuell juridisk genomgång av likvärdighetsbegreppet.

5 1 kap. 8 § första stycket skollagen.

6 1 kap. 9 § skollagen.

7 Prop. 2009/10:165, s.638.

Den tredje aspekten av likvärdighet som betonas i skollagen är att utbildningen ska sträva efter att vara *kompensatorisk*, att utbildningen ska ta hänsyn till elevers olika behov och sträva efter att uppväga skillnader i elevernas förutsättningar:

I utbildningen ska hänsyn tas till barns och elevers olika behov. Barn och elever ska ges stöd och stimulans så att de utvecklas så långt som möjligt. En strävan ska vara att uppväga skillnader i barnens och elevernas förutsättningar att tillgodogöra sig utbildningen.⁸

Den centrala frågan som behandlas i den här rapporten handlar om den tredje aspekten av likvärdighet, det vill säga skolans kompensatoriska uppdrag. I rapporten beskrivs och analyseras utvecklingen utifrån olika kvantitativa mått (indikatorer) på likvärdighet:

- **Betydelsen av elevens socioekonomiska bakgrund för skolresultaten.**
Indikatorn visar i vilken utsträckning skolväsendet klarar att uppväga skillnader i elevernas socioekonomiska förutsättningar att tillgodogöra sig utbildningen. Elevernas olika förutsättningar undersöks med hjälp av de olika socioekonomiska bakgrundsfaktorer som tidigare forskning och analyser visat har betydelse för elevernas resultat.⁹
- **Skillnader i skolornas elevsammansättning (skolsegregation).**
Denna indikator visar hur elevsammansättningen på skolorna skiljer sig åt utifrån både socioekonomisk bakgrund och migrationsbakgrund.
- **Skillnader i skolornas resultat (mellanskolvariation).**
Denna indikator visar hur resultaten varierar mellan olika skolor. Här undersöks också hur mycket av dessa resultatskillnader som kan förklaras av att elevsammansättningen skiljer sig mellan skolor (skolsegregation).
- **Betydelsen av skolans elevsammansättning för en elevs förväntade resultat (skolnivåeffekt).**
Denna indikator visar hur mycket skolresultaten skiljer sig mellan elever som har samma individuella socioekonomiska bakgrund, men som går på skolor med olika socioekonomisk elevsammansättning.

Det är viktigt att påpeka att det utifrån skollagen inte går att slå fast någon specifik *nivå* för när utbildningen har uppnått en god eller tillräcklig kompensatorisk förmåga. Det ska vara just en strävan hos de ansvariga för utbildningen. Det centrala i den här rapporten är istället att analysera eventuella *förändringar* i likvärdighet över tid, vilket görs genom att analysera utvecklingen av ovanstående indikatorer.

Generellt redovisas resultat för hela tidsperioden 1998–2016 när så är möjligt. Men av skäl som redovisas i nästa avsnitt begränsas de flesta analyser till 2000–2015.

8 1 kap. 4 § andra stycket skollagen.

9 Vi har valt att inte presentera analyser där vi försöker isolera betydelsen av migrationsbakgrund för skolresultaten. Frågan är mycket komplex och det krävs mer utrymme för att ge frågan rättvisa. Se Grönqvist & Niknami (2017a) och (2017b) samt Engzell (2016a). Däremot har migrationsbakgrund en viktig plats i våra analyser, bland annat genom att vi delar upp analyserna efter migrationsbakgrund.

Rapporten kommer också att avgränsa den huvudsakliga resultatredovisningen till ovanstående indikatorer utifrån betygen och registeruppgifter. Motsvarande indikatorer utifrån internationella kunskapsundersökningar som PISA har redovisats i respektive nationella rapporter för dessa studier, men ett begränsat antal indikatorer utifrån PISA-resultaten redovisas även i bilagan avsnitt B2.3 till den här rapporten.

Resultatmätt, bakgrundsvariabler och metod

I det här avsnittet beskrivs kortfattat de resultatmätt och bakgrundsvariabler som används i rapporten. För mer utförliga beskrivningar av de olika variablerna hänvisar vi till bilagan avsnitt B1.2. Där redogörs närmare för hur de olika variablerna har förändrats över tid. Dessutom beskrivs olika problem med variablerna och vilka metoder som används för att försöka hantera dessa problem.

Variabler som mäter elevernas skolresultat

Rapporten och dess bilaga utgår från slutbetyg från årskurs 9 i grundskolan 1998–2016, och använder två typer av resultatmätt:¹⁰

- sammanlagt meritvärde, det vill säga summan av betygen för 16 olika ämnen
- summan av betygen i matematik och engelska.

Ett potentiellt problem är att betygsutvecklingen inte nödvändigtvis speglar utvecklingen av elevernas faktiska kunskaper. Elevernas betyg ökar under samma period som kunskaperna sjunker enligt de internationella kunskapsmätningarna PISA och TIMSS, mätningar som är konstruerade just för att mäta elevernas kunskapsutveckling över tid.¹¹

För att hantera problemet med ökande betyg standardiseras betygen årsvis, så att sambandet mellan bakgrundsvariabler och betyg blir mer jämförbara över tid (vad standardisering innebär är beskrivet i bilagan avsnitt B1.2). Ett ytterligare sätt att hantera problematiken är att komplettera med analyser av betygen i matematik och engelska, två ämnen där det finns nationella prov som i någon mån förankrar betygen och motverkar eventuell betygsinflation.¹²

Ett annat potentiellt problem är att betygssättningen var avvikande de första åren efter att det målrelaterade betygssystemet infördes 1998.¹³ Lärarna var till en början försiktiga med att utnyttja de högsta och lägsta betygen i den nya betygsskalan.¹⁴ För att undvika att sådana speciella omständigheter påverkar analysen av förändringar över tid begränsas trendanalyserna till startåret 2000. Värderna för åren 1998 och 1999 finns fortfarande med i figurerna.

Ett ytterligare problem är bytet av betygsskala 2013, från en fyrgradig skala (ej nått målen, G, VG, MVG) till en sexgradig skala (F–A). De beräkningar vi

10 I bilagan avsnitt B2.3 redovisas också resultat från PISA-undersökningen.

11 Skolverket (2016b), s. 88 ff. Gustafsson m.fl. (2014) s.37 ff.

12 Ämnet svenska, som också har nationella prov, är svårare att tolka, eftersom andelen elever som i stället läser svenska som andraspråk varierar.

13 Bilagan avsnitt B1.3.

14 Skolverket (2012), s. 24 f.

har gjort visar på en begränsad effekt av bytet av betygsskala på våra analyser,¹⁵ men vi har ändå justerat figurerna för denna effekt.

I bilagan avsnitt B1.3 och B1.4 redovisas mer utförligt hur dessa frågor hanteras.

Bakgrundsvariabler för elevers socioekonomiska bakgrund och migrationsbakgrund

Den centrala uppgiften i rapporten är att undersöka hur olika mått som kännetecknar elevernas olika familjebakgrund påverkar deras resultat i slutet av grundskolan. Jämfört med tidigare likvärdighetsrapporter publicerade av Skolverket¹⁶ har vi nu tillgång till fler variabler som beskriver elevers socioekonomiska bakgrund. Vi använder följande olika bakgrundsvariabler (närmare beskrivna i bilagan avsnitt B1.2):

- föräldrarnas utbildningsnivå utifrån en 13-gradig skala, 1998–2016
- föräldrarnas inkomst, 1998–2015
- föräldrarnas grad av bidragstagande, 1998–2015
- ett socioekonomiskt index sammansatt av ovanstående tre variabler
- migrationsbakgrund, där eleverna delas upp i tre kategorier:
 - elever med svensk bakgrund
 - elever födda i Sverige med utländsk bakgrund
 - utlandsfödda elever.

Valet av variabler har sin grund i forskning om vad som har störst påverkan på skolresultaten.¹⁷ Valet grundas också på vilka variabler som är tillgängliga i register.¹⁸

En avgränsning är att vi bara analyserar bakgrundsvariabler för elevernas föräldrar, inte för några andra släktingar eller tidigare generationer.¹⁹

Elevsammansättningen utifrån bakgrundsvariablerna har förändrats över tid. En viktig förändring förändringen är den ökande invandringen det senaste decenniet. Detta hanteras genom att vi, där det är lämpligt, särredovisar resultaten uppdelat på olika migrationsbakgrunder. Även föräldrarnas utbildningsnivå, inkomst och bidragstagande har förändrats över tid. Detta måste hanteras för att vi ska kunna närma oss ett svar på om det är samhällsförändringar eller skolans kompensatoriska förmåga som ligger bakom det vi ser i analyserna. I bilagan avsnitt B1.2 beskriver vi mer om hur vi hanterar förändringar i bakgrundsvariablerna.

15 Bilagan avsnitt B1.3.

16 Skolverket (2006) Skolverket (2012).

17 Skolverket (2009), s. 98, och Sirin (2005). Det finns även registervariabler som kan ha en betydelse som inte är med i analyserna. En sådan variabel är föräldrarnas yrkesstatus. Vi har haft tillgång till en sådan variabel men bara för åren 2001–2012. Analyser vi gjort visar att våra resultat inte har förändrats på något avgörande vis om vi inkluderar variabeln för yrkesstatus. Se Excelfilen.

18 Vi inkluderar även elevens kön och elevens ålder i förhållande till typåldern för respektive årskull/kohort i vissa analyser.

19 För sådana analyser se Adermon m.fl. (2016) och Hällsten & Pfeffer (2017a) och (2017b).

Metoder och kompletterande analyser

Vi beskriver mycket kortfattat i respektive kapitel hur vi har genomfört analyserna och vilka förändringar som har gjorts sedan tidigare analyser. I övrigt hänvisar vi till bilagan avsnitt B1.6 för en mer detaljerad och teknisk beskrivning av metoderna. Det finns även en Excelfil med rapportens figurer och en rad kompletterande analyser att hämta från Skolverkets webbplats.²⁰

Rapportens innehåll och disposition

Kapitel 2 undersöker den centrala frågan om vilken betydelse elevers socioekonomisk bakgrund har för skolresultaten över tid. Analysen delas även upp för att se om den socioekonomiska bakgrundens betydelse utvecklats olika för elever med svensk bakgrund, elever födda i Sverige med utländsk bakgrund respektive för utlandsfödda elever.

Vidare analyseras vad i den socioekonomiska bakgrunden – föräldrarnas utbildningsnivå, inkomster respektive bidragstagande – som har störst betydelse och hur dessa faktorer påverkar på betygsresultaten har utvecklats över tid.

Kapitel 3 beskriver hur *skolsegregationen* har förändrats över tid, det vill säga hur eleverna är sorterade på skolor utifrån föräldrarnas utbildningsnivå, inkomstnivå och utifrån ett socioekonomiskt index. Motsvarande analys görs av skolsegregationen för migrationsbakgrund.

Kapitel 4 beskriver hur resultatskillnaderna i betyg mellan skolor, det vi kallar *mellanskolvariation*, har utvecklats över tid.

Vidare analyseras hur mycket av den ökade mellanskolvariationen som kan förklaras av en ökad skolsegregation. Kapitlet innehåller också analyser av hur elevers betygsresultat påverkas av en skolas socioekonomiska sammansättning, det vi kallar *skolnivåeffekter*, och hur detta har utvecklats över tid.

Kapitel 5 diskuterar vad de olika resultaten kan säga om utvecklingen av likvärdigheten i slutet av grundskolan.

²⁰ Excelfilen finns att ladda ner i anslutning till rapporten i Skolverkets publikationsdatabas på www.skolverket.se.

Familjebakgrundens betydelse för skolresultaten

I det här kapitlet undersöker vi hur stor betydelse familjebakgrunden har för skolresultaten och hur detta har förändrats över tid.

Betydelsen av socioekonomisk bakgrund för betygen

Skolverkets likvärdighetsrapport från 2012 undersökte sambandet mellan elevers betygsresultat och deras socioekonomiska bakgrund utifrån om föräldrarna hade en eftergymnasial utbildning eller inte.²¹ Den här rapporten använder ett betydligt bredare mått på socioekonomisk bakgrund, ett socioekonomiskt index, som är sammansatt av tre variabler som var och en mäter olika aspekter av en elevs socioekonomiska bakgrund:²²

- föräldrarnas utbildningsnivå (13-gradig skala)
- föräldrarnas samlade inkomst
- föräldrarnas grad av bidragstagande.

Ett sådant socioekonomiskt index har fördelar jämfört med att bara använda föräldrarnas utbildningsnivå. Det mest centrala är att det fångar in fler aspekter av socioekonomisk bakgrund jämfört med bara föräldrarnas utbildningsnivå. Det socioekonomiska indexet förklarar en mycket högre andel av variationen i elevernas betyg. Exempelvis förklarar det socioekonomiska indexet cirka 23 procent av variationen i meritvärde 2015, jämfört med cirka 12 procent om man använder måttet om föräldrarna har eftergymnasial utbildning eller inte.²³

Betydelsen av socioekonomisk bakgrund analyseras här med linjär regression, vilket ger två mått: *effektstorleken* eller *andel förklarad variation* (R²). Vi har valt att bara presentera det senare måttet, andel förklarad variation, här i huvudtexten.²⁴

Figur 2.1 visar hur betydelsen av socioekonomisk bakgrund, mätt med det socioekonomiska indexet, har förändrats över tid för både meritvärdet och betygen i matematik och engelska.

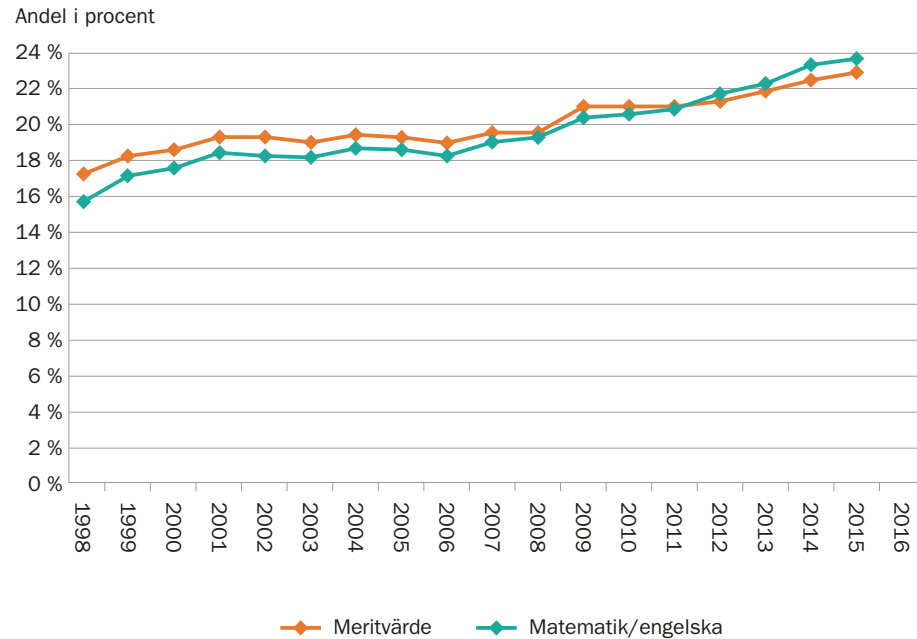
21 Skolverket (2012), kap. 4.

22 Se bilagan avsnitt B1.2 för en mer utförlig beskrivning av indexet.

23 Det socioekonomiska indexet kan förklara cirka 4–4,5 procentenheter mer (beroende på år) av variationen i betygsresultat jämfört med bara den 13-gradiga utbildningsnivåvariabeln, det vill säga R² är cirka 4–4,5 procentenheter högre. Den 13-gradiga utbildningsvariabeln förklarar i sin tur upp till nästan 7 procentenheter mer av variationen jämfört med den dikotoma (2-gradiga) utbildningsnivåvariabeln som användes i den förra likvärdighetsrapporten Skolverket (2012). Se Excel-filen.

24 Se bilagan avsnitt B1.6 för en mer detaljerad förklaring av de två måtten. Alla samband är analyserade utifrån båda måtten, men eftersom mönstren är snarlika oavsett mått håller vi oss till andel förklarad variation i huvudtexten. Motsvarande analys av effektstorlekar redovisas i Excel-filen.

Figur 2.1 Betydelsen av socioekonomisk bakgrund för meritvärdet respektive betygen i matematik och engelska. Samtliga elevgrupper.*



* Betydelsen mäts med andel förklarad variation (R2) i enkel regression med det årsvis standardiserade meritvärdet respektive betygen i matematik och engelska som utfallsvariabel och det socioekonomiska indexet som förklarande variabel. Tidsserierna är justerade för bytet av betygsskala 2013. Med samtliga elevgrupper menas elever med både svensk bakgrund och utländsk bakgrund. Däremot ingår bara elever med personnummer och där det finns uppgift om föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet).

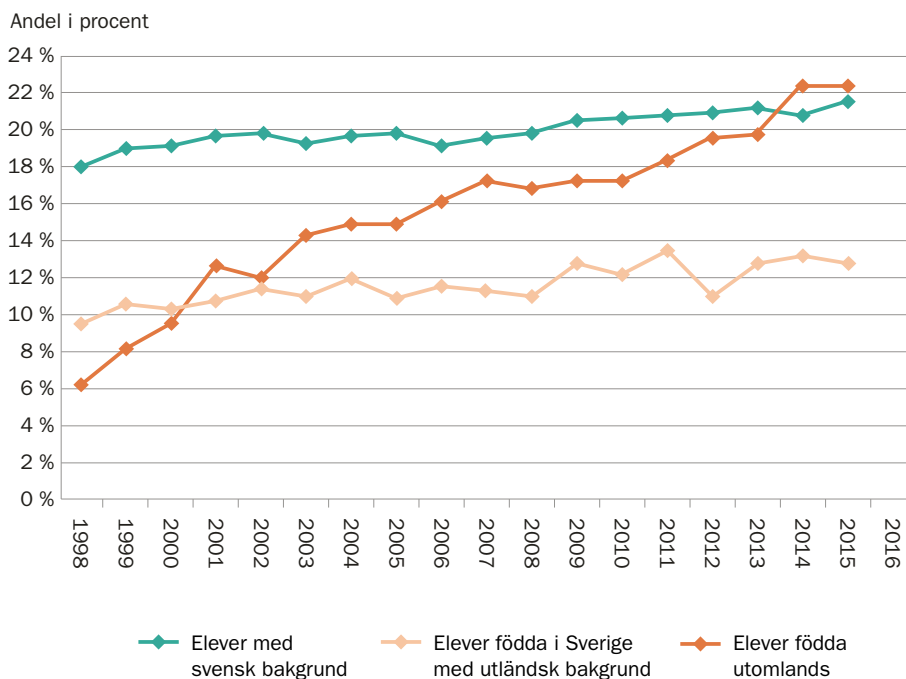
Figur 2.1 visar hur stor andel av variationen i elevers skolresultat som kan förklaras av deras socioekonomiska bakgrund. Eftersom det finns skäl att misstänka att övergången till det nya målrelaterade betygssystemet 1998 kan vara en viktig faktor för ökningen de första åren, är det rimligt att koncentrera sig på förändringar 2000–2015.²⁵ År 2000 var andelen av den totala variationen i elevers meritvärde som kunde förklaras av det socioekonomiska indexet drygt 18 procent. År 2015 hade den andelen ökat till 23 procent. För den här perioden har betydelsen av socioekonomisk bakgrund ökat med drygt 4 procentenheter för meritvärdet och drygt 6 procentenheter när det gäller betygen i matematik och engelska. Oavsett om man analyserar meritvärdet eller betygen i matematik och engelska har alltså den socioekonomiska bakgrunden fått ökad betydelse under tidsperioden. Ökningen sker framför allt från slutet av 00-talet.

²⁵ Se avsnitt 1.2 och bilagan avsnitt B1.3 om varför vi analyserar trender från 2000 i stället för från 1998.

Betydelsen av socioekonomisk bakgrund – uppdelat på migrationsbakgrund

Förändringar i befolkningens sammansättning utifrån bakgrundsvariablerna kan påverka tidsserierna, till exempel genom att andelen utlandsfödda elever har ökat under tidsperioden. Figur 2.2 redovisar därför samma analys som för meritvärdet i figur 2.1, men uppdelad utifrån elevernas migrationsbakgrund.

Figur 2.2 Betydelsen av socioekonomisk bakgrund för meritvärdet, uppdelat på migrationsbakgrund.*



* Betydelsen mäts med andel förklarad variation (R²) i enkel regression med det årsvis standardiserade meritvärdet som utfallsvariabel och det socioekonomiska indexet som förklarande variabel. Tidsserierna är justerade för bytet av betygsskala 2013. I analyserna ingår bara elever med personnummer och där det finns uppgift om föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet).

Figur 2.2 visar på en betydande ökning av den socioekonomiska bakgrundens betydelse för meritvärdet för utlandsfödda elever, från drygt 9 procent år 2000 till över 22 procent år 2015, det vill säga knappt 13 procentenheter. En slutsats utifrån figur 2.2 är att den kraftigt ökande betydelse av den socioekonomiska bakgrunden för utlandsfödda samt att andelen utlandsfödda elever också ökat under tidsperioden är en viktig förklaring till ökningen vi såg i figur 2.1.²⁶ Men även för elever med svensk bakgrund sker en ökning i betydelse sett över hela tidsperioden, men den är betydligt mindre än för utlandsfödda elever. Mellan 2000 och 2015 ökade betydelsen för elever med svensk bakgrund med drygt 2 procentenheter för meritvärdet. För elever födda i Sverige med utländsk bakgrund är utvecklingen mer ojämn, men över tid är trenden också ökande, drygt 2 procentenheter. Liknande trender syns för betygen i matematik och

26 Ökningen av utlandsfödda framgår av bilagan avsnitt B1.2 tabell B1.4.

engelska.²⁷ Möjliga orsaker till att vi ser dessa skillnader mellan migrationsgrupper diskuteras i kapitel 5.

Föräldrarnas utbildningsnivå har störst betydelse, men betydelsen av inkomst ökar

Genom att analysera de olika variablerna som ingår i det socioekonomiska indexet var för sig kan vi få en uppfattning om vilken eller vilka av dessa variabler som ligger bakom den ökade betydelsen av socioekonomisk bakgrund.²⁸ En sådan analys pekar mot att det är en ökad betydelse av föräldrarnas inkomst som huvudsakligen är den drivande faktorn.²⁹ Föräldrarnas utbildningsnivå är visserligen den faktor som har störst betydelse för elevresultaten, men betydelsen av utbildningsnivå har inte ökat över tid.³⁰

När samma analys görs uppdelat på migrationsbakgrund är resultatet för elever med svensk bakgrund likartad den för hela elevgruppen. Betygsresultaten visar en ökad betydelse av föräldrarnas inkomster, medan betydelsen av föräldrarnas utbildning är konstant över tid.³¹ För elever födda i Sverige med utländsk bakgrund syns inga tydliga mönster över tid när det gäller betydelsen av de olika socioekonomiska faktorerna.³² För utlandsfödda elever ser resultatutvecklingen annorlunda ut. Här har samtliga tre socioekonomiska bakgrundsfaktorer bidragit till den kraftigt ökade betydelsen av socioekonomisk bakgrund, men under olika tidsperioder.³³

Syskonkorrelationer och analyser av PISA

Vi har även genomfört två ytterligare analyser av hur betydelsen av socioekonomisk bakgrund påverkat skolresultaten i slutet av grundskolan. Den ena analysen har gjorts med en annan metod (så kallade syskonkorrelationer) och den andra med ett annat datamaterial (PISA). Dessa två analyser redovisas i bilagan avsnitt B2.2 och B2.3.

Syskonkorrelationer mäter hur lika syskons betyg är. Antagandet är att ju mer lika syskons betyg är, desto mer antas familjebakgrunden påverka skolresultaten och desto mindre betydelse har skolan. Resultatet från syskonkorrelationerna uppvisar ett likartat mönster som i figur 2.1 med en ökning över tid med start från slutet av 00-talet. Det ger därmed stöd för att betydelsen av socioekonomisk bakgrund ökat. Utvecklingen av syskonkorrelationerna när eleverna delas upp utifrån migrationsbakgrund visar också ett liknande mönster som i figur 2.2 ovan.

27 Se bilagan avsnitt B2.1, figur B2.1.

28 Se figur B2.2 i bilagan avsnitt B2.1.

29 Se bilagan avsnitt B1.2 där det finns en fördjupning om hur den ökade betydelsen av föräldrarnas inkomster kan tolkas.

30 Jämför Holmlund m.fl. (2014), s. 157 ff, vars analys fram till 2010 också visar en ökad betydelse av inkomst. Se Böhlmark & Holmlund (2012) s.31ff för ett liknande resultat. Se även Fredriksson & Vlachos (2011) s.53.

31 Se bilagan avsnitt B2.1, figur B2.3a.

32 Se bilagan avsnitt B2.1, figur B2.3b.

33 Se bilagan avsnitt B2.1, figur B2.3c.

När det gäller PISA analyseras elevresultaten utifrån PISA:s socioekonomiska index, ESCS. För att få en större säkerhet undersöks de olika ämnena bara utifrån när de var huvudområde i PISA, vilket innebär att tidsperioderna som undersöks varierar för de olika ämnena.³⁴ Analyserna ger inte någon samstämmig bild av utvecklingen. I *läsförståelse* och *naturvetenskap* har betydelsen av familjebakgrund ökat över tid, medan den minskat i *matematik*. Vi har dessutom gjort analyser som visar att det finns osäkerheter i PISA:s socioekonomiska index. Eftersom indexet bygger på uppgifter från eleverna, bör man tolka förändringar över tid i PISA med viss försiktighet. Resultaten för PISA 2015 motsäger i vilket fall inte bilden från registeranalyserna om en ökad betydelse av familjebakgrund för skolresultaten. Resultaten för PISA 2015 motsäger i vilket fall inte bilden från registeranalyserna om en ökad betydelse av familjebakgrund för skolresultaten.³⁵

Sammanfattning

Kapitlet visar att betydelsen av socioekonomisk bakgrund för betygen har ökat 2000–2015, framför allt från slutet av 00-talet. En del av den ökade betydelsen kan förklaras av att gruppen utlandsfödda elever har vuxit och att betydelsen av socioekonomisk bakgrund för denna grupp har ökat kraftigt. Men även för elever födda i Sverige, med svensk respektive utländsk bakgrund har betydelsen ökat, även om ökningen är betydligt mindre.

Föräldrarnas utbildningsnivå är den bakgrundsfaktor som kan förklara den största delen av betygsresultaten, men betydelsen av utbildningsnivå har inte ökat över tid. Det är framför allt föräldrarnas inkomster som har fått en ökad betydelse för elevernas betygsresultat. Men för utlandsfödda elever har även föräldrarnas utbildningsnivå och grad av bidragstagande fått en ökad betydelse.

³⁴ Se bilagan avsnitt B2.3.

³⁵ Även TIMSS-undersökningen tyder på en ökad betydelse av familjebakgrund mellan 2011–2015. Skolverket (2016d) s.58 f.

Skolornas elevsammansättning – skolsegregation

I det här kapitlet analyserar vi hur skolornas sammansättning är uppdelad efter elevernas familjebakgrund, det vi kallar skolsegregation. Skolsegregation analyseras utifrån både socioekonomisk bakgrund och migrationsbakgrund.

Skolsegregation mäts bara utifrån olika bakgrundsvariabler. Detta innebär att tidsserierna inte påverkas av införandet av olika betygsskalor och deras eventuella effekter på lärarnas betygsättning. Trendberäkningarna i det här kapitlet har därmed år 1998 som utgångspunkt i stället för år 2000 som i kapitel 2.

När det gäller analysmetod användes i tidigare rapporter måttet standardavvikelse i skolornas sammansättning, vilket är ett *absolut* mått på skolsegregation. Här används istället variansdekomposition som är ett *relativt* mått på skolsegregation, och en robust metod när det gäller att mäta skillnader över tid (se bilagan avsnitt B1.6 för en förklaring av metoden).

Mellan 1998 och 2008 ökade antalet skolor med årskurs 9-elever (från 1 139 skolor till 1 625 skolor) samtidigt som det genomsnittliga antalet elever per skola minskade.³⁶ Efter 2008 är antalet skolor ungefär detsamma.³⁷ I analyserna har skolorna viktats utifrån elevantal för att varje skola ska få samma proportionella betydelse.³⁸

Skolornas elevsammansättning utifrån socioekonomisk bakgrund

I Skolverkets rapport från 2012 undersöktes skolsegregationen med en variabel som byggde på om föräldrarna hade eftergymnasial utbildning eller inte.³⁹ Här redovisas i stället en analys med samma mer omfattande socioekonomiska variabler som i förra kapitlet.

Figur 3.1 visar skolsegregationens utveckling utifrån tre socioekonomiska variabler. En högre värde innebär att skolorna är mer segregerade, det vill säga eleverna är mer uppdelade på olika skolor utifrån den variabel som undersöks.

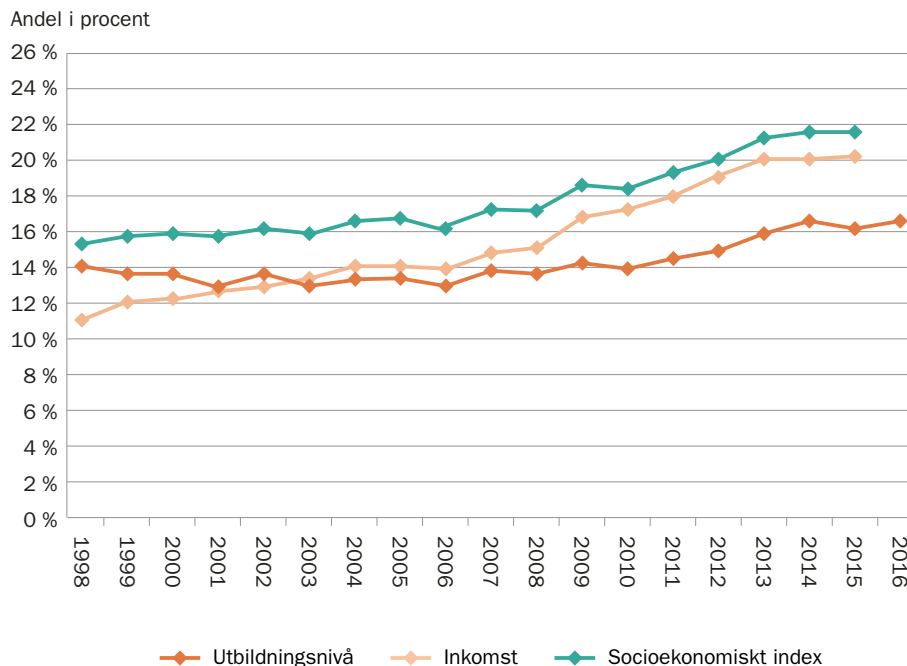
³⁶ Se bilagan avsnitt B1.5.

³⁷ Från och med den 1 juli 2011 infördes begreppet ”skolenhet” i skollagen, vilket medförde att den tidigare s.k. *skolkoden* övergick till *enbetskod*. Skolenhet är inte nödvändigtvis identiskt med den gamla definitionen av en skola, men våra analyser visar att detta inte har haft någon märkbar praktisk betydelse för antalet skolor med årskurs 9. Se bilagan avsnitt B1.5.

³⁸ När skolorna viktas efter elevantal ger variansdekomposition resultat som är oberoende av förändringar i skolstorlek och antal skolor. I bilagan avsnitt B3.1 redovisas även de oviktade resultaten.

³⁹ Skolverket (2012), s. 62 ff. Vi fann då ingen förändring av skolsegregationen mellan 1998 och 2011. Jämför Yang Hansen & Gustafsson (2016) som även de finner en oförändrad skolsegregation utifrån föräldrarnas utbildning, i två nivåer, för samma tidsperiod.

Figur 3.1 Skolsegregation utifrån socioekonomisk bakgrund, mätt med socioekonomiskt index, föräldrarnas utbildningsnivå och föräldrarnas inkomster.*



* Skolsegregationen mäts med variansdekomposition och uttrycker andelen av den totala variationen i variabeln som kan förklaras med hur mycket skolorna skiljer sig i sin genomsnittliga nivå i samma variabel. I analysen ingår samtliga elever med personnummer samt föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet). Viktat för skolstorlek.

Figur 3.1 visar att skolsegregationen har ökat för samtliga socioekonomiska variabler: föräldrarnas utbildningsnivå, föräldrarnas inkomst och det socioekonomiska indexet. Skolsegregationen utifrån föräldrarnas utbildningsnivå var ganska konstant under 00-talet för att sedan börja öka under 2010-talet. Totalt sett har skolsegregationen i utifrån föräldrarnas utbildningsnivå ökat med drygt 2 procentenheter från 1998 till 2016. Skolsegregation utifrån föräldrars inkomst har däremot ökat under i stort sett hela tidsperioden, från drygt 11 procent 1998 till drygt 20 procent 2015.⁴⁰ Den har däremot planat ut under de allra sista åren.

När det gäller det socioekonomiska indexet så påminner utvecklingen mycket om den för utbildningsnivån, vilket beror på att utbildningsnivån är den komponent som väger tyngst i indexet. Skolsegregationen utifrån det socioekonomiska indexet har ökat med drygt 6 procentenheter mellan 1998 och 2015.

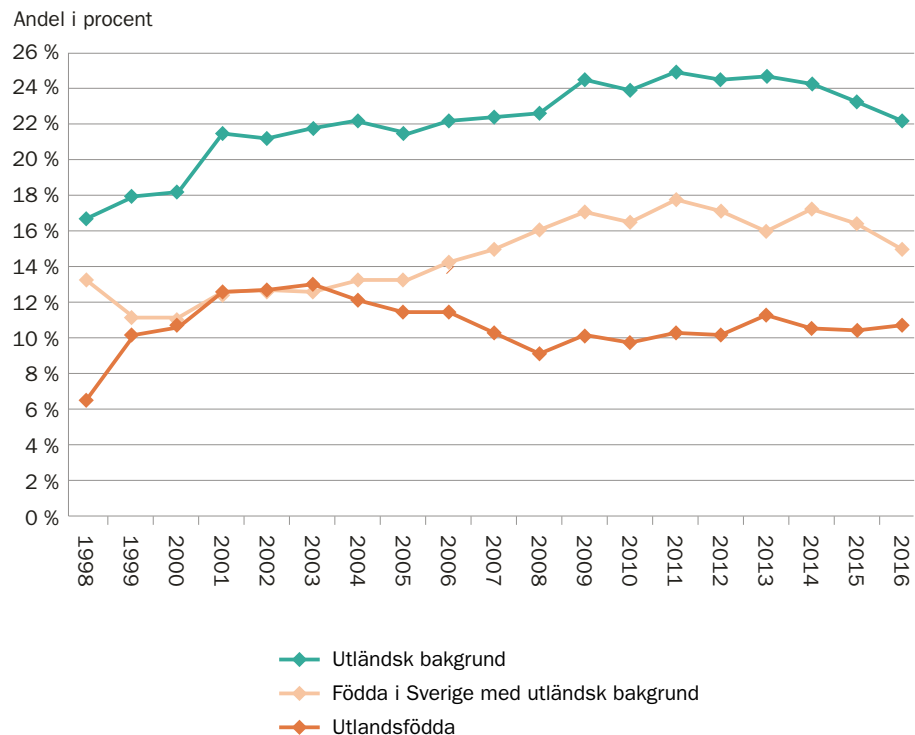
⁴⁰ Vi har även replikerat analysen i Holmlund m.fl. (2014) s. 88 ff. Den redovisar skolsegregationen på ett annat vis, men visar samma utveckling över tid. Se Excel-filen.

Skolornas elevsammansättning utifrån migrationsbakgrund

Figur 3.2 visar skolsegregationen utifrån migrationsbakgrund. Ett högre värde i figuren innebär att skolorna är mer segregerade, det vill säga att eleverna är mer uppdelade på olika skolor utifrån sin migrationsbakgrund.

Observera att elevgrupperna är delvis överlappande, eftersom både utlandsfödda och elever födda i Sverige med utländsk bakgrund ingår i gruppen utländsk bakgrund.⁴¹

Figur 3.2 Skolsegregation utifrån migrationsbakgrund. Elever med någon form av utländsk bakgrund, elever födda i Sverige med utländsk bakgrund och utlandsfödda elever.*



* Skolsegregationen mäts med variansdekomposition och uttrycker andelen av den totala variationen i variabeln som kan förklaras med hur mycket skolorna skiljer sig i sin genomsnittliga nivå i samma variabel. I analyserna ingår bara elever med personnummer och där det finns uppgift om migrationsbakgrund samt föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet). Viktat för skolstorlek.

Figur 3.2 visar att skolsegregation för elever med någon form av utländsk bakgrund ökade från cirka 17 till 25 procent mellan 1998 och 2011. De senaste fem åren har den avtagit, och låg 2016 på ungefär 22 procent, vilket motsvarar en ökning med cirka 6 procentenheter sedan 1998. Skolsegregationen för elever som är födda i Sverige med utländsk bakgrund följer i stort sett samma mönster, men på en lägre nivå. Totalt sett har skolsegregationen för denna grupp ökat

⁴¹ I bilagan avsnitt B3.1 redovisas även skolsegregationen över tid för undergrupperna *nyanlända* elever och *elever som invandrat efter sju års ålder*. För båda dessa grupper syns en ökad skolsegregation sedan slutet av 00-talet.

med cirka 2 procentenheter. För utlandsfödda elever ökade skolsegregationen under slutet av 1990-talet och början av 00-talet, men har sedan avtagit och legat ganska stabilt på drygt 10 procent de senaste åren. Det framgår också av figuren att skolsegregationen sedan mitten av 00-talet är högre för elever som är födda i Sverige med utländsk bakgrund än för utlandsfödda elever.

Sammanfattning

Kapitlet visar att det skett en ökad skolsegregation utifrån både socioekonomisk bakgrund och migrationsbakgrund under 1998–2016. Men skolsegregation utifrån migrationsbakgrund har avtagit eller stabiliserats under senare år. Även skolsegregationen utifrån socioekonomisk bakgrund verkar ha stabiliserats de allra senaste åren.

Resultatskillnader mellan skolor

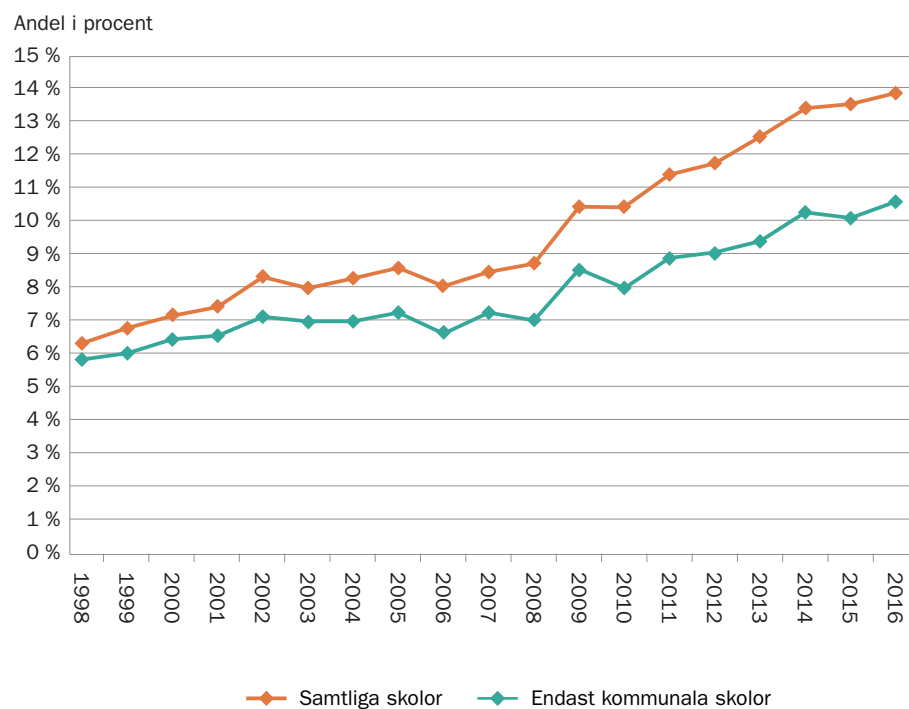
Förra kapitlet visade hur elevsammansättningen utifrån socioekonomisk bakgrund och migrationsbakgrund skiljer sig mellan olika skolor. Det här kapitlet analyserar hur betygsresultaten skiljer sig mellan olika skolor, den så kallade *mellanskolvariationen*, och hur den har förändrats över tid. Vidare analyseras hur stor del av förändringen i mellanskolvariationen som beror på ökad skolsegregation, det vill säga förändrad elevsammansättning mellan skolor. Slutligen analyseras så kallade skolnivåeffekter, det vill säga hur skolans socioekonomiska sammansättning påverkar elevers betygsresultat.

Mellanskolvariation

När det gäller mellanskolvariation så mäts även den, liksom skolsegregationen, med det statistiska måttet variansdekomposition.⁴²

Figur 4.1 visar hur mellanskolvariationen utifrån meritvärdet har utvecklats över tid, dels för samtliga skolor, dels när bara kommunala skolor ingår i analysen.⁴³

Figur 4.1 Mellanskolvariation i meritvärde för samtliga skolor respektive enbart kommunala skolor.*



* Mellanskolvariationen är mätt med variansdekomposition i det årsvis standardiserade meritvärdet. Den uttrycker andelen av den totala variationen i det standardiserade meritvärdet som kan förklaras med hur mycket skolorna skiljer sig i sin genomsnittliga nivå i samma variabel. Skolorna är viktade efter skolstorlek. Samtliga elever med personnummer ingår i analysen.

⁴² Se bilagan avsnitt B1.6 för en förklaring.

⁴³ Se Excel-filen för redovisning av mellanskolvariationen för bara fristående skolor. Se Excel-filen för en oviktad analys av mellanskolvariationen. Mellanskolvariationen ökar något mer med en oviktad analys, men mönstret är detsamma.

Figuren visar hur stor andel av den totala variationen i elevers meritvärde som kan förklaras av variation i skolornas genomsnittliga meritvärde. En större andel (ett högre värde i figuren) innebär att eleverna är mer uppdelade på olika skolor utifrån deras meritvärde.

Figur 4.1 visar att mellanskolvariationen i meritvärden för samtliga skolor ökat över hela tidsperioden, från drygt 7 procent år 2000 till knappt 14 procent år 2016, nästan en fördubbling.⁴⁴ Ökningstakten är större från slutet av 00-talet.

När analysen bara omfattar kommunala skolor visar figur 4.1 dels att nivån på mellanskolvariationen är lägre, dels att den inte ökar i samma omfattning som med samtliga skolor. Figur 4.1 ska inte tolkas som ett bevis för att fristående skolor bidragit till en ökad mellanskolvariation och att denna betydelse ökar. Analysen tar inte hänsyn till ett alternativt scenario där alla elever i stället hade gått i kommunala skolor. Däremot finns forskning som visar att etableringen av fristående skolor har lett till en ökad mellanskolvariation.⁴⁵

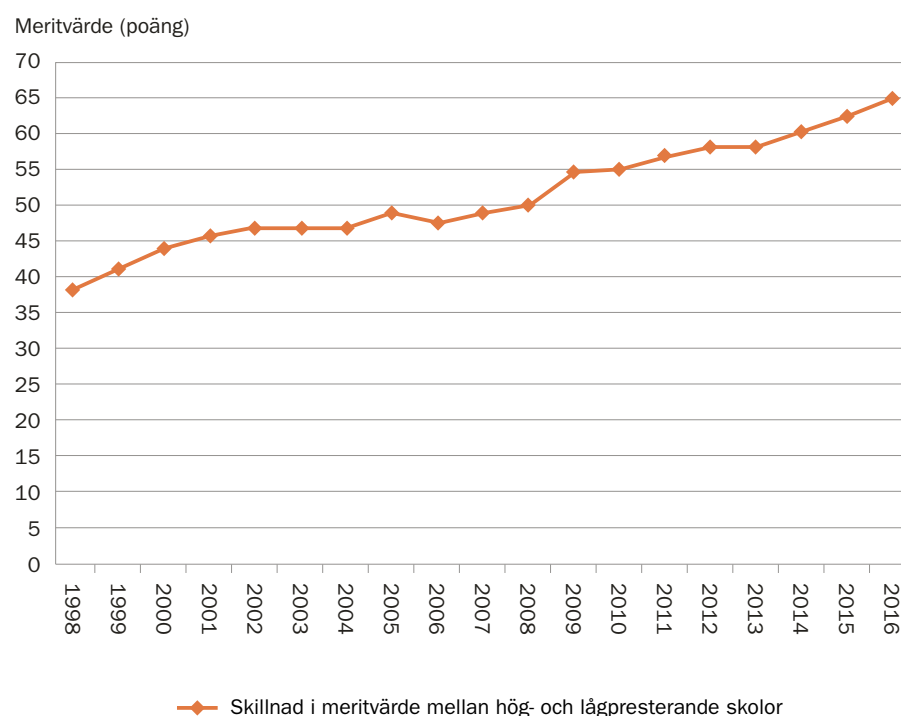
Samtidigt kan konstateras att när det gäller endast kommunala skolor har mellanskolvariationen ökat med mer än 4 procentenheter under tidsperioden, vilket indikerar att fristående skolor knappast kan vara den enda orsaken till den ökande mellanskolvariationen.

Ett förenklat och mindre abstrakt mått på resultatskillnader mellan skolor presenteras i figur 4.2. Här visas för varje år *skillnaden* i meritvärde mellan de högst respektive lägst presterande skolorna. De högst och lägst presterande skolorna utgörs av de 25 procent av skolorna som har högst genomsnittligt meritvärde respektive de 25 procent av skolorna som har lägst genomsnittligt meritvärde. Ett högre värde i figuren indikerar alltså större resultatskillnader mellan hög- och lågpresterande skolor.

44 När mellanskolvariationen analyseras utifrån betygen i matematik och engelska är utvecklingen likartad. Se Excel-filen. Värdena för mellanskolvariationen är lägre i den här figuren än i motsvarande figur i den förra likvärdighetsrapporten, figur 3.1 i Skolverket (2012). Det beror på att värdena i den förra rapporten inte var viktade efter skolstorlek. Men det angavs felaktigt i den diagramrubriken att värdena var viktade. Om skolorna inte viktas efter elevantal blir både nivån och ökningen av mellanskolvariationen högre. Se Excel-filen. Genom att skolornas viktas efter elevantal så tar analysen hänsyn till det faktum att antalet skolor ökat och blivit mindre sett till elevantal (se kapitel 3.1). Om nytillkomna skolor har en systematiskt annorlunda sammansättning, till exempel om de är mer homogena när det gäller skolresultaten, kommer mellanskolvariationen att öka.

45 Böhlmark & Holmlund (2011) s. 41 och Holmlund m.fl. (2014) kap. 10 och s. 344, visar att mellanskolvariationen ökat mer i kommuner där andelen elever i fristående skolor vuxit mer. Se även Böhlmark m.fl. (2015) och (2016), Östh m.fl. (2013), Malmberg m.fl. (2013) och Andersson m.fl. (2012) för en diskussion om skolval och friskolor.

Figur 4.2 Skillnad i meritvärde mellan hög- och lågpresterande skolor.*



* Utryckt som differensen i genomsnittligt meritvärde (ej standardiserat) mellan den översta respektive den understa kvartilen skolor. Data är viktat för skolstorlek. Samtliga elever med personnummer ingår i analysen.

Figur 4.2 visar att år 2000 var skillnaden i meritvärde mellan de högst respektive lägst presterande skolorna drygt 40 poäng. Den skillnaden har sedan ökat kontinuerligt och var år 2016 cirka 65 poäng. Utvecklingen över tid mot större resultat-skillnader mellan skolor visar sig därmed även med detta förenklade mått.

Resultatskillnader mellan klasser inom skolor och mellan kommuner

Analys har också gjorts av resultatskillnader mellan klasser/undervisningsgrupper (inom skolor) respektive resultatskillnader mellan kommuner (bilagan avsnitt B4.1).

När det gäller skillnader mellan klasser/undervisningsgrupper har vi bara tillgång till uppgifter från 2004 och framåt.⁴⁶ Analysen visar att det har skett en ökning av *mellanklassvariationen* i betygsresultat mellan åren 2004 och 2016.⁴⁷

Även när det gäller skillnader mellan kommuner, *mellankommunvariation* i betygsresultat, har det skett en ökning under 2010-talet. Det går även att se en tydlig effekt av att det stora antalet nyanlända som kom 2015 fördelades ojämnt på olika kommuner.

⁴⁶ Uppgifterna om klass/undervisningsgrupp är av olika skäl ganska osäkra, men vi har gjort en del kontroller som tyder på att det ändå inte skett några större förändringar som påverkar trenderna över tid. Se Excel-filen.

⁴⁷ Holmlund m.fl. (2014) s. 134 har tillgång till klassuppgifter längre tillbaka i tiden och kan visa på en ökning av mellanklassvariationen som startar långt tidigare.

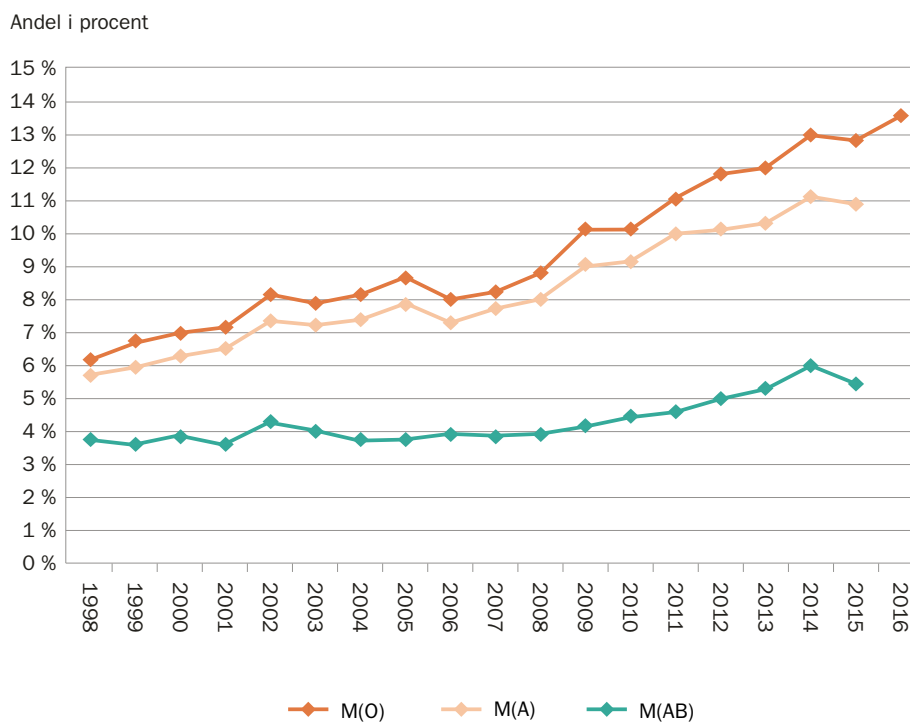
Skolsegregation som orsak till ökade resultatskillnader mellan skolor

Att skolorna har blivit mer segregerade utifrån både socioekonomisk bakgrund och utländsk bakgrund beskrevs i kapitel 3, och kapitel 2 visade att det finns ett samband mellan elevers familjebakgrund och resultat. Därmed kan det förväntas att en ökning av skolsegregationen med avseende på dessa bakgrundsfaktorer också bör leda till ökade skillnader i skolors resultat, alltså ökad mellanskolvariation.

I detta avsnitt går vi vidare och analyserar sambandet mellan den ökade mellanskolvariationen i resultat och den ökade skolsegregationen utifrån elevers bakgrund. Den viktigaste frågan är hur mycket av den ökade variationen i betygsresultat mellan skolor som kan förklaras av den ökade skolsegregationen.

Figur 4.3 visar hur mellanskolvariationen påverkas när hänsyn tas till elevsammansättningen genom att kontrollera för olika bakgrundsvariabler.

Figur 4.3 Återstående mellanskolvariation efter att hänsyn tagits till elevers bakgrund.*



* M(O): Mellanskolvariation utan kontroll för elevers bakgrund. M(A): Mellanskolvariation efter kontroll för elevers kön och migrationsbakgrund. M(AB): Mellanskolvariation efter kontroll för elevers kön, migrationsbakgrund och socioekonomiska bakgrund.

* I analysen ingår bara de elever som har personnummer och där vi har registeruppgifter om migrationsbakgrund samt föräldrarnas utbildningsnivå, inkomst och grad av bidragstagande. I modell M(A) ingår förutom kön och migrationsbakgrund även variabler som indikerar om eleven är yngre respektive äldre än typåldern för respektive årskull. I modell M(AB) baseras den socioekonomiska bakgrunden på det socioekonomiska indexet som i sin tur baseras på föräldrarnas utbildningsnivå, inkomst och grad av bidragstagande. Tidsserien för M(A) respektive M(AB) är begränsad till 2015, eftersom variablerna inkomst och bidragstagande bara är tillgängliga fram till 2015. Data viktat för skolstorlek.

Utgångspunkten för analysen är den mellanskolvariation som visades i figur 4.1 och som i figur 4.3 nedan motsvaras av den blå kurvan, $M(0)$.⁴⁸ $M(0)$ representerar alltså mellanskolvariationen när ingen kontroll görs för elevers bakgrund.

I ett första steg kontrolleras för elevernas kön, migrationsbakgrund samt om eleverna är yngre eller äldre än typåret för respektive årskull. Med *kontrollera* menas att vi tar hänsyn till att en del av variationen mellan skolor beror på att skolornas sammansättning skiljer sig med avseende på dessa variabler. Om dessa variabler har betydelse för att förklara en elevs skolresultat och skolorna samtidigt har blivit mer sorterade efter dessa variabler, kommer den återstående mellanskolvariationen $M(A)$ att vara lägre än den ursprungliga $M(0)$. I praktiken gäller det framför allt variabeln migrationsbakgrund.⁴⁹

Figur 4.3 visar att $M(A)$ -kurvan ligger något lägre än $M(0)$ -kurvan och ökar något mindre. Detta ska tolkas som att en viss del av den ökade mellanskolvariationen kan förklaras av en ökad skolegregation utifrån migrationsbakgrund. Men fortfarande kvarstår den största delen av den ökade mellanskolvariationen.

Nästa steg är att vi även kontrollerar för elevernas socioekonomiska bakgrund, i form av det socioekonomiska indexet, vilket representeras av den gröna kurvan $M(AB)$. Nu sjunker mellanskolvariationen betydligt och ökningen är också betydligt svagare. Tolkningen är att en betydande del av de resultatskillnader som finns mellan skolor kan förklaras av att skolorna är segregerade utifrån elevers socioekonomiska bakgrund, eftersom $M(AB)$ -kurvan är betydligt lägre och ökar mindre än $M(A)$ -kurvan.⁵⁰

Sammantaget visar analysen att en ökad sortering av elever utifrån socioekonomisk bakgrund och migrationsbakgrund kan förklara det mesta av den ökande mellanskolvariationen fram till år 2008, men därefter viker den gröna linjen av uppåt. Det betyder att det därefter skett en viss ökning i mellanskolvariation även när hänsyn tas till ökad sortering/skolegregation i våra variabler för socioekonomisk bakgrund och migrationsbakgrund.

Figur 4.4 redovisar hur stor andel av ökningen i mellanskolvariation som tillsammans kan förklaras av förändrad sammansättning på skolorna utifrån migrationsbakgrund respektive socioekonomisk bakgrund, sett både över hela tidsperioden och uppdelat på 00-talet respektive 2010-talet separat.⁵¹

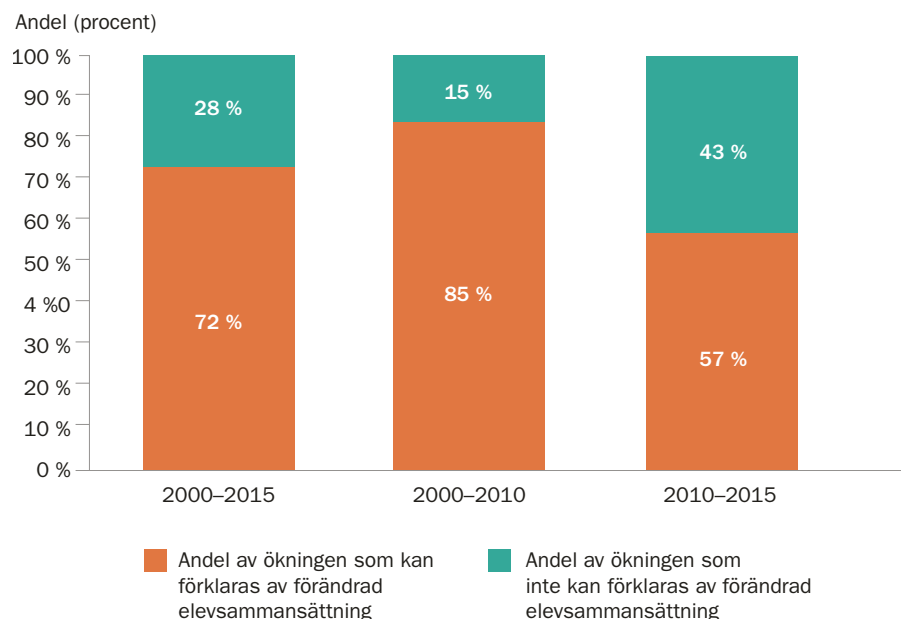
48 Observera att mellanskolvariationen i figur 4.3 inte är helt identisk med den i figur 4.1. Mellanskolvariationen i figur 4.1 är baserad på samtliga elever med personnummer. I figur 4.3 ingår bara de elever som har personnummer och där det dessutom finns uppgifter om migrationsbakgrund samt föräldrarnas utbildningsnivå, inkomst och grad av bidragstagande. Det innebär att det är något färre elever som ligger till grund för analysen jämfört med den i figur 4.1, och därmed blir den kurva som representerar $M(0)$ marginellt annorlunda.

49 Även om kön kan förklara skillnader i resultat mellan elever, är skolorna inte segregerade i någon större utsträckning efter kön. Kön förklarar därmed i praktiken ingenting av mellanskolvariationen, till skillnad från migrationsbakgrunden.

50 Detta är en sanning med viss modifikation. När vi lägger till elevens socioekonomiska bakgrund i modellen kan den återstående mellanskolvariationen minska dels på grund av ökad sortering i denna variabel, dels på grund av en ökad betydelse för resultaten av denna variabel. Vi visade i kapitel 2 att betydelsen av socioekonomisk bakgrund har ökat under tidsperioden och därmed kan detta även förklara en viss men antagligen mindre del av den ökade mellanskolvariationen.

51 Valet av 2010 som brytpunkt, i stället för till exempel 2008, motiveras med att resultaten blir mer jämförbara med tidigare studier som använt samma analysansats, Böhlmark & Holmlund (2012) kap. 3.3 och Holmlund m.fl. (2014) s. 140 ff. Men resultatmönstret är detsamma om brytpunkten i stället är 2008.

Figur 4.4 Andel av ökningen i mellanskolvariation som kan förklaras av ökad skolsegregation.*



* Andel av ökningen som kan förklaras av skolors sammansättning i migrationsbakgrund, kön, om eleven är yngre eller äldre än typåret samt socioekonomisk bakgrund (bestående av föräldrarnas utbildningsnivå, inkomst och grad av bidragstagande).

Observera att i den förklarade andelen ingår även eventuell ökad betydelse för resultaten av dessa variabler. Eftersom kapitel 2 visade att betydelsen av socioekonomisk bakgrund ökat över tidsperioden, så bidrar denna ökning också till att förklara en del av ökningen i mellanskolvariation över samma tidsperiod.

Sett över hela tidsperioden 2000–2015 kan drygt 70 procent av ökningen i mellanskolvariation förklaras av att skolsegregation utifrån migrationsbakgrund och socioekonomisk bakgrund har ökat. Knappt 30 procent av ökningen kan därmed inte förklaras av ökad skolsegregation i dessa variabler.⁵²

Det finns en påtaglig skillnad mellan olika tidsperioder. Under 2000–2010 kan nästan all ökning i mellanskolvariation förklaras av ökad skolsegregation.⁵³ När vi analyserar tidsperioden efter 2010 ändras den bilden. Andelen av de ökade resultatskillnaderna mellan skolor som ökad skolsegregation kan förklara minskar nu betydligt. En betydande andel, 43 procent av ökningen, måste därmed ha andra orsaker än en ökad skolsegregation av eleverna utifrån de variabler vi använder.

52 Utifrån denna modell är det inte möjligt att redovisa exakt hur mycket socioekonomisk bakgrund respektive migrationsbakgrund förklarar var för sig, eftersom detta delvis beror på i vilken ordning man lägger till variablerna. Men oavsett ordning så förklarar den socioekonomiska skolsegregationen (i kombination med ökad betydelse av socioekonomisk bakgrund) åtminstone dubbelt så stor andel av ökningen i mellanskolvariation jämfört med ökad skolsegregation i migrationsbakgrund.

53 Detta resultat överensstämmer med andra analyser som använt samma metod. Böhlmark & Holmlund (2012), kap. 3.3, finner att en stor del men inte all ökad mellanskolvariation fram till 2010 kan förklaras med ökad skolsegregation, det vill säga en ökad sortering av elever utifrån de bakgrundsfaktorer som de använder i sin analys. Holmlund m.fl. (2014) s.140 ff finner att nästan all ökning i mellanskolvariation fram till 2010 kan förklaras med en ökad sortering av elever.

Att samma bakgrundsvariabler förklarar allt mindre av mellanskolvariationen tyder på att det måste finnas andra orsaker till att resultatkillnaderna mellan skolor ökar. En diskussion om andra möjliga orsaker förs i kapitel 5.⁵⁴

Skolnivåeffekter – betydelsen av skolans sammansättning för elevens förväntade resultat

Det här avsnittet undersöker hur elevens resultat påverkas av skolans olika socioekonomiska elevsammansättning, det vi kallar *skolnivåeffekter*. Mer specifikt undersöks om betygsresultaten för elever med samma individuella socioekonomiska bakgrund varierar beroende på om de går på skolor med olika fördelaktig socioekonomisk elevsammansättning.

För att underlätta förståelsen av skolnivåeffekter görs här en förenklad framställning där man kan se hur de faktiska meritvärdespoängen varierar beroende på skolans sammansättning med en viss socioekonomisk bakgrund. För att ytterligare tydliggöra framställningen redovisas resultaten uppdelade på olika elevkategorier och olika skolkategorier.⁵⁵

I figur 4.5a–b illustreras *skolnivåeffekten utifrån socioekonomisk bakgrund* på följande vis. Eleverna har kategoriserats efter hur fördelaktig deras egen individuella socioekonomiska bakgrund är:

- Elevgrupp A: den fjärdedel av eleverna som har bäst socioekonomiska förutsättningar (högst socioekonomiskt index).
- Elevgrupp B: de 50 procent av eleverna som tillhör mittensegmentet på skalan för det socioekonomiska indexet.
- Elevgrupp C: den fjärdedel av eleverna som har sämst socioekonomiska förutsättningar (lägst socioekonomiskt index).

På samma sätt rangordnas skolorna i tre olika grupper efter skolornas socioekonomiska elevsammansättning:

- A-skolor: den fjärdedel av skolorna som har högst genomsnittlig socioekonomisk sammansättning
- B-skolor: de 50 procent av skolorna som tillhör mittensegmentet på skalan för den genomsnittliga socioekonomiska sammansättningen
- C-skolor: den fjärdedel av skolorna som har lägst genomsnittlig socioekonomisk sammansättning.

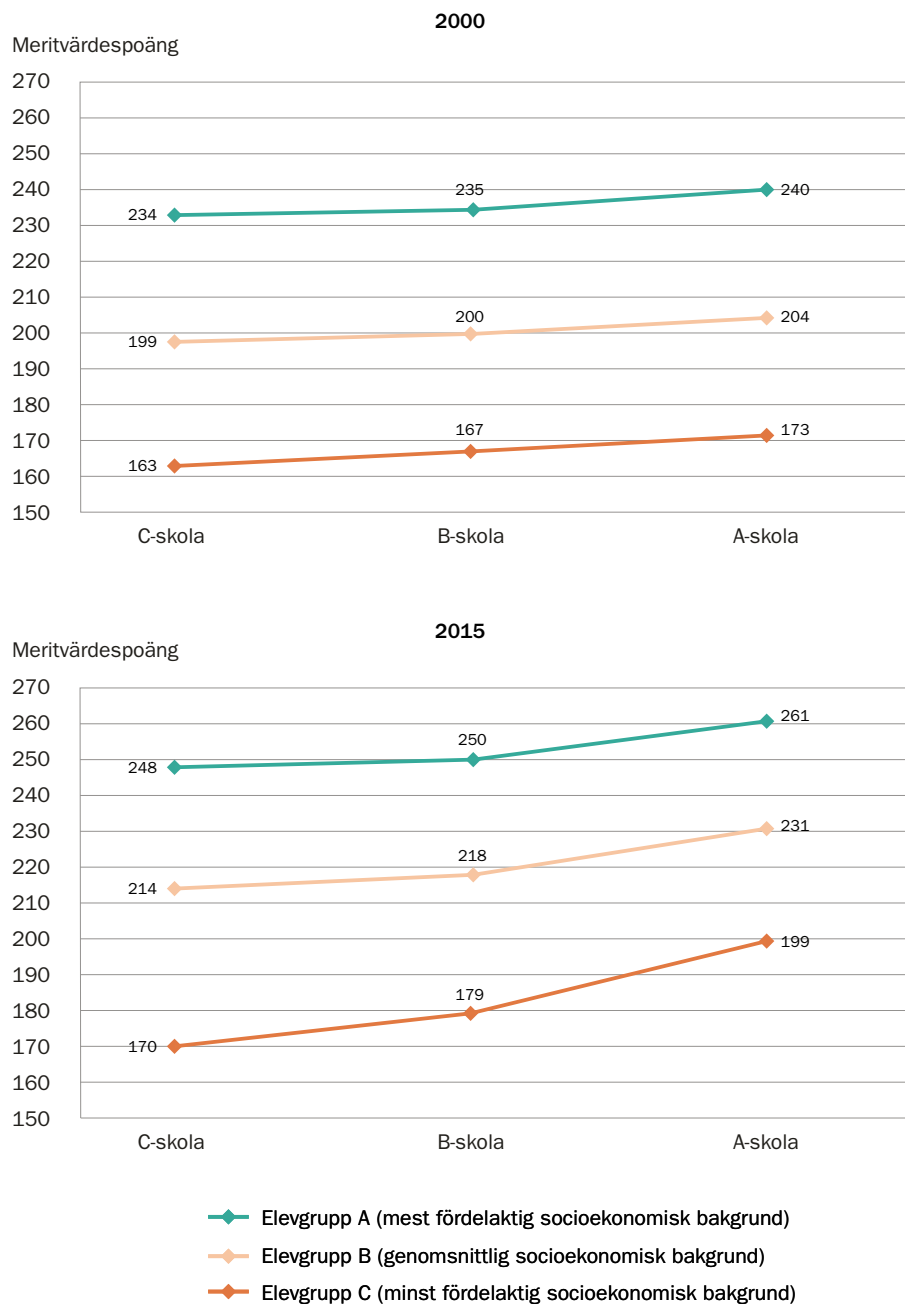
I figur 4.5a–b redovisas resultaten i form av meritvärdet för varje kombination av elevgrupp och skolgrupp, för år 2000 respektive 2015.

⁵⁴ Se även bilagan avsnitt B4.3.

⁵⁵ Se bilagan avsnitt B4.2 för en mer formell redovisning av skolnivåeffekten.

I förra likvärdighetsrapporten, Skolverket (2012), kap. 6, redovisades skolnivåeffekten på ett annorlunda sätt. Den var inte uppdelad på olika elevgrupper av socioekonomisk bakgrund. Dessutom utgjordes måttet av regressionskoefficienten i en regressionsmodell i två nivåer. Slutligen baserades den socioekonomiska bakgrunden på den dikotoma (2-gradiga) utbildningsvariabeln i stället för det socioekonomiska indexet.

Figur 4.5a–b Genomsnittligt meritvärde beroende på elevens socioekonomiska bakgrund i kombination med skolans socioekonomiska sammansättning, 2000 respektive 2015.*



* Eleverna är indelade efter de 25 procent med lägst, 50 procent mitten respektive 25 procent med högst socioekonomiskt index. Skolorna är indelade efter de 25 procent skolor med lägst, 50 procent mitten respektive 25 procent högst genomsnittligt socioekonomiskt index. I analyserna ingår endast elever med personnummer och där det finns uppgift om föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet).

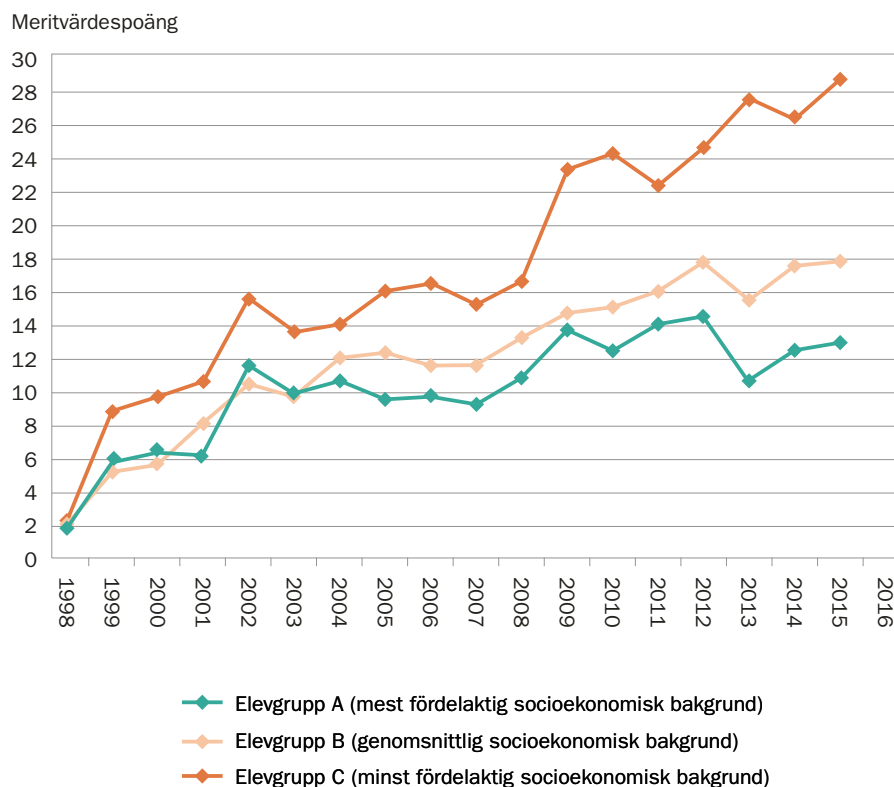
I figur 4.5a–b kan resultaten jämföras för elever som tillhör samma elevgrupp, det vill säga har samma socioekonomiska bakgrund, men som går på olika skolor utifrån skolgrupp, det vill säga olika socioekonomisk sammansättning. Detta görs genom att följa en given kurva i figurerna.

Figur 4.5a visar då att år 2000 för elever i elevgrupp C (den lägsta kurvan), skilde det i genomsnitt 10 meritvärdespoäng beroende på om de gick i en A-skola eller i en C-skola ($173-163 = 10$ poäng). För elever i elevgrupp B och A var skillnaderna mindre, bara cirka 5–6 poäng. Detta innebär att det år 2000 fanns vissa skolnivåeffekter, men att de var relativt små, 5–10 poäng beroende på elevgrupp.

År 2015 (figur 4.5b) visar på större skolnivåeffekter. För elever i elevgrupp C är nu skillnaden 29 poäng ($199-170 = 29$) beroende på om de går på en skola med en relativt fördelaktig socioekonomisk sammansättning (A-skola) eller en skola med en relativt ofördelaktig socioekonomisk sammansättning (C-skola). Även för de andra två elevgrupperna har motsvarande skolnivåeffekter ökat och är nu 13 respektive 17 poäng, men alltså fortfarande lägre än för elever med de relativt sämsta socioekonomiska förutsättningarna.

Figur 4.6 visar skolnivåeffekternas utveckling för varje enskilt år från 1998 till 2015 i en och samma figur, definierade utifrån samma koncept som i figur 4.5a–b men med skillnaden att det nu bara är själva skillnaden mellan A-skolor och C-skolor som visas för respektive elevgrupp. Ett exempel. Den blå kurvan representerar elever i elevgrupp C, det vill säga elever med relativt ofördelaktig socioekonomisk bakgrund. Värdena är skillnaden i genomsnittligt resultat för dessa elever beroende på om de går i en skola med en relativt fördelaktig socioekonomisk sammansättning (A-skola) eller en skola med en relativt ofördelaktig socioekonomisk sammansättning (C-skola), det vill säga 10 poäng år 2000 och 29 poäng år 2015. Ett högre värde i figuren innebär en större skolnivåeffekt, det vill säga en större skillnad mellan elever med samma socioekonomiska bakgrund men som går på skolor med olika socioekonomisk sammansättning.

Figur 4.6 Skolnivåeffekter över tid, uppdelade efter olika elevgrupper.*



* Skolnivåeffekterna är beräknade som skillnaden i genomsnittligt meritvärde mellan elever som går i skolor med relativt fördelaktig respektive relativt mindre fördelaktig socioekonomisk sammansättning, uppdelat efter elevens egen socioekonomiska bakgrund.

* Eleverna är indelade efter de 25 procent med lägst, 50 procent mitten respektive 25 procent med högst socioekonomiskt index. Värdena avser skillnaden i genomsnittligt meritvärde för de 25 procent skolor med högst respektive de 25 procent med lägst genomsnittligt socioekonomiskt index. I analyserna ingår bara elever med personnummer och där det finns uppgift om föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet).

Figur 4.6 visar att skolnivåeffekten utifrån socioekonomisk bakgrund har ökat mer eller mindre kontinuerligt för samtliga tre elevgrupper fram till 2012. Därefter bryts uppgången för skolnivåeffekter för elever med medel- och hög socioekonomisk bakgrund, medan den fortsätter att öka för elever med låg socioekonomisk bakgrund.

De skolnivåeffekter som redovisas i såväl figur 4.5a–b som 4.6 beskriver det *totala* genomslaget på elevernas resultat av att skolor skiljer sig med avseende på socioekonomisk sammansättning, det vill säga den samlade betydelsen av att gå på en skola med en relativt fördelaktig socioekonomisk sammansättning jämfört med en skola med en relativt ofördelaktig socioekonomisk sammansättning. Det totala genomslaget eller den samlade betydelsen beror på två saker:

- *effektstorleken* av skolnivåeffekten
- *skolsegregationens storlek*, det vill säga hur mycket skolorna skiljer sig i socioekonomisk sammansättning.

Effektstorleken uttrycker hur mycket det förväntade betygsresultatet skiljer sig *per enhets* skillnad mellan skolorna i det socioekonomiska indexet.

Men det totala genomslaget beror även på hur stor skolegregationen är, dvs hur många enheter som det skiljer sig i det socioekonomiska indexet mellan de 25 procent skolor med mest fördelaktig socioekonomisk sammansättning och de 25 procent skolor med minst fördelaktig socioekonomisk sammansättning. Om skolegregationen var noll, det vill säga om sammansättningen såg exakt likadan ut på alla skolor, skulle genomslaget eller betydelsen bli noll, oavsett hur stor effektstorleken är. Ju större skolegregationen är, desto större blir genomslaget för en given effektstorlek, det vill säga desto större blir betydelsen av vilken skola en elev går på.

Eftersom vi i kapitel 3 visade att skolegregationen utifrån det socioekonomiska indexet har ökat, beror en del av den ökning i det totala genomslaget som vi ser i figur 4.4a–b och 4.5 på att skillnaderna i skolornas genomsnittliga socioekonomiska sammansättning har ökat. I bilagan avsnitt B4.2 redovisas effektstorleken för skolnivåeffekten. Den visar också på en ökning under tidsperioden, även om den har avtagit något under senare år.⁵⁶

En diskussion om vad som kan ligga bakom de ökande skolnivåeffekterna förs i bilagan avsnitt B4.2 och i kapitel 5.

Sammanfattning

Kapitlet visar att resultatskillnaderna mellan skolor, den så kallade mellanskolvariationen nästan har fördubblats under tidsperioden. När bara kommunala skolor analyseras syns fortfarande en betydande ökning, men den är svagare än när även fristående skolor ingår.

Kapitlet visar också att en betydande andel av den ökade mellanskolvariationen i betygsresultat kan förklaras av att skolorna i ökad utsträckning skiljer sig i sin elevsammansättning med avseende på elevernas socioekonomiska bakgrund och migrationsbakgrund, det vill säga en ökad skolegregation. Under 00-talet kan det mesta av ökningen i mellanskolvariation förklaras med ökad skolegregation. Men under 2010-talet kan bara drygt hälften av de ökade resultatskillnaderna förklaras på samma sätt. De senaste årens ökning i resultatskillnader mellan skolor måste alltså även bero på något annat än en ökad skolegregation.

Kapitlet visar slutligen att elever som går i skolor med en mer fördelaktig socioekonomisk sammansättning tenderar att prestera bättre skolresultat, även efter att hänsyn har tagits till elevens egen socioekonomiska bakgrund. Vi använder begreppet skolnivåeffekter för att beskriva dessa skillnader i resultat. Skolnivåeffekterna utifrån socioekonomisk bakgrund har ökat över tidsperioden och det gäller i synnerhet för elever med sämre socioekonomiska förutsättningar.

⁵⁶ Se bilagan avsnitt B4.3 för en analys av hur mycket skolnivåeffekten bidrar till att förklara ökningen i mellanskolvariation.

Diskussion av resultaten

I det här avslutande kapitlet sammanfattar vi de viktigaste resultaten av vår analys och diskuterar hur resultaten kan tolkas.

Har skolans kompensatoriska förmåga försämrats?

Rapporten visar, liksom tidigare rapporter och forskning, att den socioekonomiska bakgrunden har en stor betydelse för en elevs skolresultat. Analyserna av socioekonomiska bakgrundsvariabler från register visar att elevernas socioekonomiska bakgrund har fått större betydelse för betygsresultaten i årskurs 9, en ökning som startade någon gång i slutet av 00-talet. Även måttet syskonkorrelationer visar på en likartad utveckling när det gäller en ökad betydelse av familjebakgrunden.⁵⁷

Den ökande betydelsen av familjebakgrund kan, under förutsättningen att vi lutar på att våra variabler och metoder är tillförlitliga, bero på två orsaker. Det kan bero på att skolan blivit sämre på sitt kompensatoriska uppdrag, det vill säga att skolan inte förmår att stödja elever med sämre förutsättningar lika bra som tidigare. Den andra möjliga orsaken är att elevernas förutsättningar, som de har med sig hemifrån, har blivit mer olikartade och gjort det kompensatoriska uppdraget svårare.

När resultaten delas upp efter elevernas migrationsbakgrund ser vi tydliga tecken på att det senare orsaksförhållandet sannolikt har haft en betydelse. Den socioekonomiska bakgrunden har fått en kraftigt ökad betydelse för utlandsfödda elever under hela tidsperioden och syskonkorrelationerna för utlandsfödda elever visar också på en betydande ökning.

Den ökade betydelsen av socioekonomisk bakgrund för utlandsfödda elever samt det faktum att andelen utlandsfödda elever i årskurs 9 ökat under 2010-talet (efter att ha minskat sedan början av 00-talet⁵⁸) är sammantaget en viktig förklaring till att den socioekonomiska betydelsen har ökat när vi analyserar hela elevgruppen. Att den socioekonomiska betydelsen ökat för utlandsfödda elevers betygsresultat behöver däremot inte nödvändigtvis betyda att utlandsfödda elever får en sämre anpassad undervisning eller ett sämre stöd i skolan idag jämfört med tidigare. Det kan också betyda att gruppen utlandsfödda elever startar skolan med sämre förutsättningar och att skolan därmed har en svårare uppgift idag jämfört med för ett antal år sedan. Det finns en hel del som talar för det senare.

57 De forskare som analyserat utvecklingen har tidsserier som sträcker sig längre bak i tiden – men inte så långt fram i närtid som i den här rapporten. För liknande analyser som i den här rapporten, se Holmlund m.fl. (2014) s.153 ff och Böhlmark & Holmlund (2012). Fredriksson & Vlachos (2011) s. 53 undersöker också utvecklingen med uppgifter om föräldrars utbildning och inkomst. Gustafsson & Yang Hansen (2017) analyserar bara betydelsen av föräldrars utbildningsnivå över tid.

58 Se bilagan avsnitt B1.2 tabell B15a samt avsnitt B2.4 figur B2.7.

Utlandsfödda elever har fått sämre grundförutsättningar

De utlandsfödda elevernas genomsnittliga invandringsålder har ökat från 7 till 9 år, vilket naturligtvis ger svårare förutsättningar att nå målen i slutet av grundskolan.⁵⁹ Dessutom kommer de utlandsfödda eleverna i allt högre grad från länder med sämre utbildningssystem etc. Både de utlandsfödda eleverna själva och deras föräldrar har sämre förutsättningar att nå goda resultat jämfört med tidigare årskullar.⁶⁰ I genomsnitt har de utlandsfödda eleverna själva och skolornas lärare fått en allt större utmaning att nå upp till samma betygsresultat som tidigare årskullar utlandsfödda elever. Samtidigt är det viktigt att konstatera att den ökade betydelsen av socioekonomisk bakgrund för utlandsfödda elever skett under hela tidsperioden, även före det att invandringsåldern ökade och invandringen från länder med sämre förutsättningar skedde (figur 2.2). Men om skolorna blivit bättre eller sämre på att undervisa de utlandsfödda eleverna går inte att svara på utifrån våra analyser.

En annan diskussion gäller om en ökad andel nyanlända elever påverkar skolresultaten för eleverna som är födda i landet, till exempel genom att stödresurser går till den ökande andelen nyanlända på bekostnad av elever födda i Sverige. Det är en fråga som diskuterats i en tidigare Skolverksrapport och där det konstaterades att forskningen inte kommit till något entydigt svar på den frågan.⁶¹ Några analyser finns än så länge inte för svenska förhållanden. Det är mycket komplicerat att isolera sådana indirekta effekter på elevers resultat – och det är knappast lättare att mäta indirekta effekter när det gäller den socioekonomiska betydelsen för skolresultaten. Vi kan alltså inte uttala oss om det finns några sådana effekter.

Men våra analyser visar också att den socioekonomiska bakgrundens betydelse ökat även för elever med svensk bakgrund – även om denna ökning är betydligt mindre än för utlandsfödda elever. När vi identifierar vilka faktorer som fått en ökad betydelse visar det sig att det är en ökad betydelse av familjeinkomsten som är den starkaste drivande faktorn bakom den ökande betydelsen av socioekonomisk bakgrund. Detta gäller även för elever med svensk bakgrund. Föräldrarnas utbildningsnivå är den faktor som alltså har störst betydelse, men det är inkomstvariabeln som står för huvuddelen av den ökade betydelsen.⁶²

Att föräldrarnas inkomst fått en ökad betydelse för skolresultaten är ett resultat som överensstämmer med det som Holmlund m.fl. (2014) kom fram till i sin analys som sträckte sig fram till 2010.⁶³ Men Holmlund m.fl. (2014) drog slutsatsen att ökningen skulle kunna förklaras av att inkomstmättet fått en ändrad innebörd över tid och därför inte reflekterar en faktisk ökad betydelse för skolresultaten. De baserar den slutsatsen på att syskonkorrelationerna i deras analys inte uppvisar någon ökad betydelse av familjebakgrund.⁶⁴ Våra analyser,

59 Se bilagan avsnitt B2.4 figur B2.10.

60 Se bilagan avsnitt B2.4.

61 Skolverket (2016a) s. 21. Gustafsson m.fl. (2016) s. 28. En ny amerikansk studie som visar på att en plötslig invandringsvåg inte ledde till några indirekta effekter för infödda elever är Figlio & Özek (2017).

62 Att vi skriver "huvuddelen" beror på att för gruppen utlandsfödda elever så ser vi att även utbildningsnivå och grad av bidragstagande haft en ökad betydelse.

63 Holmlund m.fl. (2014) s. 157 ff. Se Böhlmark & Holmlund (2012) s.31ff för ett liknande resultat. Se även Fredriksson & Vlachos (2011) s.53.

64 Holmlund m.fl. (2014) s. 154 och 161.

som sträcker sig längre fram i tiden, visar däremot på ökade syskonkorrelationer de senaste åren.⁶⁵ Sammantaget finns därmed anledning att utgå från att den socioekonomiska bakgrundens betydelse för skolresultaten har ökat i den svenska grundskolan.⁶⁶

Sverige ett genomsnittsländ i internationella kunskapsmätningar när det gäller familjebakgrundens betydelse för skolresultat

Även om det finns osäkerheter⁶⁷ i de mått som används för att mäta familjebakgrundens betydelse i internationella kunskapsmätningar så måste vi använda dessa om vi vill jämföra likvärdigheten i utbildningen i Sverige med andra länder. I de senaste mätningarna framstår Sverige som ett genomsnittsländ bland deltagande jämförbara länder när det gäller familjebakgrundens betydelsen för skolresultaten i slutet av grundskolan.⁶⁸ I PISA 2015 ligger Sverige på OECD-genomsnittet när det gäller hur elevernas socioekonomiska bakgrund påverkar resultaten i naturvetenskap, matematik och läsförståelse, och är det land i Norden där den socioekonomiska bakgrunden verkar ha störst betydelse.⁶⁹ Även i TIMSS 2015 årskurs 8 i matematik och naturvetenskap är Sverige ett genomsnittsländ när det gäller hur den socioekonomiska bakgrunden påverkar resultaten. Betydelsen är större än i Norge som är det enda nordiska land som finns att jämföra med.⁷⁰ Vad gäller ICCS-undersökningen 2016, som mäter elevernas kunskaper i medborgar- och demokratifrågor i årskurs 8, framstår det som Sverige hamnar över genomsnittet av de deltagande länderna när det gäller betydelsen av socioekonomisk bakgrund.⁷¹

Ökade skillnader mellan skolor och ökade skolnivåeffekter

Vi har kunnat konstatera att det skett en ökad socioekonomisk skolegregation utifrån föräldrarnas inkomster och utbildningsnivå 1998–2015. När det gäller migrationsbakgrund har det också skett en ökad skolegregation, men mönstret är inte helt entydigt när olika undergrupper studeras. Det har skett en tydligt ökad skolegregation utifrån elever med någon form av utländsk bakgrund, men denna segregation har minskat något de allra senaste åren. För elever födda i Sverige med utländsk bakgrund syns ett liknande mönster. För de delvis

65 Dessutom har vi gjort olika kontrollanalyser som visar att ökade inkomstskillnader i samhället troligen inte kan förklara den ökade betydelse som inkomst har fått för betygsresultaten. Se bilagan avsnitt B1.2 för en redovisning av simuleringar på ökade inkomstskillnader och dess konsekvenser på regressions samband mellan inkomst och betyg.

66 Se även bilaga B5 för en analys av hur ökad betydelse av socioekonomisk bakgrund påverkar den totala variationen i betygsresultat.

67 Se bilagan avsnitt B2.3.

68 Även i de internationella mätningar (PIRLS och TIMSS) som jämför elevernas kunskaper i läsförståelse, matematik och naturvetenskap i årskurs 4 framstår Sverige som ett genomsnittsländ när det gäller familjebakgrundens betydelse. Skolverket (2017b) kapitel 3.2. Skolverket (2016d) s.58.

69 Skolverket (2016c) s. 35, 40. OECD (2017) s. 76ff. Samt table 1.6.12a–c.

70 Skolverket 2016e s. 58.

71 Schulz m.fl. (2017) s. 66. När det gäller betydelsen föräldrars utbildning och yrkesstatus hamnar Sverige på ett genomsnitt med betydelsen av av antal böcker i hemmet ser ut att ha det en större betydelse än i de flesta deltagande länder. Här finns dock osäkerheter (se bilagan avsnitt B2.3, i denna rapport.). En multivariat analys i Schultz m.fl. (2017) s.185, ger ett resultat där Sverige hamnar över ett genomsnitt när det gäller betydelsen av socioekonomisk bakgrund. Sverige är ungefär på samma nivå som Danmark och Norge medan Finland har en något lägre betydelse av familjebakgrund.

överlappande elevgrupperna nyanlända respektive elever som invandrat efter skolstart syns en ökning i segregationen sedan slutet av 00-talet. Men för hela gruppen utlandsfödda har skolsegregationen hållit sig på en ganska stabil nivå under de senaste 6–7 åren.

Vi kan också konstatera att det skett en fördubbling av betygsskillnaderna mellan skolor (mellanskolvariation) 2000–2016. Mellanskolvariationen ökar även mellan enbart kommunala skolor, men när man inkluderar fristående skolor i analysen ökar skillnaderna ytterligare. Det är inget absolut bevis för att fristående skolor bidragit till ökande resultatskillnader mellan skolor, men befintlig forskning ger stöd för att det förhåller sig på det viset.⁷²

Vi har inte gjort några egna analyser om vad som ligger bakom den ökade skolsegregationen. Men befintlig forskning ger en samstämmig bild av orsakerna i så måtto att det är en kombination av ökad boendesegregation och föräldrars val av skola som ligger bakom utvecklingen – hur mycket som beror på vad skiljer sig åt mellan olika studier.⁷³

Våra analyser visar att den ökade skolsegregationen, det vill säga en ökad sortering av elever utifrån deras socioekonomiska bakgrund och migrationsbakgrund, ligger bakom huvuddelen av ökningen i resultatskillnader mellan skolor fram till år 2010. 2010–2015 kan däremot bara drygt hälften av ökningen i mellanskolvariation förklaras med en socioekonomisk sortering med våra bakgrundsvariabler.

Vi kan också se att den så kallade skolnivåeffekten, det vill säga hur mycket skolans sammansättning påverkar en elevs förväntade resultat, har ökat betydligt. År 2000 var den förväntade skillnaden mellan två elever som båda tillhörde den fjärdedel med lägst socioekonomisk bakgrund 10 meritvärdespoäng om den ena eleven gick på en skola med en mindre gynnsam socioekonomisk sammansättning och den andra på en skola med en gynnsam socioekonomisk sammansättning. År 2015 hade denna skillnad ökat till nästan 30 meritvärdespoäng.⁷⁴

Har skolans sammansättning blivit viktigare för en elevs resultat?

Vi kan inte vara helt säkra på att de ökade skolnivåeffekterna verkligen betyder att skolans socioekonomiska sammansättning fått en ökad betydelse för den enskilda elevens resultat. Det finns några alternativa förklaringar som skulle kunna innebära att skolans betydelse för elevernas skolresultat *inte* nödvändigtvis ökat.

72 Böhlmark & Holmlund (2011) s. 41 och Holmlund m.fl. (2014) kap. 10 och s. 344 visar att mellanskolvariationen ökar mer i kommuner där friskolesektorn vuxit mycket. Se även Böhlmark m.fl. (2015) och (2016) samt Östh m.fl. (2013), Malmberg m.fl. (2013) och Andersson m.fl. (2012) för en diskussion om skolval och friskolor.

73 Östh m.fl. (2013) kommer fram till att huvuddelen av de ökade skolskillnaderna åren 2000–2006 beror på skolval. Holmlund m.fl. (2014) s. 346 drar slutsatsen att ungefär hälften av den ökande mellanskolvariationen beror på en ökad boendesegregation och hälften på skolvalet. Böhlmark m.fl. 2016 och Böhlmark m.fl. (2015) kommer fram till att ungefär 18 procent av den ökade skolsegregationen beror på val till friskolor och att huvuddelen beror på ökad boendesegregation. Se även Yang Hansen & Gustafsson (2016). Om den etniska boendesegregationen har ökat eller inte diskuteras i detalj i Malmberg m.fl. (2016). Åtgärder för att minska boendesegregation beskrivs i Bergsten & Holmqvist (2013).

74 Även i PISA ser vi ökande skolnivåeffekter, Skolverket (2016c) s. 37 ff. OECD (2016) tabell 1.6.12a–c. Skolnivåeffekten, när vi ökar det socioekonomiska indexet ESCS med en standardavvikelse för en skola, ligger på runt 60–67 PISA-poäng i matematik, läsförståelse och naturvetenskap för Sverige. Det är inte signifikant skilt från OECD-genomsnittet, men runt dubbelt så mycket som för de övriga nordiska länderna. Det indikerar att skolans sammansättning kan ha en större betydelse för en elevs skolresultat i Sverige än i övriga nordiska länder.

(a) En tänkbar förklaring skulle kunna vara att betygssättningen blivit mer generös över tid i skolor med en mer gynnsam socioekonomisk elevsammansättning. Detta skulle i så fall gynna elever som går i skolor med en gynnsam socioekonomisk sammansättning, men det betyder inte att eleverna lär sig mer på dessa skolor – bara att betygen sätts generösare. Vår bedömning är att det inte är troligt att en förändrad betygssättning kan förklara de ökade skolnivåeffekterna. Till exempel har vi inte sett någon markant förändring över tid när det gäller skillnader i skolors betygssättning i förhållande till de nationella proven.⁷⁵

(b) En annan förklaring skulle kunna vara att en ökad invandring ligger bakom de ökade skolnivåeffekterna, till exempel genom att nyanlända i högre grad koncentreras till vissa skolor. Inte heller detta är en trolig förklaring och detta av två skäl. Dels ökade skolnivåeffekten minst lika mycket när andelen utlandsfödda elever minskade under första delen av 00-talet. Dels ser vi även en tydligt ökande skolnivåeffekt när vi enbart analyserar elever med svensk bakgrund.⁷⁶

(c) En annan möjlig förklaring är att vi saknar en del bakgrundsvariabler för eleverna, sådana som till exempel Holmlund m.fl. (2014) haft tillgång till, som föräldrarnas ålder vid födseln, antal syskon, ordning i syskonskaran, uppgifter om föräldrarna är skilda.⁷⁷ En förutsättning för att detta skulle förklara den ökade skolnivåeffekten är i så fall att elever med dessa ytterligare egenskaper skulle ha sorterats på skolor *utöver* de socioekonomiska variabler vi redan undersökt. Detta är knappast troligt.

(d) Mer troligt är i så fall att det finns andra bakgrundsfaktorer som inte är möjliga att fånga in via register som samvarierar med skolans socioekonomiska sammansättning. Det kan ha skett en ökad sortering av elever utifrån ”dolda” faktorer som inte finns i några register, till exempel elevernas studiemotivation, stöd från föräldrarna etc.⁷⁸ Om sortering utifrån dolda bakgrundsfaktorer förklarar hela ökningen i skolnivåeffekter så skulle det i så fall innebära att skolan inte fått en ökad betydelse för en elevs skolresultat. Om detta verkligen skett går av naturliga skäl inte att svara på eftersom dessa faktorer just är dolda.

Men det kan också vara så att ingen av ovanstående förklaringar kan förklara de ökande skolnivåeffekterna. Det finns en risk att de ökade skolnivåeffekterna verkligen speglar att skolans sammansättning har fått en större betydelse för en elevs resultat, att den ökande skolsegregationen lett till att det över tid spelar större roll vilken skola en elev går i.⁷⁹ Skolverket och olika forskare undersöker skolskillnader just eftersom det finns en *risk* för att ökad skolsegregation skulle kunna försvåra skolsystemets förmåga att kompensera för elevernas olika förutsättningar.⁸⁰

75 Skolverket (2016b) s. 83 ff.

76 Se Excelfilen.

77 Holmlund m.fl. (2014) s.140.

78 Skolverket (2012) s.77

79 Se Fredriksson & Vlachos (2011) s.33f för ett liknande resultat med en annan metod.

80 Skolverket 2015s s.10f. SOU 2017:35 s.291f. Se även t ex Holmlund m.fl. (2014) kap2.5. Holmlund (2016) s.24ff. Gustafsson, Sörlin, Vlachos 2016 s.76ff. Böhlmark & Holmlund (2012) kap 3.3.

Risker med ökad skolsegregation

Vi avslutar rapporten med att nämna ett antal risker med ökad skolsegregation. En risk är att utbildningskvaliteten blir sämre på skolor med en mer ogynnsam elevsammansättning genom att mer erfarna lärare och rektorer söker sig från sådana skolor och söker sig till skolor med en mer gynnsam elevsammansättning.⁸¹ Detta problem kan förvärras med en ökad lärarbrist.⁸² En ökad skolsegregation förutsätter också att resursfördelningen ”hänger med” och blir mer omfördelade för att inte den kompensatoriska förmågan ska försämrats. Det finns en risk för att resursfördelningen till skolor inte klarat att kompensera för den ökade skolsegregationen.⁸³ Även om det är svårt att mäta finns också en risk att såväl skolkamraters som lärares förväntningar varierar utifrån skolans sammansättning av elever. Olika förväntningar kan påverka elevernas skolprestationer men även deras attityder, som val av gymnasieprogram och yrkesplaner, risk för kriminalitet etc.⁸⁴ En ökad skolsegregation kan med tiden göra att sådana här så kallade kamrateffekter och lärarförväntningar får ett större genomslag.⁸⁵ Dessutom kan klassrumsklimatet vara sämre (stökigare) på skolor med en mindre gynnsam socioekonomisk sammansättning, något som till exempel PISA-undersökningen pekar på.⁸⁶

Vid sidan av effekten på elevernas skolresultat och framtidsval kan det finnas andra skäl att eftersträva en allsidig sammansättning på skolorna när det gäller elevernas socioekonomiska bakgrund och deras migrationsbakgrund. I 10 kap. 2 § tredje stycket skollagen står under de allmänna bestämmelserna för grundskolan och under rubriken Utbildningens syfte: ”Utbildningen ska främja allsidiga kontakter och social gemenskap och ge en grund för ett aktivt deltagande i samhällslivet.”⁸⁷ I läroplanen finns också formuleringar om skolan som en ”social och kulturell mötesplats”.⁸⁸

81 Hansson & Gustafsson (2016). Holmlund (2016). Karbownik & Martinsson (2014). Skolverket (2014) s.21 ff. Skolverket (2016c) s.43.

82 Skolverket (2017a) kap 4.

83 I Skolverkets lägesbedömning 2015 gjordes bedömningen att den kompensatoriska resursfördelningen inte förändrats tillräcklig för att möta utvecklingen med en ökande skolsegregation, Skolverket (2015) s.59ff. Holmlund m.fl. (2014) s.12, 342, drar en liknande slutsats. År 2014 infördes i skollagen en bestämmelse, 2 kap 8a §, som förtydligar kravet på kommuner att fördela resurser efter elevers behov och förutsättningar samt även en bestämmelse, 2 kap 10 §, som ställer samma krav på rektorer inom skolenheten.

84 Holmlund m.fl. (2014) s. 26f. Björklund m.fl. (2010a) kapitel 8.1.4. Parker m.fl. (2016). Rubie-Davies (2010).

85 Skolverket (2012) s. 78 ff. Böhlmark & Holmlund (2012) s. 39. Chetty m.fl. (2016).

86 Skolverket (2016) s. 45.

87 I propositionstexten för paragrafen framhålls bland annat: ”Det är också av stort värde mot bakgrund av grundskolans fostrande uppgift att elever med olika bakgrund och skilda förutsättningar kan mötas i skolan.” Prop. 2009/10:165 s.367.

88 Förordning om läroplan för grundskolan, förskoleklassen och fritidshemmet (SKOLF 2010:37), Skolans värdegrund och uppdrag.

Skolverket uttrycker i sin senaste lägesbedömning 2017 att en allsidig elevsammansättning på skolorna är viktig för att utveckla elevernas tolerans och förståelse:

Skolverket vill betona vikten av att skolväsendet har ett styrsystem som bidrar till att kompensera för skillnader i barns och elevers förutsättningar och som främjar en allsidig elevsammansättning. För att skolan ska vara en sammanhållande kraft i ett demokratiskt samhälle är det viktigt att elever med olika bakgrund möts och får möjligheter att skapa gemensamma erfarenheter och förståelse för varandra.⁸⁹

En alltför omfattande skolsegregation kan alltså medföra ökade svårigheter för skolväsendet att genomföra både sitt kompensatoriska uppdrag och sitt värdegrundsuppdrag.

⁸⁹ Skolverket (2017a) s. 9.

Bilaga B1

B1.1 Datakällor och exkluderingar

Betygsdata för grundskolan är hämtad från Skolverkets databaser som innehåller uppgifter om elevers betyg, kön, ålder, migrationsbakgrund och föräldrarnas utbildningsnivå. Dessa databaser har sedan länkats med SCB:s RAKS-databas, där bland annat uppgifter om föräldrars inkomst, bidragstagande och yrkesstatus har hämtats. För länkningen används elevens löpnummer, som kan ses som en anonymiserad version av personnumret.⁹⁰

Datarensningar och bortfall

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

Tabell B1.1 Antal elever som ingår i analyserna.

År	med betyg	samt löpnummer	samt utbildningsnivå	samt socioekonomiskt index	Bortfall vid analys där socioindex ingår i relation till elever med betyg
1998	96 926	96 926	93 392	93 392	3,6 %
1999	96 313	96 257	95 392	95 392	1,0 %
2000	99 478	99 366	98 617	98 617	0,9 %
2001	104 255	104 099	101 906	101 906	2,3 %
2002	106 715	106 562	104 352	104 346	2,2 %
2003	108 824	108 684	106 648	106 645	2,0 %
2004	115 732	115 493	113 829	113 828	1,6 %
2005	119 273	118 996	117 400	117 399	1,6 %
2006	125 709	125 546	123 816	123 815	1,5 %
2007	124 952	124 800	122 933	122 933	1,6 %
2008	122 759	122 551	120 938	120 936	1,5 %
2009	118 495	118 273	116 675	116 674	1,5 %
2010	114 176	113 909	112 199	112 199	1,7 %
2011	106 799	106 506	104 826	104 826	1,8 %
2012	100 161	99 844	98 184	98 184	2,0 %
2013	95 491	95 129	93 138	93 138	2,5 %
2014	97 088	96 619	94 473	94 472	2,7 %
2015	97 474	96 459	94 268	93 743	3,8 %
2016	105 600	100 238	97 855	-	-

⁹⁰ Nyckeln mellan löpnummer och personnummer har endast SCB tillgång till.

B1.2 Variabler som ingår i analyserna

Utfallsvariabler

Meritvärde

Summan av elevens 16 betyg. Kan anta värden från 0 till 320 poäng.

Betyg i matematik och engelska

Summan av betyget i matematik och engelska i poängform. Kan anta värden från 0 till 40 poäng.

Standardiserat meritvärde respektive betyg i matematik och engelska

För de flesta analyser används det standardiserade meritvärdet [Z_{meritv}] respektive det standardiserade maen-betyget [Z_{maen}]. Genom att standardisera betygen årsvis blir sambandet mellan betygen och en bakgrundsvariabel mer jämförbara mellan olika år.

$$Z_{\text{meritv}}_{\text{elev } i} = \frac{\text{meritv}_{\text{elev } i} - \bar{X}_{\text{meritv } \text{år } x}}{SD_{\text{meritv } \text{år } x}}$$

Där:

$Z_{\text{meritv}}_{\text{elev } i}$ = standardiserat meritvärde för eleven

$\text{meritv}_{\text{elev } i}$ = meritvärdet för eleven

$\bar{X}_{\text{meritv } \text{år } x}$ = nationellt medelvärde för meritvärdet år x

$SD_{\text{meritv } \text{år } x}$ = nationell standardavvikelse för meritvärdet år x

Tabell B1.2 Deskriptiv bild i form av antal, medelvärde (medelv.) och standardavvikelse (SD) för respektive år för de utfallsvariabler som ingår i analyserna.

	A. Max antal observationer						B. endast elever med löpnummer					
	Meritvärde			maen-betyg			Meritvärde			maen-betyg		
	Antal	Medelv.	SD	Antal	Medelv.	SD	Antal	Medelv.	SD	Antal	Mean	SD
1998	96 926	198,91	59,57	96 926	24,31	8,21	96 926	198,91	59,57	96 926	24,31	8,21
1999	96 313	199,72	62,22	96 313	24,18	8,45	96 257	199,78	62,15	96 257	24,18	8,44
2000	99 478	200,41	64,51	99 478	24,27	8,71	99 366	200,51	64,41	99 366	24,28	8,70
2001	104 255	200,30	66,00	104 255	24,30	8,88	104 099	200,45	65,83	104 099	24,32	8,86
2002	106 715	201,91	65,77	106 715	24,36	8,84	106 560	202,07	65,59	106 560	24,37	8,82
2003	108 824	202,90	65,67	108 824	24,70	8,85	108 684	203,02	65,53	108 684	24,71	8,84
2004	115 732	204,39	65,52	115 732	24,65	8,96	115 493	204,58	65,31	115 493	24,67	8,93
2005	119 273	203,67	65,66	119 273	24,51	8,95	118 996	203,87	65,42	118 996	24,54	8,92
2006	125 709	204,42	65,08	125 709	24,72	8,92	125 546	204,55	64,93	125 546	24,73	8,90
2007	124 952	204,72	65,76	124 952	24,64	9,03	124 800	204,87	65,57	124 800	24,65	9,01
2008	122 759	206,75	66,20	122 759	24,97	9,10	122 549	206,93	65,99	122 549	24,99	9,08
2009	118 495	206,88	67,08	118 495	25,03	9,20	118 273	207,10	66,81	118 273	25,05	9,17
2010	114 176	206,18	67,44	114 176	24,94	9,33	113 909	206,43	67,17	113 909	24,97	9,30
2011	106 799	208,12	66,28	106 799	25,17	9,38	106 506	208,45	65,88	106 506	25,21	9,33
2012	100 161	209,35	66,22	100 161	25,27	9,38	99 844	209,71	65,79	99 844	25,32	9,33
2013	95 491	210,23	64,39	95 414	26,33	9,32	95 101	211,03	63,70	95 052	26,39	9,25
2014	97 088	211,83	65,80	97 059	26,12	9,25	96 619	212,78	64,90	96 590	26,21	9,15
2015	97 474	213,54	68,69	97 434	26,11	9,70	96 459	215,49	66,58	96 419	26,31	9,48
2016	105 600	208,60	79,51	105 487	25,65	10,70	100 238	218,51	67,88	100 160	26,87	9,39

	C. Endast elever med löpnummer och uppgift om socioekon. bakgrund											
	Meritvärde			maen-betyg			Zmeritv*			Zmaen*		
	Antal	Medelv.	SD	Antal	Medelv.	SD	Antal	Medelv.	SD	Antal	Mean	SD
1998	93 392	200,66	58,02	93 392	24,59	7,94	93 392	0,00	1,00	93 392	0,00	1,00
1999	95 392	200,46	61,43	95 392	24,27	8,35	95 392	0,00	1,00	95 392	0,00	1,00
2000	98 617	201,11	63,78	98 617	24,36	8,62	98 617	0,00	1,00	98 617	0,00	1,00
2001	101 906	201,23	65,31	101 906	24,43	8,78	101 906	0,00	1,00	101 906	0,00	1,00
2002	104 346	202,89	64,94	104 346	24,49	8,73	104 346	0,00	1,00	104 346	0,00	1,00
2003	106 645	203,73	65,02	106 645	24,82	8,76	106 645	0,00	1,00	106 645	0,00	1,00
2004	113 828	205,22	64,79	113 828	24,76	8,86	113 828	0,00	1,00	113 828	0,00	1,00
2005	117 399	204,45	65,03	117 399	24,62	8,86	117 399	0,00	1,00	117 399	0,00	1,00
2006	123 815	205,19	64,40	123 815	24,82	8,83	123 815	0,00	1,00	123 815	0,00	1,00
2007	122 933	205,63	64,90	122 933	24,76	8,92	122 933	0,00	1,00	122 933	0,00	1,00
2008	120 936	207,59	65,40	120 936	25,09	9,00	120 936	0,00	1,00	120 936	0,00	1,00
2009	116 674	207,88	66,06	116 674	25,16	9,07	116 674	0,00	1,00	116 674	0,00	1,00
2010	112 199	207,33	66,31	112 199	25,10	9,18	112 199	0,00	1,00	112 199	0,00	1,00
2011	104 826	209,39	64,97	104 826	25,35	9,21	104 826	0,00	1,00	104 826	0,00	1,00
2012	98 184	210,57	65,00	98 184	25,45	9,22	98 184	0,00	1,00	98 184	0,00	1,00
2013	93 117	212,44	62,16	93 066	26,60	9,04	93 117	0,00	1,00	93 066	0,00	1,00
2014	94 472	214,16	63,44	94 447	26,41	8,95	94 472	0,00	1,00	94 447	0,00	1,00
2015	93 743	217,77	64,04	93 712	26,61	9,18	93 743	0,00	1,00	93 712	0,00	1,00
2016	0	-	-	0	-	-	0	-	-	0	-	-

* Standardiserat meritvärde respektive standardiserat maen-betyg -Uppgift om socioekonomisk bakgrund endast tillgänglig fram till 2015.

Tabell B1.3 Elevens kön och hans ålder i förhållande till typåldern för årskursen.*

År	antal	Kön		elev tidig		elev sen	
		Medelv.	SD	Medelv.	SD	Medelv.	SD
1998	93 392	48,7 %	50,0 %	0,7 %	8,6 %	3,2 %	17,7 %
1999	95 392	48,6 %	50,0 %	0,8 %	8,8 %	3,9 %	19,3 %
2000	98 617	48,8 %	50,0 %	1,8 %	13,1 %	3,9 %	19,3 %
2001	101 906	48,6 %	50,0 %	2,6 %	16,1 %	3,6 %	18,7 %
2002	104 346	48,8 %	50,0 %	3,2 %	17,7 %	3,4 %	18,2 %
2003	106 645	49,2 %	50,0 %	3,5 %	18,4 %	3,6 %	18,7 %
2004	113 828	48,7 %	50,0 %	3,3 %	17,9 %	3,5 %	18,3 %
2005	117 399	48,8 %	50,0 %	3,3 %	17,9 %	3,7 %	18,8 %
2006	123 815	48,9 %	50,0 %	3,1 %	17,3 %	3,7 %	19,0 %
2007	122 933	48,4 %	50,0 %	2,6 %	16,0 %	4,2 %	20,1 %
2008	120 936	48,8 %	50,0 %	2,2 %	14,6 %	4,2 %	20,0 %
2009	116 674	48,8 %	50,0 %	1,9 %	13,8 %	4,6 %	20,9 %
2010	112 199	49,0 %	50,0 %	1,8 %	13,1 %	4,6 %	21,0 %
2011	104 826	48,6 %	50,0 %	1,6 %	12,5 %	5,0 %	21,9 %
2012	98 184	49,0 %	50,0 %	1,6 %	12,5 %	5,3 %	22,4 %
2013	93 138	48,7 %	50,0 %	1,5 %	12,3 %	5,5 %	22,7 %
2014	94 472	48,5 %	50,0 %	1,7 %	12,9 %	5,4 %	22,5 %
2015	93 743	48,8 %	50,0 %	1,6 %	12,5 %	5,4 %	22,6 %
2016*	100 238*	48,2 %	50,0 %	1,5 %	12,1 %	5,8 %	23,3 %

*Siffror för 2016 är inte jämförbara med tidigare år. För 1998–2015 gäller siffrorna elever som har uppgift om både [edu13], [Pink] och [bidrag]. För 2016 krävs bara uppgift om migrationsbakgrund, eftersom ingen data finns för [pink] och [bidrag] för 2016.

Bakgrundsvariabler

Kön

Anges som 0 för pojke och 1 för flicka.

Ålder i förhållande till typåldern för årskursen

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

Tabell B1.3 visar andelen elever efter kön och andelen som är yngre respektive äldre än typåldern.

Tabell B1.4 Elevers migrationsbakgrund.*

År	Antal	Svensk bakgrund	Född i Sverige med utländsk bakgrund	Utlandsfödd	Utlandsfödd och Nyanländ
		Andel	Andel	Andel	Andel
1998	93 392	88,5 %	5,2 %	6,3 %	0,1 %
1999	95 392	85,9 %	5,0 %	9,2 %	1,0 %
2000	98 617	85,5 %	5,2 %	9,3 %	1,1 %
2001	101 906	86,9 %	5,5 %	7,6 %	1,2 %
2002	104 346	87,0 %	5,5 %	7,5 %	1,2 %
2003	106 645	86,8 %	5,8 %	7,4 %	1,3 %
2004	113 828	87,1 %	6,0 %	6,9 %	1,4 %
2005	117 399	87,2 %	6,1 %	6,7 %	1,4 %
2006	123 815	87,3 %	6,7 %	6,0 %	1,3 %
2007	122 933	87,0 %	6,9 %	6,1 %	1,4 %
2008	120 936	86,6 %	7,7 %	5,6 %	1,7 %
2009	116 674	85,3 %	9,0 %	5,6 %	2,0 %
2010	112 199	84,3 %	9,4 %	6,3 %	2,5 %
2011	104 826	82,8 %	10,0 %	7,2 %	2,5 %
2012	98 184	81,6 %	10,0 %	8,4 %	2,8 %
2013	93 138	80,6 %	10,1 %	9,3 %	2,8 %
2014	94 472	79,7 %	10,4 %	9,9 %	3,0 %
2015	93 743	79,4 %	10,1 %	10,5 %	3,4 %
2016*	100 238*	76,4 %	9,8 %	13,8 %	5,6 %

* Siffror för 2016 är inte jämförbara med tidigare år. För 1998–2015 gäller siffrorna elever som har uppgift om både [edu13], [Pink] och [bidrag]. För 2016 krävs bara uppgift om migrationsbakgrund, eftersom ingen data finns för [pink] och [bidrag] för 2016. Som nyanlända räknas som elever som blivit folkbokförda i Sverige inom de senaste fyra åren.

Migrationsbakgrund

Eleverna är uppdelade i tre migrationskategorier:

1. Elever med svensk bakgrund: eleven är född i Sverige och har minst en förälder som är född i Sverige
2. Elever födda i Sverige med utländsk bakgrund: eleven är född i Sverige och båda föräldrarna är födda utomlands.
3. Utlandsfödda elever: eleven är född utomlands (oavsett var föräldrarna är födda).⁹¹

Ovanstående definition följer den officiella definitionen som används i den offentliga statistiken. Men den följer inte den definition som används av OECD i PISA-rapporterna.⁹²

Till skillnad från SCB antas här alla elever och föräldrar vara utlandsfödda om uppgift saknas om födelseland.⁹³

Tabell B1.4 visar andelen elever efter migrationsbakgrund.

91 Observera att bland utlandsfödda elever ingår även elever vars båda föräldrar är födda i Sverige, t.ex. barn som är födda när föräldrarna arbetar utomlands.

92 Enligt PISA-definitionen är kategori 2 identisk. Kategori 1 definieras däremot som alla elever där minst en förälder är född i Sverige (oavsett var eleven är född). Till kategori 3 räknas därmed bara elever som dels är utlandsfödda och där också båda föräldrarna är utlandsfödda.

93 SCB placerar i stället alla elever som inte kan definieras på grund av bortfall i födelseland i kategori 1, eftersom det är den största kategorin. Vi anser att det är suboptimalt, eftersom gruppen utlandsfödda (kategori 3) då underskattas med motiveringen att det är högst sannolikt att en individ som saknar uppgift om födelseland är född i utlandet.

Tabell B1.5a Andel elever uppdelade efter föräldrarnas genomsnittliga utbildningsnivå [edu13], samtliga elevgrupper.*

År	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5	6,0	6,5	7,0
1998	2,8 %	2,6 %	10,6 %	13,8 %	19,6 %	10,4 %	10,8 %	7,8 %	7,0 %	5,7 %	7,8 %	1,0 %	0,2 %
1999	2,9 %	2,5 %	10,5 %	13,6 %	19,8 %	10,5 %	11,0 %	7,8 %	6,9 %	5,8 %	7,7 %	1,0 %	0,2 %
2000	2,5 %	2,3 %	9,8 %	14,1 %	20,1 %	10,1 %	11,3 %	7,8 %	7,1 %	5,8 %	7,9 %	1,0 %	0,2 %
2001	2,0 %	1,7 %	8,2 %	13,3 %	21,7 %	11,0 %	11,9 %	8,3 %	7,3 %	5,9 %	7,6 %	1,0 %	0,2 %
2002	1,9 %	1,5 %	7,5 %	13,0 %	22,1 %	11,5 %	12,1 %	8,5 %	7,1 %	5,9 %	7,6 %	1,0 %	0,2 %
2003	1,7 %	1,3 %	6,7 %	12,6 %	22,2 %	12,0 %	12,4 %	8,9 %	7,4 %	6,0 %	7,5 %	1,0 %	0,2 %
2004	1,6 %	1,2 %	6,1 %	12,3 %	22,7 %	12,3 %	12,6 %	9,1 %	7,3 %	6,1 %	7,4 %	1,0 %	0,2 %
2005	1,6 %	1,1 %	5,7 %	11,7 %	22,7 %	12,6 %	12,9 %	9,5 %	7,6 %	6,2 %	7,3 %	1,0 %	0,2 %
2006	1,5 %	0,9 %	5,1 %	11,2 %	22,5 %	13,1 %	13,0 %	10,0 %	7,7 %	6,3 %	7,5 %	0,9 %	0,2 %
2007	1,5 %	0,9 %	4,8 %	10,6 %	22,3 %	13,3 %	13,0 %	10,5 %	7,8 %	6,4 %	7,6 %	1,0 %	0,3 %
2008	1,5 %	0,9 %	4,4 %	10,1 %	21,5 %	13,5 %	13,2 %	11,0 %	8,1 %	6,7 %	8,0 %	1,0 %	0,3 %
2009	1,7 %	0,8 %	4,2 %	9,4 %	21,0 %	13,8 %	13,3 %	11,2 %	8,3 %	6,8 %	8,2 %	1,0 %	0,3 %
2010	1,7 %	0,8 %	4,0 %	8,7 %	20,5 %	13,9 %	13,5 %	11,5 %	8,7 %	7,1 %	8,2 %	1,1 %	0,3 %
2011	1,9 %	0,8 %	4,0 %	8,1 %	19,6 %	14,0 %	13,7 %	11,6 %	8,8 %	7,3 %	8,7 %	1,1 %	0,3 %
2012	2,3 %	0,8 %	3,8 %	7,6 %	18,6 %	14,0 %	13,7 %	11,8 %	9,1 %	7,4 %	9,2 %	1,3 %	0,4 %
2013	2,4 %	0,8 %	3,8 %	7,2 %	17,8 %	14,0 %	13,8 %	11,8 %	9,4 %	7,8 %	9,6 %	1,3 %	0,4 %
2014	2,7 %	0,9 %	3,7 %	6,8 %	17,0 %	13,8 %	13,7 %	11,9 %	9,6 %	8,0 %	10,1 %	1,4 %	0,4 %
2015	2,9 %	0,9 %	3,7 %	6,4 %	16,1 %	13,7 %	13,9 %	12,3 %	9,7 %	8,1 %	10,4 %	1,5 %	0,4 %
2016*	3,1 %	0,9 %	3,7 %	5,9 %	14,9 %	13,6 %	13,8 %	12,2 %	10,1 %	8,4 %	11,3 %	1,5 %	0,5 %

* Siffror för 2016 är inte jämförbara med tidigare år. För 1998–2015 gäller siffrorna elever som har uppgift om både [edu13], [Pink] och [bidrag]. För 2016 krävs bara uppgift om [edu13], eftersom ingen data finns för [Pink] och [bidrag] för 2016.

Föräldrarnas utbildningsnivå

Utgångspunkten är den 7-gradiga variabeln [SUN2000_old]:

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

Variabeln [edu13] har skapats genom att ta genomsnittet av båda föräldrarnas 7-gradiga utbildningsnivå, vilket genererar en variabel med 13 steg (1: 1,5. 2: 2,5: 3: 3,5: 4: 4,5: 5: 5,5: 6: 6,5: 7).

Om uppgift saknas för en förälder baseras variabelvärdet på den förälder som det finns uppgift för.

Den 13-gradiga variabeln [edu13] har kodats om till 12 dikotoma (0, 1) variabler, för att få en bättre modellpassning.⁹⁴

Tabellerna B1.5a–b visar utbildningsvariabeln deskriptivt.⁹⁵

⁹⁴ Skillnaden mellan att använda utbildningskategorierna som en kontinuerlig variabel jämfört med att använda 12 dikotoma kategorier är mycket marginell, eftersom [edu13] uppvisar en i stort sett linjär ökning i resultat för varje stegs ökning i variabeln. För kategori 7,0 är dock det genomsnittliga betyget i stort sett på samma nivå som för kategori 6,5.

⁹⁵ Se Excel-filen för deskriptiv redovisning av utbildningsnivå efter migrationsbakgrund.

Tabell B1.5b Medelvärde och spridning för utbildningsvariabeln [edu13], samtliga elevgrupper.*

År	Antal	Medelvärde	Standardavvikelse
1998	93 392	3,56	1,34
1999	95 392	3,56	1,33
2000	98 617	3,59	1,33
2001	101 906	3,65	1,28
2002	104 346	3,67	1,27
2003	106 645	3,70	1,25
2004	113 828	3,72	1,24
2005	117 399	3,74	1,22
2006	123 815	3,77	1,22
2007	122 933	3,80	1,21
2008	120 936	3,84	1,21
2009	116 674	3,86	1,22
2010	112 199	3,89	1,22
2011	104 826	3,92	1,23
2012	98 184	3,95	1,25
2013	93 138	3,98	1,26
2014	94 472	4,01	1,28
2015	93 743	4,03	1,29
2016*	97 855*	4,07	1,31

* Siffror för 2016 är inte jämförbara med tidigare år. För 1998–2015 gäller siffrorna elever som har uppgift om både [edu13], [Pink] och [bidrag]. För 2016 krävs bara uppgift om [edu13], eftersom ingen data finns för [pink] och [bidrag] för 2016.

Som redan nämnts finns det för samtliga bakgrundsvariabler en grundläggande problematik i att de förändras över tid. Vad gäller utbildningsnivå sker en ökning över tid – allt fler har högskoleutbildning och allt färre har enbart grundskoleutbildning (tabell B1.5a). Med en dikotom variabel eller en variabel med få utbildningssteg kan en förskjutning av den genomsnittliga utbildningsnivån eventuellt påverka variabelns förmåga att förklara resultaten, beroende på var kategorigränserna går. En variabel med fler utbildningssteg minskar till viss del detta problem.

Men även med en mer fingradig variabel för utbildningsnivå kan det faktum att utbildningsnivån inte är konstant över tid eventuellt påverka analyser av betydelsen för resultat över tid. Ett sätt att hantera detta är att standardisera utbildningsvariabeln för varje år.⁹⁶ Vi gör också en kontroll av analyserna genom att ta bort de lägsta utbildningskategorierna, som eventuellt har fått en ökad betydelse för resultaten genom att de representeras av en allt mer marginell grupp.⁹⁷ En ytterligare kontroll som görs är att vikta data, så att respektive

96 Vi använder årsvis standardiserade regressionskoefficienter i alla analyser av effektstorlek på skolresultat av socioekonomisk bakgrund. Se Excel-filen.

97 Se Excel-filen.

utbildningsnivå hålls konstant över tid.⁹⁸ Slutligen delar vi också upp analysen på olika migrationskategorier för att bättre förstå utvecklingen.

Föräldrarnas inkomst

Måttet bygger på ett genomsnitt av föräldrarnas årsvisa percentilrankade samlade inkomst för de 7 sista åren fram till och med det år eleven går ut åk 9.⁹⁹ Genomsnittet bygger på de år som det finns observationer för. För vissa elever kan därmed föräldrarnas inkomst byggas på färre än 7 års uppgifter.¹⁰⁰ Om uppgift finns för en förälder men inte den andra ett år, har inkomsten för den förälder som det saknas uppgift från antagits vara 0.

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

Tabell B1.6 visar den percentilrankade inkomstvariabeln [Pink] deskriptivt.

98 Se Excelfilen.

99 Det finns nya analyser där forskare undersökt hur flera familjgenerationer påverkar elevernas skolresultat. Hällsten & Pfeffer (2017a) och (2017b).

100 Inkomstuppgifter finns dessutom bara tillgängliga fram till och med 2013. För 2014 är den senaste inkomstuppgiften därmed 1 år gammal och för 2015 är den 2 år gammal. Detta innebär sannolikt att sambandet mellan betyg och inkomst underskattas något för dessa två år (jämfört med övriga år), vilket i så fall leder till en viss underskattning av en eventuell förändring i sambandet över tid. Om man tar bort de två mest aktuella inkomståren för varje avgångsår så att tidserierna blir mer jämförbara, sjunker nivån på sambandet marginellt samtidigt som förändringen över tid ökar marginellt. Denna problematik gäller även bidragsvariabeln. Vi har ändå valt att hålla fast vid 7-årsgenomsnitt där det går (1998–2013) i stället för ett 5-årsgenomsnitt.

Tabell B1.6 Föräldrarnas percentilrankade inkomst, samtliga elevgrupper*

År	Antal	Medelvärde	Standardavvikelse
1998	93 392	50,06	27,06
1999	95 392	49,95	27,26
2000	98 617	49,97	27,26
2001	101 906	49,95	27,25
2002	104 346	49,96	27,20
2003	106 645	49,92	27,17
2004	113 828	49,91	27,18
2005	117 399	49,91	27,22
2006	123 815	49,91	27,22
2007	122 933	49,87	27,25
2008	120 936	49,79	27,33
2009	116 674	49,74	27,43
2010	112 199	49,71	27,47
2011	104 826	49,69	27,52
2012	98 184	49,65	27,64
2013	93 138	49,74	27,64
2014	94 472	49,69	28,03
2015	93 743	49,39	28,15

* Årsvis deskriptiv bild av den percentilrankade inkomstvariabeln [Pink] för samtliga elever med uppgift om föräldrars inkomst, utbildningsnivå och bidragstagande.

Observera att inkomstvariabeln [Pink] bygger på genomsnittet för de 7 senaste åren av föräldrarnas sammanlagda inkomst, där inkomsten för varje enskilt år är percentilrankad. Detta är inte samma sak som att percentilranka genomsnittet för 7 år. Därför är de årsvisa medelvärdena för [Pink] inte identiska för varje enskilt år.¹⁰¹

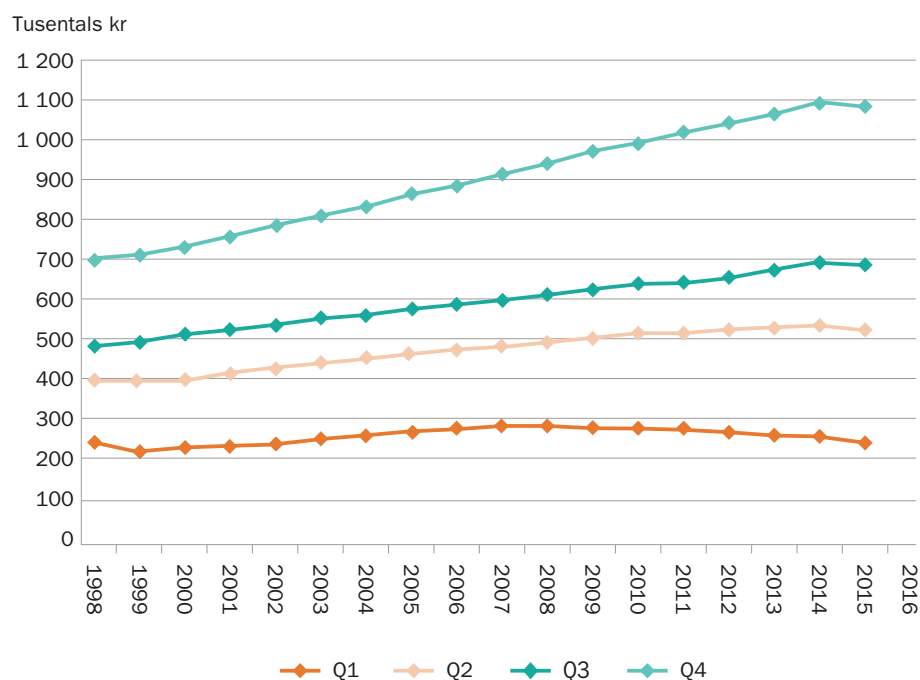
Uppgifter om föräldrars inkomster är ny jämfört med Skolverkets tidigare likvärdighetsrapporter, men har använts tidigare av svenska forskare.¹⁰² Generellt har föräldrars inkomst förändrats över tid, dels genom att realinkomsten i genomsnitt har ökat, dels genom att inkomstskillnaderna har ökat.¹⁰³ Figur B1.1 visar hur inkomsten i reella termer (fasta priser = kontrollerat för inflation) har utvecklats för olika kvartiler i inkomstfördelningen.

101 I Excel-filen redovisas den icke percentilrankade inkomsten.

102 Holmlund m.fl. (2014). Böhlmark & Holmlund (2012). Fredriksson & Vlachos (2011).

103 Se Excel-filen. Se Holmlund m.fl. (2014) s. 74 ff för en utförlig beskrivning av inkomstskillnaderna fram till år 2009. Se även SCB 2017.

Figur B1.1 Realinkomstens utveckling för olika kvartilgrupper.*



* Realinkomsten bygger på elevens föräldrars samlade inkomst i genomsnitt för de 7 senaste åren fram till avgångsåret. Q1 är de 25 procent av elever med den lägsta familjeinkomsten, medan Q4 är de 25 procent av eleverna med den högsta familjeinkomsten. Inkomsten är uttryckt i tusentals kronor och i 2016 års priser.

Percentilranking, som betyder att inkomsterna rangordnas varje år, kan minska problemet med att medelvärde och spridning förändras över tid.¹⁰⁴

Percentilranking bygger på ett antagande om att det inte är skillnader i den reella inkomsten som har betydelse, utan istället ”platsen i kön” som är det avgörande.

Under detta antagande är den percentilrankade inkomsten robust även när inkomstskillnaderna ökar. Detta innebär att om det underliggande sambandet mellan skolresultat och inkomst är konstant, så kommer inte ökade inkomstskillnader i sig att leda till förändringar i det uppmätta sambandet så länge inkomstvariabeln percentilrankas.¹⁰⁵

Men om det antagandet inte är uppfyllt är det mer problematiskt. Om betydelsen av inkomst inte bara beror på platsen i inkomstkönen utan faktiskt även på det reella avståndet i inkomst (efter kontroll för inflation), hur fungerar det percentilrankade inkomstmåttet då? Figur B1.1 visar att inkomstskillnaderna har ökat över tid, så att den lägsta kvartilens realinkomster i stort sett legat still under tidsperioden medan den översta kvartilen har haft en substantiell ökning i realinkomst. Samtidigt har resultatskillnaderna mellan dessa grupper ökat. Om det är så att elevernas resultat påverkas av föräldrarnas reella inkomst, kommer den informationen att suddas ut om inkomsten percentilrankas. Betydelsen av inkomst kommer då att överskattas, eftersom platsen i kön är oförändrad men resultatskillnaderna är större. I så fall kan det finnas skäl att använda

104 Percentilrankingen görs på samtliga elever. Detta innebär att när spridningen i den percentilrankade inkomsten analyseras uppdelat på olika migrationsgrupper kan den uppvisa förändringar över tid. För hela gruppen ska den dock inte ändras mer än ytterst marginellt.

105 Se Excel-filen.

realinkomsten i stället för den *percentilrankade* inkomsten.¹⁰⁶ Vi har genomfört analyser av betydelsen av det socioekonomiska indexet med denna realinkomstvariabel [Rink], och det visar sig då att mönstren kvarstår men att ökningen av betydelsen blir något mindre. Även för elever med svensk bakgrund kvarstår resultatet att betydelsen ökar över tid, men ökningen är mindre.¹⁰⁷

Föräldrarnas grad av bidragstagande

Vi har även med en ny variabel som mäter i vilken utsträckning någon förälder har eller har haft bidrag som huvudsaklig inkomstkälla under de senaste 7 åren.¹⁰⁸ För vart och ett av de 7 åren fram till eleven gått ut årskurs 9 anges om minst en av föräldrarna varit bidragstagare (haft bidrag som huvudsaklig inkomstkälla). Variabeln är ett genomsnitt av dessa 7 år, vilket innebär att den inte är en dikotom variabel utan varierar mellan 0 och 1. Ett värde på 1 innebär att minst en av föräldrarna haft bidrag som huvudsaklig inkomstkälla under samtliga 7 år. Ett värde på till exempel 0,42 (3/7) innebär att någon av föräldrarna varit bidragstagare under 3 av dessa 7 år.

Tabell B1.7 redovisar redovisas bidragsvariabeln deskriptivt.

Tabell B1.7 Föräldrarnas grad av bidragstagande.*

År	Antal	Medelvärde	Standardavvikelse
1998	93 392	0,177	0,301
1999	95 392	0,196	0,319
2000	98 617	0,192	0,316
2001	101 906	0,189	0,314
2002	104 346	0,180	0,305
2003	106 645	0,174	0,299
2004	113 828	0,164	0,292
2005	117 399	0,165	0,292
2006	123 815	0,161	0,291
2007	122 933	0,161	0,293
2008	120 936	0,160	0,295
2009	116 674	0,165	0,302
2010	112 199	0,165	0,306
2011	104 826	0,160	0,303
2012	98 184	0,153	0,300
2013	93 138	0,150	0,299
2014	94 472	0,149	0,308
2015	93 743	0,149	0,307

* Årsvis deskriptiv bild av variabeln [bidrag] för samtliga elever med uppgift om föräldrars inkomst, utbildningsnivå och bidragstagande.

106 Realinkomsten fås genom att dividera den nominella inkomsten för varje år med konsumentprisindex för det aktuella året, så att inkomsten uttrycks i fasta (jämförbara) priser mellan olika år.

107 Se Excel-filen.

108 Uppgifter om bidragstagande finns bara fram till och med 2013.

Kommentar: Ovanstående nationella medelvärden kan tolkas på olika sätt. Den skulle kunna tolkas så här: År 2015 har 14,9 procent av alla elever minst en förälder som haft bidrag som huvudsaklig inkomstkälla under samtliga de senaste 7 åren fram till 2015. Men eftersom bidragsvariabeln [bidrag] inte är en dikotom (0 eller 1) variabel, så kan det likaväl handla om att en större andel elever haft minst en förälder som haft bidrag som huvudsaklig inkomstkälla men bara under några av de 7 åren fram till 2015.

Bidragsvariabeln visar en nedåtgående trend i graden av bidragstagande, både för samtliga elever och för de olika migrationsgrupperna. Spridningen ligger relativt konstant, men har minskat något för elever med svensk bakgrund och ökat för utlandsfödda elever.

Socioekonomiskt index

För flera av analyserna används ett index som bygger på ovanstående tre redovisade variabler: *föräldrarnas utbildningsnivå*, *föräldrarnas inkomst* och *föräldrarnas grad av bidragstagande*. Indexet är konstruerat så att varje komponent (variabel) har en del i indexet som motsvarar dess relativa betydelse för att förklara resultaten. Exempelvis utgör föräldrarnas genomsnittliga utbildningsnivå störst andel av indexet, eftersom den variabeln är den som förklarar mest av elevernas resultat av de tre variablerna. Grad av bidragsberoende är i stället den variabel som kan förklara minst av elevers resultat, och därmed utgör den också den minsta andelen av det socioekonomiska indexet. Inkomst intar en mellanposition.¹⁰⁹ Indexet har standardiserats årsvis på samma sätt som meritvärdet.

Tabell B1.8 visar det socioekonomiska indexet deskriptivt.

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

Tabell B1.8 Socioekonomiskt index.

År	Antal	socioindex*		standardiserat socioindex**	
		Medelvärde	Standard- avvikelse	Medelvärde	Standard- avvikelse
1998	93 392	0,029	0,402	0,000	1,000
1999	95 392	0,011	0,421	0,000	1,000
2000	98 617	0,009	0,426	0,000	1,000
2001	101 906	0,012	0,437	0,000	1,000
2002	104 346	0,012	0,436	0,000	1,000
2003	106 645	0,011	0,435	0,000	1,000
2004	113 828	0,010	0,440	0,000	1,000
2005	117 399	0,009	0,440	0,000	1,000
2006	123 815	0,010	0,434	0,000	1,000
2007	122 933	0,012	0,438	0,000	1,000
2008	120 936	0,010	0,439	0,000	1,000
2009	116 674	0,012	0,452	0,000	1,000
2010	112 199	0,013	0,452	0,000	1,000
2011	104 826	0,014	0,454	0,000	1,000
2012	98 184	0,013	0,457	0,000	1,000
2013	93 138	0,022	0,458	0,000	1,000
2014	94 472	0,021	0,465	0,000	1,000
2015	93 743	0,034	0,462	0,000	1,000

* Inte standardiserat, men baserat på det standardiserade meritvärdet.

** Standardiserat och baserat på det standardiserade meritvärdet (den variabel som använts i analyserna).

Det är svårt att tolka indexets utveckling över tid i absoluta termer. Men en redovisning av hur genomsnittet och spridningen i indexet förändrats över tid för olika migrationsgrupper, ger en relativ indikation på hur olika elevgruppers förutsättningar har förändrats.¹¹⁰

Yrkesstatus

För tidsperioden 2001–2012 har vi även tillgång till en registervariabel som klassificerar föräldrarnas yrkesstatus. På grund av att den inte finns tillgänglig för hela tidsperioden 2000–2015 och att bortfallet är något större, har vi valt att inte ta med den i våra huvudresultat eftersom det skulle begränsa tidseriernas längd. Men vi har inkluderat den i separata analyser för den begränsade tidsperioden, och resultaten från dessa analyser redovisas i Excel-filen. Tabell B1.9 visar hur yrkesstatusvariabeln [hyseg] kategoriseras.

¹¹⁰ Se Excel-filen.

Tabell B1.9 Yrkesstatusvariabeln [hyseg]

Yrkesstatus (yrkesbaserad socioekonomisk grupp), 10 kategorier:
01 Höga tjänstemän och ledande befattningar
02 Kvalificerade tjänstemän
03 Övriga tjänstemän
04 Småföretagare exkl. lantbrukare
05 Lantbrukare m.fl.
06 Arbetsledare och tekniker
07 Yrkesutbildade inom handel, service och omsorg
08 Yrkesutbildade arbetare
09 Övriga arbetare
10 Ej förvärvsarbetande

Som framgår har dessa kategorier inte någon entydig hierarki. Det är inte så att elever med föräldrar som är lantbrukare skulle ha en mer gynnsam socioekonomisk bakgrund jämfört med elever vars föräldrar är yrkesutbildade arbetare. Därför använder vi denna variabel som en så kallad multipel dummyvariabel, det vill säga vi skapar 9 dummyvariabler (utifrån 10 kategorier eftersom en kategori utgör referenskategori) och som tillförs den regression varifrån det socioekonomiska indexet skapas (de predicerade värdena). När yrkesstatus tillförs det socioekonomiska indexet ökar förklaringsgraden för betydelsen av det socioekonomiska indexet för betygsresultaten (med drygt en procentenhet), men det övergripande mönstret över tid kvarstår.

B1.3 Anpassning till byte av betygssystem och betygsskala

Byte av betygssystem 1998

De resultat som beskrivs och analyseras i rapporten baseras på tidsperioden 1998–2016, vilket inte omfattar perioden med de gamla relativa betygen. Men även om inte tidserier redovisas som sträcker sig bakåt och förbi tidpunkten för övergången till det målrelaterade betygssystemet 1998, finns en risk att denna övergång ändå påverkar tidserierna efter 1998.

Ett nytt betygssystem kan leda till osäkerhet hos lärarna om hur betygssättningen ska fungera. Det skulle därför kunna vara så att lärarna inledningsvis sätter betyg mer i mitten av betygsskalan och med tiden, när de blir säkrare, använder mer av ytterlighetsbetygen. Mycket talar för att det var vad som hände när betygssystemet förändrades 1998. Det framgår till exempel av betygssättningen de första åren efter bytet. Det går att tala om en *lärareffekt* på betygen. Dessutom fanns inte fastställda betygskriterier för betyget MVG förrän 2000, något som ytterligare ökade osäkerheten bland lärarna.

Under förutsättning att en sådan lärareffekt fanns på betygen under de första åren efter att det nya betygssystemet infördes, kommer de indikatorer som presenteras i den här rapporten att påverkas på så sätt att de riskerar att inte vara jämförbara över tid. Till exempel kommer skillnaden i resultat mellan elever med en mer respektive mindre fördelaktig socioekonomisk bakgrund att minska, eftersom elever med en mindre fördelaktig socioekonomisk bakgrund tenderar att ha fler låga betyg och färre höga betyg jämfört med elever med en mer fördelaktig socioekonomisk bakgrund. Därmed kommer betydelsen av socioekonomisk bakgrund inledningsvis att underskattas, för att sedan öka när lärarna börjar använda ej godkänt och MVG i större utsträckning. Konsekvensen blir att åtminstone en del av en eventuell ökning under de allra första åren kan förklaras med lärarnas betygssättning och inte nödvändigtvis en faktisk ökad betydelse av den socioekonomiska bakgrunden.

Att försöka skatta denna *lärareffekt* på betygssättningen är enligt vår mening inte möjligt. Det är dessutom svårt att veta hur länge det tog innan betygssystemet ”satte sig”. För att i möjligaste mån inte låta denna anpassningsperiod snedvråda våra analyser över tid, kommer vi inte att beakta de första två åren med de målrelaterade betygen, dvs. 1998 och 1999. Vi kommer i rapporten visserligen att redovisa alla resultat över hela tidsperioden 1998–2016, men alla tolkningar och trendberäkningar görs med år 2000 som utgångspunkt. På så sätt menar vi att vi på ett rimligt sätt försöker ta hänsyn till, och minska effekten av, osäkerheten i betygssättning under dessa första år i våra tolkningar av likvärdighetsindikatorernas utveckling under det längre tidsperspektivet.

Byte av betygsskala 2013

Vårterminen 2013 infördes en ny och mer fingradig betygsskala. Denna förändring kan potentiellt ha påverkat möjligheten att jämföra resultaten från de två betygsperioderna (1998–2012 respektive 2013 och framåt). Att det har blivit fler betygssteg bör teoretiskt leda till en högre korrelation mellan socioekonomisk bakgrund och betyg. Detta eftersom fler betygssteg fångar upp mer av variationen i betyg mellan elever. Det kan då leda till att en ökning i

betydelsen av socioekonomisk bakgrund överskattas, eftersom den till viss del kan förklaras av bytet av betygsskala, *skaleffekt*. Den andra hypotetiska effekten är att en ny skala i kombination med ett nytt sätt att uttrycka kraven för betygsnivåerna (från betygskriterier till kunskapskrav) även det leder till en viss osäkerhet för lärarna, liknande den som sannolikt uppstod vid bytet av betygssystem 1998. Denna *lärareffekt* skulle om den existerade när betygsskalan ändrades, leda till en *lägre korrelation* mellan socioekonomisk bakgrund och betyg. Det är alltså två potentiellt motverkande effekter av bytet.

Vi har gjort flera analyser av dessa två hypoteser. Slutsatsen är att vi inte finner något som tyder på att lärarna blivit mer försiktiga att använda alla betygsstegen, även ytterlighetsbetygen, i den nya betygsskalan. Att andelen med A i betyg 2013 var betydligt lägre än vad andelen med MVG var 2012 kan helt och hållet förklaras av själva skalbytet, inte av att lärarna av försiktighet höll igen med de högsta betygen. Det är också logiskt med tanke på att den nya skalan egentligen inte är en helt ny skala, utan mer en justering av den gamla skalan, genom att två mellanbetyg – B och D tillkommit. Det går naturligtvis inte att helt utesluta att det kan ha funnits lärare som blev mer återhållsamma i betygssättningen men det är inte rimligt att tro att denna *lärareffekt* skulle vara så omfattande som den verkar ha varit 1998, när det var ett fundamentalt nytt betygssystem som infördes.

Däremot visar våra analyser att det finns en så kallad *skaleffekt* av att det har blivit fler betygssteg. Men den är inte stor. När det gäller förklaringsgraden mellan betyg och socioekonomisk bakgrund är den ungefär 0,1 till 0,5 procentenheter.¹¹¹ För att kunna göra tidserierna jämförbara över både den gamla och den nya betygsskalan, har vi justerat ned tidserierna för denna effekt.

Simulering av effekter på ny betygsskala 2013

Det här avsnittet beskriver mer i detalj hur vi har gått till väga för att skatta eventuella effekter som byte av betygsskala kan tänkas ha på några av de mest centrala likvärdighetsindikatorerna som presenteras i den här rapporten. Framför allt gäller analysen hur sambandet mellan resultat och socioekonomisk bakgrund påverkas vid byte av betygsskala.

Vi har här använt simuleringar för att försöka modellera hur ett byte av betygsskala kan påverka förklaringsgraden samt den standardiserade effektstorleken i linjär regression mellan resultatmåttens meritvärde, respektive det kombinerade betyget i matematik och engelska och den socioekonomiska indexvariabeln som används i rapporten och som bygger på de tre variablerna [edu13], [Pink] och [bidrag]. De skattade skaleffekterna har sedan använts för att justera de tidserier som presenteras i rapporten för sambandet mellan resultat och socioekonomisk bakgrund.

Metod

Utgångspunkten är ett dataset för 2012 (året före skalbytet) med faktiska data om elevers betyg och bakgrundsvariabler (föräldrarnas utbildningsnivå [edu13], inkomst [Pink], bidragstagande [bidrag], elevens kön [sex], migrationsbakgrund

¹¹¹ Skaleffekten skattas till cirka 0,1 procentenhet förrän det gäller meritvärdet och 0,5 procentenheter förrän det gäller det kombinerade betyget i matematik- och engelskaengelskabetyget.

[migswe], [migfor], [invefter7] samt om eleven är yngre eller äldre än typåldern för årskursen [elev_tidig], [elev_sen]). Bara elever med uppgift för samtliga variabler ingår i analyserna. Elever som får ett samlat NO- eller SO-betyg har uteslutits.

Linjära regressioner har skattats för vart och ett av de faktiska betygen i samtliga 16 ämnen som utfallsvariabel och samtliga nio faktiska bakgrundsvariabler.¹¹² Därifrån har de predicerade värdena och residualerna från respektive regression sparats. Dessutom har sambandet mellan samtliga ämnens residualer skattats. Sedan har slumpstermer skapats och som bygger på residualsambanden mellan ämnena i syfte att återskapa/simulera den naturliga variation som finns runt de predicerade värdena.¹¹³

En approximativt kontinuerlig skala skapas/simuleras för varje ämne genom summan av respektive predicerat värde och motsvarande simulerad slumpsterm. Detta genererar 16 simulerade betygsfördelningar på en kontinuerlig skala och som var och en kan tolkas som en underliggande *kunskapsfördelning*.

För varje simulerad kontinuerlig kunskapsfördelning (en för respektive ämne) härleds *två* separata *diskreta* betygsskalor, en som representerar den *gamla* betygsskalan (1998–2012) och en som representerar den *nya* betygsskalan (2013–), [ma_betyg_old] respektive [ma_betyg_new] enligt följande:

Tabell B1.10 Brytpunkter för respektive betygsskala.

Intervall på kontinuerlig skala	[betyg_old]	Intervall på kontinuerlig skala	[betyg_new]
Upp till 4,9999	0	Upp till 4,9999	0
5–12,4999	10	5–11,2499	10
		11,25–13,7499	12,5
12,5–17,4999	15	13,75–16,2499	15
		16,25–18,7499	17,5
17,5 och över	20	18,75 och över	20

Vi utgår alltså från den kontinuerliga kunskapsskalan och projicerar den på två olika trunkeringsregler (enligt tabell ovan), en för den gamla betygsskalan och en för den nya. Idén är att för en given identisk underliggande kunskapskala få ett betyg enligt den gamla skalan och ett enligt den nya.

Ett meritvärde respektive kombinerat betyg i matematik och engelska skapas för både den gamla och den nya betygsskalan: [mv_old], [mv_new], [maen_old], [maen_new] och dess standardiserade motsvarigheter.

Sedan utförs linjära regressioner med ovanstående utfallsvariabler och med det socioekonomiska indexet som förklarande variabel.¹¹⁴ Andelen förklarad variation (R²) samt den standardiserade effektstorleken jämförs för varje utfalls-

112 Matematik, engelska, biologi, fysik, kemi, samhällskunskap, geografi, historia, religion, bild, musik, slöjd, teknik, idrott och hälsa, hemkunskap samt svenska eller svenska som andraspråk. De förklarande variablerna utgörs av [edu13], [Pink], [bidrag], [sex], [migfor], [migswe], [invefter7], [elev_tidig] och [elev_sen].

113 Simuleringsprocessen är mer komplicerad och bygger bl.a. på en iterativ process.

114 Egentligen görs ett steg innan, där det socioekonomiska indexet skapas som det predicerade värdet i en regression med samma utfallsvariabler men med [edu13], [Pink] och [bidrag] som förklarande variabler.

mått mellan den gamla och den nya betygsskalan och skillnaden är en skattning av skaleffekten vid byte av betygsskala.

Hela simuleringsprocessen har replikerats fem gånger och den slutliga skattningen av skaleffekten utgörs av genomsnittet av dessa fem simuleringar.¹¹⁵

Observera att det socioekonomiska indexet inte är simulerat, utan utgörs av den faktiska variabeln som skattats utifrån den multipla regressionen. Utfallsvariablerna, [mv_old] och [mv_new] bygger däremot på de predicerade värdena i kombination med de simulerade slumptermerna. De utgör alltså simulerade värden. Konceptet är att utifrån en gemensam simulerad kontinuerlig skala projicera betygsgränser för de två diskreta betygsskalorna och sedan jämföra hur sambandet skiljer sig när utfallsmåttet från den gamla respektive nya skalan används i en i övrigt identisk regression. Därmed fås en skattning av den rena *skaleffekten*, dvs. den effekt som uppstår som ett resultat av den rent mekaniska övergången från en skala till en annan, allt annat lika.¹¹⁶

Resultat från simuleringarna

Tabell B1.11 Skattade skaleffekter, samtliga elever.*

		95 % konfidensintervall			
		skaleffekt	s.e.	undre	övre
standard- avvikelse	Zmaen	-0,025	0,00045	-0,026	-0,024
	Zmv	-0,015	0,00026	-0,0153	-0,014
R2	Zmaen	0,52 %	0,063 %	0,39 %	0,64 %
	Zmv	0,10 %	0,020 %	0,06 %	0,14 %
Standardi- serad effektstorlek	Zmaen	0,006	0,0002	0,005	0,0064
	Zmv	0,001	0,0007	0,000	0,003

* Standardavvikelsen avser spridningen i respektive utfallsmått för hela populationen elever. R2 (andel förklarad variation) och standardiserad effektstorlek gäller enkla regressioner med respektive utfallsmått och det socioekonomiska indexet.

Resultaten visar att:

- R2: Det uppstår en positiv ”snedvridning” vid skalbytet på 0,52 procentenheter för maen-betyget och 0,10 procentenheter för meritvärdet.
- Standardiserad effektstorlek: en positiv ”snedvridning” vid skalbytet på 0,006 för maen-betyget och 0,001 för meritvärdet.

115 Här har precis som i rapportens huvudtext betygen kodats som (0, 10, 15, 20) respektive (0, 10, 12.5, 15, 17.5, 20). Tester med den alternativa kodningen (1.5, 3, 4, 5) respektive (1.5, 3, 3.5, 4, 4.5, 5) har också genomförts men resultaten blir mycket likartade.

116 Hur väl de simulerade *gamla* betygen stämmer överens med de faktiska betygen från 2012 redovisas nedan, men en avgörande faktor är hur väl vi lyckats återskapa sambandet mellan samtliga ämnens residualer. Vi har använt en iterativ process som gett tillfredställande resultat.

Justering av tidserierna

En positiv skaleffekt innebär att även om betydelsen av socioekonomisk bakgrund inte förändrades i sak mellan 2012 och 2013 så kommer skalbytet innebära att andel förklarad variation (R²) ändå ökar med 0,10 procentenheter när betydelsen av socioekonomisk bakgrund mäts mot meritvärdet. Detta innebär att varje R²-värde i tidserierna från och med 2013 ska justeras ned med en konstant som motsvarar den skattade skaleffekten (0,10 procentenheter). För maen-betyget är motsvarande nedjustering i stället 0,52 procentenheter. Även de standardiserade effektstorlekarna justeras ned enligt samma koncept utifrån deras skattade skaleffekter.

På motsvarande sätt har även skaleffekter skattats för de enskilda migrationsgrupperna.

Lärareffekter

Lärarna agerade försiktigt när betygssystemet förändrades från det relativa till det kriterierelaterade systemet 1997–1998. Detta var ett fundamentalt byte där lärarna för första gången skulle ställa elevprestationer mot fastställda kunskapskrav/betygskriterier. Det framstår som rimligt att lärarna då valde att vara mer sparsamma med de lägsta respektive högsta betygen (ej godkänt och MVG).

Men det finns avgörande skillnader mellan det byte som skedde mellan 1997 och 1998 jämfört med det som skedde mellan 2012 och 2013. För det första som ovan nämnts byte man från ett relativt till ett kriterierelaterat system. Det senaste bytet är mer att betrakta som en finjustering av skalan i ett redan existerande kriterierelaterat system. Kunskaperna bedöms fortfarande utifrån fastställda betygskriterier eller kunskapskrav. Eftersom lärarna nu borde ha en viss vana vid ett kriterierelaterat system, borde de inte ha haft samma anledning att hålla igen som vid förra bytet, i alla fall inte i samma grad.

Dessutom är själva skalförändringen annorlunda på minst två sätt. Det ena är att den nya skalan framförallt är avsedd att nyansera gränslandet mellan G och VG respektive VG och MVG. Så även om skalan har ändrats till rena bokstavs-betyg, så känner lärarna igen sig när de ser hur betygen översätts till meritvärde-poäng, från 0, 10, 15, 20 till 0, 10, 12,5, 15, 17,5, 20. Skalan har alltså inte förändrats i grunden, utan har bara blivit mer nyanserad. Slutligen går antalet skalsteg åt motsatt håll. Mellan 1997 och 1998 minskade antalet skalsteg från 5 till 4 medan det från 2012 till 2013 gick från 4 till 6, alltså ökade skalans fingradighet i stället för att förra gången ha minskat.

Hur väl lyckas simuleringen för övergången mellan 2012 och 2013 motsvara den förändring som verkligen skedde? I tabell B1.12 jämförs de faktiska och simulerade betygsfördelningarna utifrån en av de fem simuleringarna. Eftersom den nya skalan framför allt påverkade de högre betygen och inte de lägre, kan man förvänta sig att det är där vi ska se förändringarna. Enligt tabell B1.12 visar simuleringen i engelska att andelen elever med 20 poäng ska sjunka med 6,5 procentenheter. I verkligheten sjönk den med 6,9 procentenheter. Andelen elever med 0 poäng ska enligt simuleringen vara oförändrad, men i verkligheten sjönk den med 0,4 procentenheter. I matematik fungerar simuleringen också bra men med lite större avvikelse. Andelen med 20 poäng simuleras att minska med 4,5 procentenheter, medan den faktiska nedgången stannade på 2,8 procentenheter. Andelen med 0 poäng ska vara oförändrad enligt simuleringen,

men minskade med 1,4 procentenheter i verkligheten.¹¹⁷ Observera att avvikelserna mellan den simulerade och faktiska förändringen i andel med högsta betyg går åt olika håll för matematik och engelska, vilket är en indikation på att eventuella lärareffekter (som förväntas vara återhållande) knappast kan vara av någon större betydelse. Observera också att vi inte kan utesluta att det faktiskt skett skillnader i betygen som inte beror på någon lärareffekt som kan härledas till själva bytet av betygsskala, eftersom vi jämför de faktiska betygen för både 2012 och 2013.¹¹⁸

Ur ett teoretiskt perspektiv bedömer vi att förändringen 1997–1998 var betydligt mer problematisk än motsvarande förändring 2012–2013. Den empiriska analysen stödjer de teoretiska argumenten för att ifrågasätta förekomsten av en betydande *lärareffekt*, eftersom alla faktiska förändringar i ändarna av betygsskalan 2012–2013 i stort sett kan förklaras av mekaniken i skalbytet, så som de förutspås av simuleringarna.

117 Observera att denna jämförelse bara bygger på en simulering medan vi tidigare har använt fem simuleringar. Överensstämmelsen är dock god även för övriga fyra simuleringar.

118 När vi skattar skaleffekten tar vi fram en helt teoretisk fördelning, och sedan låter vi de två trunkeringsreglerna generera två betygsutfall. Skattningen av skaleffekten kan därmed sägas vara relativt oberoende av de faktiska betygen, så länge våra simulerade *gamla* betyg någorlunda överensstämmer med de *gamla* faktiska 2012-betygen från 2012. Det går inte att skatta lärareffekten på samma sätt, utan här kan vi bara subjektivt bedöma att skalbytet på det stora hela förutsäger det utfall vi fick 2013.

Tabell B1.12 Jämförelse av simulerade och faktiska betyg för både gamla och nya betygsskalan (2012 respektive 2013), matematik och engelska.

Betygspoäng	ma_2012_gamla skalan_faktisk		ma_gamla skalan_simulerad	
	Antal	Procent	Antal	Procent
,0	7 596	9,7	7 414	9,5
10,0	38 265	48,8	35 875	45,7
15,0	21 958	28,0	24 552	31,3
20,0	10 621	13,5	10 599	13,5
Total	78 440	100,0	78 440	100,0

Betygspoäng	ma_2013_nya skalan_faktisk		ma_nya skalan_simulerad	
	Antal	Procent	Antal	Procent
,00	7 744	8,3	7 414	9,5
10,00	29 502	31,7	28 267	36,0
12,50	17 058	18,4	14 960	19,1
15,00	17 422	18,7	12 380	15,8
17,50	11 265	12,1	8 348	10,6
20,00	9 937	10,7	7 071	9,0
Total	92 928	100,0	78 440	100,0

Betygspoäng	en_2012_gamla skalan_faktisk		en_gamla skalan_simulerad	
	Antal	Procent	Antal	Procent
,0	5 499	7,0	4 898	6,2
10,0	24 719	31,5	28 346	36,1
15,0	30 668	39,1	26 721	34,1
20,0	17 554	22,4	18 475	23,6
Total	78 440	100,0	78 440	100,0

Betygspoäng	en_2013_nya skalan_faktisk		en_nya skalan_simulerad	
	Antal	Procent	Antal	Procent
,00	6 089	6,6	4 898	6,2
10,00	14 432	15,5	21 648	27,6
12,50	15 994	17,2	13 667	17,4
15,00	23 972	25,8	13 729	17,5
17,50	18 041	19,4	11 119	14,2
20,00	14 400	15,5	13 379	17,1
Total	92 928	100,0	78 440	100,0

B1.4 Takeffekter

Även om betygen standardiseras årsvis för att göra dem mer jämförbara, löser det inte alla problem med betygsinflation. Det kan fortfarande finnas problem med så kallade *takeffekter*. I det här fallet betyder det att en allt högre andel elever får det högsta betyget och därmed ”slår i taket”, vilket leder till att betygsskalan blir sämre på att diskriminera mellan olika kunskapsnivåer. Till exempel har andelen elever som fått MVG i stort sett fördubblats i ämnena teknik, slöjd, musik, idrott och hälsa samt bild mellan 1998 och 2012, vilket framgår av tabell B1.13.

Tabell B1.13 Andel elever som får högsta betyg för respektive ämne 1998 jämfört med 2012.

Ämne	Andel elever med betyget MVG Sorterat efter procentuell ökning		
	1998	2012	Förändring 1998-2012 %
Teknik	7,1	15	111
Slöjd	10,9	21,4	96
Musik	11,3	21,2	88
Idrott och hälsa	14,5	26,7	84
Bild	11,8	21,2	80
Engelska	13,2	22,6	71
Hem- och konsumentkunskap	13,9	23,6	70
Geografi	10,6	16,9	59
NO (samlad betyg)	10,6	16,2	53
Samhällskunskap	11,7	17,5	50
SO (samlad betyg)	14,5	21,5	48
Religionskunskap	12,5	18,4	47
Historia	13,5	19,5	44
Matematik	9,7	13,7	41
Svenska	12,3	16,7	36
Svenska som andraspråk	4,1	5,2	27
Biologi	12,9	16	24
Modersmål	36,7	45,3	23
Fysik	11,9	14,4	21
Kemi	11,2	12,7	13

Vad som händer är att takeffekten trycker ihop den totala variationen i ämnena när allt fler elever ”slår i taket”. Variationen bland de mest högpresterande eleverna minskar över tid, när gruppen MVG-elever blir allt större. Effekten blir större för enskilda ämnen än för meritvärdet som består av 16 ämnen. Betygsinflationen kan också få en effekt om andelen elever som får ej godkänt i betyg minskar och fler får G i betyg. Detta eftersom steget mellan ej godkänt och G är 10 poäng, och därmed större jämfört med övriga betygssteg där avståndet är 5 poäng. Så förhåller det sig i viss mån när det gäller elever födda i Sverige även om förändringen är måttlig. För elever som invandrat efter skolstart har det tvärtom blivit betydligt fler som får ej godkänt (numera F) över tid.¹¹⁹ Vi har gjort beräkningar på hur dessa betygsförändringar påverkat våra analyser, och bedömer att effekten är begränsad och inte ändrar slutsatserna.¹²⁰

119 Skolverket 2016b s. 25.

120 Se Excel-filen.

B1.5 Förändring av antalet skolor och byte av skolkod

Förändring av antalet skolor

Tabell B1.14 visar att antalet skolor med årskurs 9-elever ökade från 1 140 skolor år 1998 till som mest 1 694 skolor år 2014 och 1 675 skolor år 2016, en ökning med nästan 50 procent. Under samma tidsperiod har antalet elever ökat fram till 2006 men sedan åter minskat, och ligger på ungefär samma nivå 2016 som 1998. Detta innebär att det genomsnittliga antalet årskurs 9-elever per skola, eller skostorleken, har minskat från 85 årskurs 9-elever per skola till 60 elever sett över hela tidsperioden.

En fråga som väckts är hur detta påverkar mellanskolvariationen? Mer specifikt, kan den ökade mellanskolvariationen som skett mellan 1998 och 2016 delvis förklaras med att skolorna blivit mindre och därmed lett till fler skolor med extrema värden? Logiken bygger på att ju mindre en skola är, desto mer känsligt blir skolmedelvärdet för enskilda elevers resultat. Detta kan leda till att det blir fler skolor med både ovanligt höga och låga resultat.

Tabell B1.14 Antal skolor och genomsnittligt elevantal per skola.

År	Antal skolor (med årskurs 9-elever)	Genomsnittligt antal årskurs 9-elever per skola	Antal elever per avgångsår
1998	1 140	85	96 926
1999	1 165	83	96 257
2000	1 221	81	99 366
2001	1 267	82	104 099
2002	1 335	80	106 560
2003	1 410	77	108 684
2004	1 488	78	115 493
2005	1 533	78	118 996
2006	1 582	79	125 546
2007	1 601	78	124 800
2008	1 624	75	122 549
2009	1 654	72	118 273
2010	1 632	70	113 909
2011	1 642	65	106 506
2012	1 631	61	99 844
2013	1 659	57	95 129
2014	1 694	57	96 619
2015	1 684	57	96 459
2016	1 675	60	100 238

I tabell B1.15 har skolorna kategoriserats efter elevstorlek, och tabellen visar hur proportionerna mellan små och stora skolor förändrats mellan 1998 och 2012. De allra minsta skolorna (<20 elever) har blivit dubbelt så många, men det är de ”halvsmå” skolorna (20–39 elever) som ökat mest, från 96 till 300.

Tabell B1.15 Antal skolor och elever grupperat efter skolstorlek (4 kategorier), faktiska data, 1998 och 2012.

År	Skolstorlek	Antal skolor	Andel skolor	Antal elever	Andel elever
1998	< 20 elever	157	13,5 %	1 164	1,2 %
	20-39 elever	96	8,3 %	2 815	2,9 %
	40-59 elever	139	12,0 %	7 155	7,3 %
	60 elever eller fler	771	66,3 %	86 227	88,6 %
	Totalt	1163	100,0 %	97 361	100,0 %
2012	< 20 elever	316	19,1 %	2 904	2,9 %
	20-39 elever	300	18,1 %	8 534	8,5 %
	40-59 elever	285	17,2 %	14 114	14,0 %
	60 elever eller fler	757	45,7 %	75 011	74,6 %
	Totalt	1658	100,0 %	100 563	100,0 %

För att besvara frågan om hur mellanskolvariationen påverkas av den här utvecklingen, har en omfattande simuleringsstudie genomförts. Syftet är att efterlikna en situation där antalet skolor ökar men det totala elevantalet är konstant, och dessutom att förändringen efterliknar de förändrade proportionerna mellan små och stora skolor som skett i verkligheten enligt tabell B1.15.

Resultatet av dessa simuleringsanalyser visar två saker:

1. Om data inte viktas för skolstorlek, kommer ökningen av antalet skolor leda till en *minskad* mellanskolvariation. Men effekten är inte stor, cirka 0,2–0,3 procentenheter.
2. Om data viktas för skolstorlek, kommer ökningen i antalet skolor inte att påverka mellanskolvariationen.

Slutsatsen är att så länge data viktas för skolstorlek är variansdekomposition, som är den metod som används för att mäta mellanskolvariation, robust för förändringar i skolstorlek och antalet skolor.

Nedan följer en mer detaljerad redovisning av hur simuleringsstudien genomförts och dess resultat.

För att försöka besvara frågan har vi utgått från 2006 års faktiska betygsdata från årskurs 9. Datasetet har trimmats ned till 1 000 skolor (slumpmässigt urval med en ny skolkod), och sedan har skolor med mindre än 10 elever tagits bort. Detta leder till ett dataset med 78 918 elever i 912 skolor. Utifrån detta dataset beräknades mellanskolvariationen först för samtliga skolor och elever (U0 oviktat respektive U0 viktat), vilket då speglar hela populationen innan någon förändring.

För att på bästa sätt försöka simulera den utveckling som skett mellan 1998 och 2012 (se tabell B1.15) har simuleringen genomförts på följande sätt:

Utifrån de kategorier som redovisas i tabell B1.15 har 50 procent av skolorna i de tre största kategorierna valts ut slumpmässigt. Dessa skolor har sedan delats upp i två lika stora skolor (slumpmässig fördelning av eleverna). Sedan har mellanskolvariationen beräknats för den nya uppdelningen av skolor. Observera att med denna ansats är den totala elevpopulationen oförändrad, men antalet skolor ökar avsevärt och den genomsnittliga skolstorleken minskar. Denna

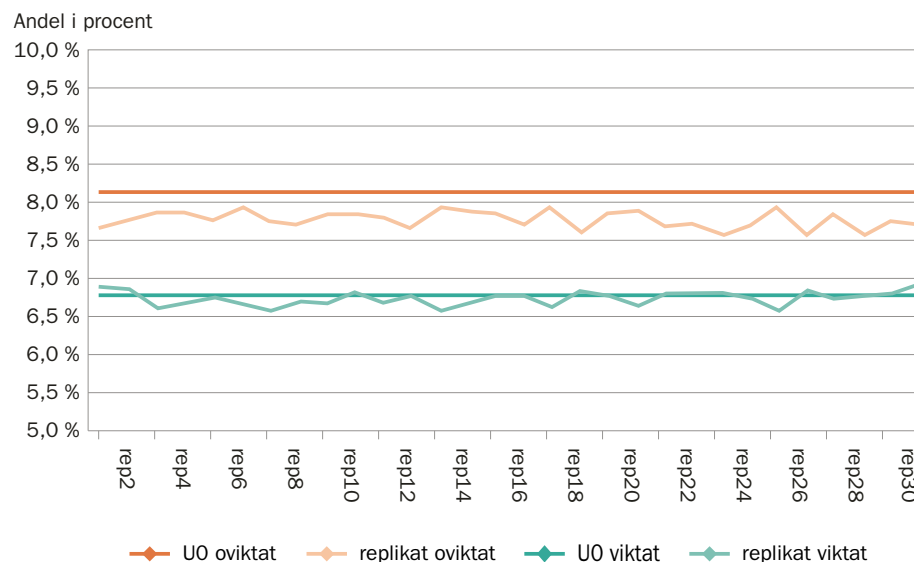
slumpmässiga splittring av skolor, liksom den slumpmässiga omfördelningen av elever i dessa skolor, har sedan upprepats 30 gånger (30 replikat).

Tabell B1.16 visar hur antalet skolor och elever i skolor förändras enligt det första replikatet. Antalet skolor ökar från 912 till 1 342, och antalet skolor i kategori 1 och 2 (de minsta skolorna) tredubblas. Antalet skolor i den största kategorin minskar från 585 till 530. En liknande utveckling syns när andelen elever i olika skolkategorier analyseras – andelen elever i kategori 1 och 2 tredubblas och andelen elever i kategori 3 fördubblas. Andelen elever i de största skolorna minskar från 85 till 67 procent. Detta motsvarar väl den faktiska utveckling som skett mellan 1998 och 2012 (enligt tabell B1.15).

Tabell B1.16 Antal skolor och elever grupperat efter skolstorlek, simulerade data, före respektive efter omfördelning.

År	Skolstorlek	Antal skolor	Andel skolor	Antal elever	Andel elever
Före omfördelning	< 20 elever	60	6,6 %	892	1,1 %
	20-39 elever	109	12,0 %	3 137	4,0 %
	40-59 elever	158	17,3 %	7 713	9,8 %
	60 elever eller fler	585	64,1 %	67 176	85,1 %
	Totalt	912	100,0 %	97 361	100,0 %
Efter omfördelning (rep 1)	< 20 elever	192	14,3 %	2 762	3,5 %
	20-39 elever	336	25,0 %	9 678	12,3 %
	40-59 elever	284	21,2 %	13 946	17,7 %
	60 elever eller fler	530	39,5 %	52 532	66,6 %
	Totalt	1 342	100,0 %	100 563	100,0 %

Figur B1.2 Mellanskolvariation för 30 oberoende replikat före respektive efter omfördelning, oviktat respektive viktat.*



* UO oviktat respektive UO viktat är den mellanskolvariation som skattats före omfördelningen och representeras av de horisontella linjerna, och därmed utgör referenslinjer för mellanskolvariationen. Rep oviktat respektive rep viktat utgör den skattade mellanskolvariationen (med och utan viktning) för vart och ett av de 30 replikaten efter omfördelning (oberoende simuleringar).

I figur B1.17 är mellanskolvariationen före omfördelning 8,1 procent för oviktat respektive 6,7 procent för viktat. Men för de oviktade replikaten är mellanskolvariationen systematiskt lägre än innan omfördelningen, dvs. utvecklingen mot fler och mindre skolor har enligt de oviktade analyserna lett till att mellanskolvariationen blir lägre, inte högre som hypotesen antydde. När i stället data viktas (efter skolstorlek) visar de 30 replikaten på en identisk mellanskolvariation innan omfördelningen. Det betyder att det inte har någon som helst effekt på skattningen av mellanskolvariationen att antalet skolor har ökat och den genomsnittliga skolstorleken blivit mindre. Metoden som används för att skatta mellanskolvariationen är variansdekomposition och simuleringen visar att denna metod är robust för förändringar i skolstorlek och antalet skolor så länge data viktas efter skolstorlek.

Varför blir den oviktade mellanskolvariationen lägre efter omfördelningen? Det visar sig att spridningen mellan skolors betygsmedelvärden är större bland små skolor, vilket är helt logiskt eftersom enskilda elevers resultat får större genomslag på medelvärdet i mindre skolor. När de stora skolorna delas ökar deras relativa vikt (det blir fler av dem). Men deras spridning ökar inte tillräckligt mycket, trots att de blir mindre, för att öka den totala spridningen, vilket beror på att de delats upp slumpmässigt. Hade de däremot delats upp mer systematiskt, hade vi fått annorlunda resultat. Men om eleverna delas upp mer systematiskt så är det en form av skolegregation utifrån resultat, och hela idén med denna simulering är att skatta effekten av minskade skolor som uppstår utan att eleverna delas upp systematiskt efter resultaten.

Ändrad skolkod

Från och med 2014 slutade Skolverket att registrera skolornas så kallade skolkod och övergick till ett nytt indentifikationssystem som går under namnet skolenhetskod. Den här förändringen kan potentiellt försämra jämförbarheten för de tidserier som mäter både mellanskolvariation och skolegregation. Teoretiskt innebär skolenhetskoden att det inte är enskilda skolor som nödvändigtvis identifieras utan skolenheter, och där en skolenhet inte nödvändigtvis är identiskt med den gamla skolan enligt skolkoden. En skola skulle kunna utgöras av flera skolenheter.

Men en närmare analys av åren för övergången visar att problemet antagligen inte har någon praktisk effekt på våra mått, speciellt som förra avsnittet också visat att variansdekomposition är ett robust mått när det gäller förändringar i antalet skolor och skolstorlek (så länge uppdelningen sker slumpmässigt). Tabell B1.18 visar hur de olika koderna kan jämföras för de år de överlappar.

Tabell B1.17 Antalet skolor beroende på val av skolidentifikationsvariabel.

	sorterat och aggregerat på:			
	[skolkod]	[skolid]	[skolenkod]	[skom]+[snamn]
2013	1 657	1657	1	1 657
2014	1	1611	1 697	1 697
2015	1	1518	1 685	1 685
2016	1 464	1461	1 685	1 685

Tabellen visar att för 2013 stämmer [skolkod] och [skolid] överens, två identifikationsvariabler som slutade att tilldelas till nya skolor efter 2013. Genom att dessutom införa ett eget identifikationssystem som bygger på skolans namn inom respektive kommun (kolumnen längst till höger) framgår hur vanligt det är att en skola med ett givet namn består av flera skolenheter. Tabellen visar att för 2013 stämmer antalet skolor överens, oavsett om de identifieras med [skolkod] eller [skom + snamn]. För 2014 övergår vi till skolenhetskod [skolenkod], och även då stämmer antalet skolor överens med [skom+snamn]. Detta gäller även för 2015 och 2016. Eftersom variabeln [skom+snamn] ser ut att fungera som ett substitut för både [skolkod] och [skolenkod], verkar det inte som att skolorna har delat upp sig efter skolenheter i någon märkbar utsträckning. En invändning är att de då också skulle ha kunnat ändrat sitt skolnamn, men en genomgång av ett urval av skolor över flera år ger ingen antydning om att så skulle vara fallet.

Att antalet skolor bara är 1 464 enligt [skolkod] år 2016 men 1 685 enligt [skolenkod] beror på att alla nytillkomna skolor bara tilldelas en skolenhetskod men ingen skolkod efter 2013.¹²¹

Slutsatsen är att vi kan generera en egen påhittad skolenhetskod för 2013 som är kompatibel med skolenhetskoden för 2014–2016. Det spelar ingen roll att den inte stämmer överens för en enskild skola, eftersom samtliga analyser genomförs separat för varje år. Det viktiga är att antalet skolor inte skiljer sig drastiskt beroende på vilken identifikationskod som används.

121 Variabeln [skolkod] finns kvar i registren, men uppdateras inte med för nya skolor.

B1.6 Statistiska metoder som används i analyserna

Betydelsen av socioekonomisk bakgrund

Sambandet mellan betygsresultat och socioekonomisk bakgrund mäts med regressionsanalys (OLS).

Eftersom den socioekonomiska bakgrunden mäts med ett socioekonomiskt index, utgörs regressionsanalysen av en enkel linjär regression. För enkel linjär regression finns ett samband mellan andelen förklarad variation och den standardiserade effektstorleken:

$$R^2 = \beta^2$$

R^2 = andel förklarad variation

β = standardiserad effektstorlek.

Om till exempel R^2 ökar över tid, så kommer även den standardiserade effektstorleken att öka över tid. Vi har valt att även redovisa den standardiserade effektstorleken i Excel-filen för de läsare som är intresserade av det måttet.

Skolsegregation och mellanskolvariation

Skolsegregation och mellanskolvariation mäts med variansdekomposition. Variansdekomposition innebär att man jämför hur stor andel av en variabels totala variation (varians) på elevnivå som kan förklaras av variation (varians) i samma variabel på skolnivå, dvs. hur mycket av till exempel den totala variationen (variansen) i meritvärde som kan förklaras av variation (varians) i skolornas genomsnittliga meritvärde.

$$\text{Mellanskolvariation} = \frac{\text{Variation mellan skolor}}{\text{Total variation}}$$

Exempel: Om det finns en viss variation (varians) i meritvärde *inom* skolor samtidigt som alla skolor har samma genomsnittliga meritvärde blir mellanskolvariationen noll, eftersom ingenting av variationen (variansen) i elevernas meritvärde kan förklaras av variation (varians) i skolornas genomsnittliga meritvärde. Om däremot alla elever på varje skola har samma meritvärde så att variationen *inom* skolor är noll men att skolorna skiljer i genomsnittligt meritvärde, kan hela (100 %) av den totala variationen (variansen) i meritvärde förklaras av variation (varians) *mellan* skolorna i meritvärde.

Observera att variansdekomposition är ett relativt mått, eftersom den sätter variationen mellan skolor i proportion till den totala variationen. Om till exempel den totala variationen i meritvärde ökar men variationen mellan skolor är konstant kommer mellanskolvariationen att minska, eftersom variationen mellan skolor nu utgör en mindre proportion i förhållande till den totala variationen. I likvärdighetsrapporten 2012 användes ett absolut mått på skolsegregation, standardavvikelsen i skolornas genomsnittliga sammansättning.¹²²

122 Skolverket (2012) kap 5.3.

Syskonkorrelationer

Även när vi analyserar syskonkorrelationer använder vi metoden variansdekomposition. Skillnaden är att vi då delar upp den totala variationen (variansen) i resultat (tex. meritvärde) i variation *inom* familjer (mellan syskon inom samma familj) och variation *mellan* familjer (mellan olika syskonpar eller syskongrupper). Syskonkorrelationen kan således tolkas som resultatvariationen mellan familjer som andel av den totala resultatvariationen, där den totala resultatvariationen är lika med summan av resultatvariation mellan familjer och resultatvariation inom familjer.

Skolnivåeffekter

Skolnivåeffekterna beräknas på två olika sätt i rapporten. I huvudtexten redovisas en förenklad variant som fångar upp det totala genomslaget av skolnivåeffekten. I bilagan redovisas en mer konventionell variant som bara fångar upp skolnivåeffekten per enhet i förändrad skolsammansättning.

1. Det totala genomslaget av skolnivåeffekten

Här kategoriseras eleverna relativt sin socioekonomiska bakgrund enligt det socioekonomiska indexet i 3 kategorier; A, B respektive C, där A utgör de 25 procent elever med den mest fördelaktiga socioekonomiska bakgrunden, B de 50 procent elever i mitten av den socioekonomiska fördelningen samt C de 25 procent elever med den minst gynnsamma socioekonomiska bakgrunden. På samma sätt kategoriseras skolorna efter sin genomsnittliga sammansättning enligt samma socioekonomiska index i A, B respektive C-skolor. Därefter beräknas skolnivåeffekten för respektive elevgrupp som differensen i resultat mellan de elever som går på A-skolor respektive C-skolor utifrån det genomsnittliga meritvärdet för den specifika elevgruppen. Vi får således en skolnivåeffekt för C-elever som endast baseras på resultaten för C-elever, en för B-elever samt en för A-elever. Beräkningarna är gjorda i SPSS.

2. Den konventionella skolnivåeffekten

Det mest konventionella sättet att beräkna skolnivåeffekten är att i en flernivåmodell beräkna den skattade koefficienten för skolans socioekonomiska sammansättning (det socioekonomiska indexet på skolnivå).

Modell:

Elevnivå:

$$[\text{Meritv}] = b_0 + b_1[\text{sex}] + b_2[\text{migswe}] + b_3[\text{migfor}] + b_4[\text{inv_efter7}] + b_5[\text{elev_tidig}] + b_6[\text{elev_sen}] + b_7[\text{socioindex}] + r$$

Skolnivå:

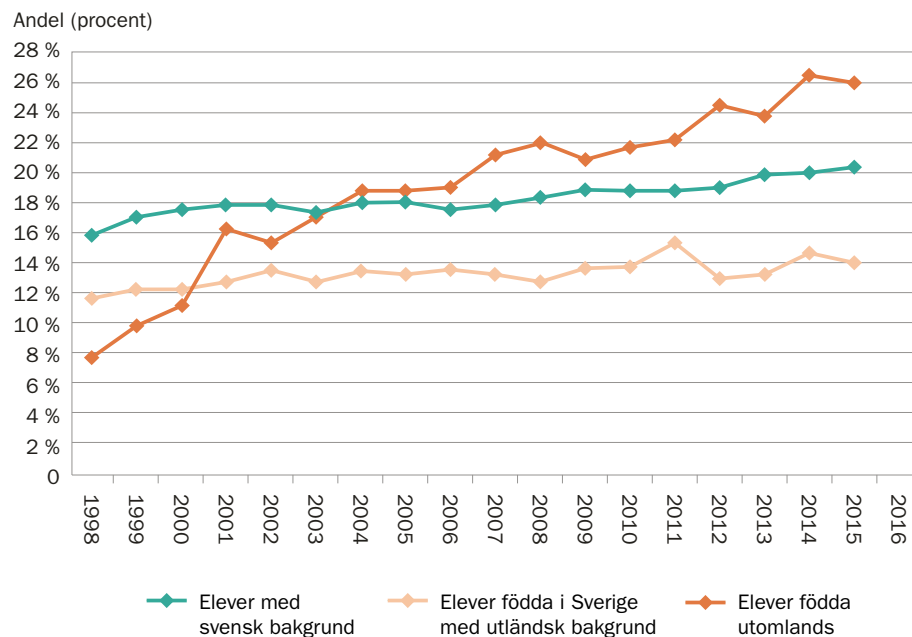
$$b_0 = g_0 + g_1[\text{mean_socioindex}] + u$$

Skolnivåeffekten utgörs av den skattade g_1 . Beräkningarna är gjorda i hlm och viktade för skolstorlek.

Bilaga B2

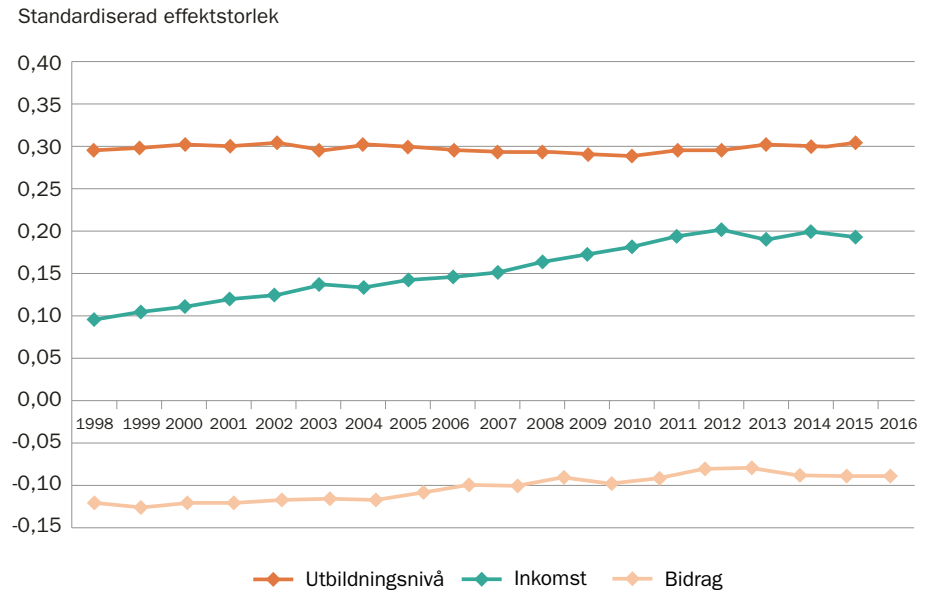
B2.1 Kompletterande analyser för betydelsen av socioekonomisk bakgrund

Figur B2.1 Betydelsen av socioekonomisk bakgrund för betygen i matematik och engelska, uttryckt som andel förklarad variation, uppdelat på migrationsbakgrund.*



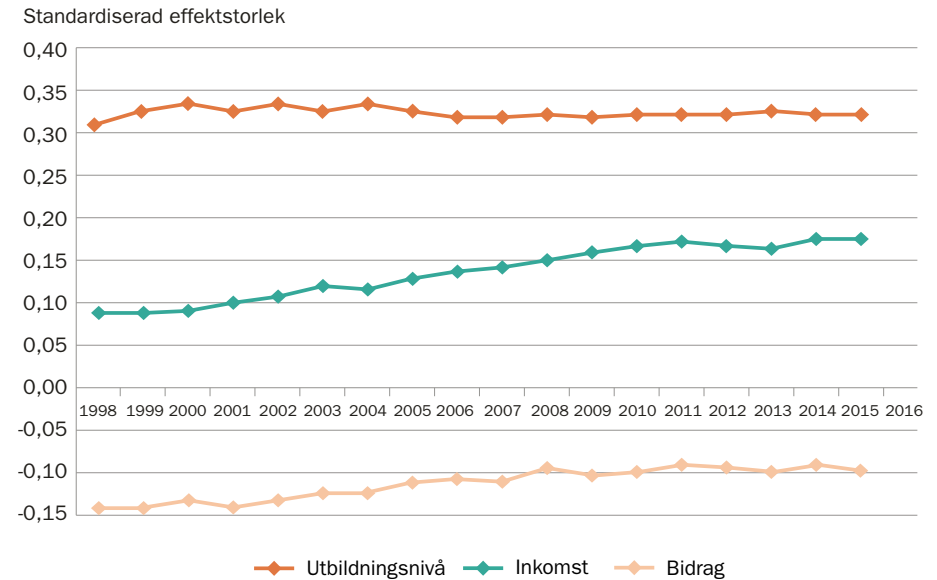
* Betydelsen mäts med andel förklarad variation (R²) i enkel regression, med det årsvis standardiserade betyget i matematik och engelska som utfallsvariabel och det socioekonomiska indexet som förklarande variabel. Tidserierna är justerade för bytet av betygsskala 2013. I analyserna ingår bara elever med personnummer och där det finns uppgift om föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet).

Figur B2.2 Betydelsen av socioekonomisk bakgrund för meritvärdet, uttryckt som effektstorlek, utifrån de bakgrundsvariabler som utgör det socioekonomiska indexet, samtliga elevgrupper.*

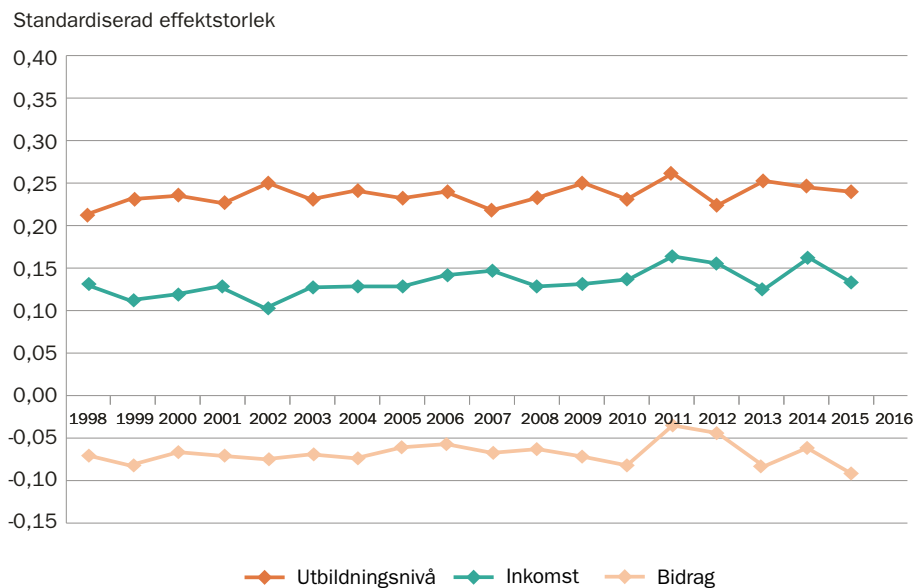


* Betydelsen mäts med standardiserad effektstorlek i multipel regression, med det årsvis standardiserade meritvärdet som utfallsvariabel och de variabler som utgör det socioekonomiska indexet, föräldrarnas utbildningsnivå, inkomst och grad av bidragstagande som förklarande variabler. Observera att tidserierna inte är justerade för bytet av betygsskala 2013. I analyserna ingår bara elever med personnummer och där det finns uppgift om föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet).

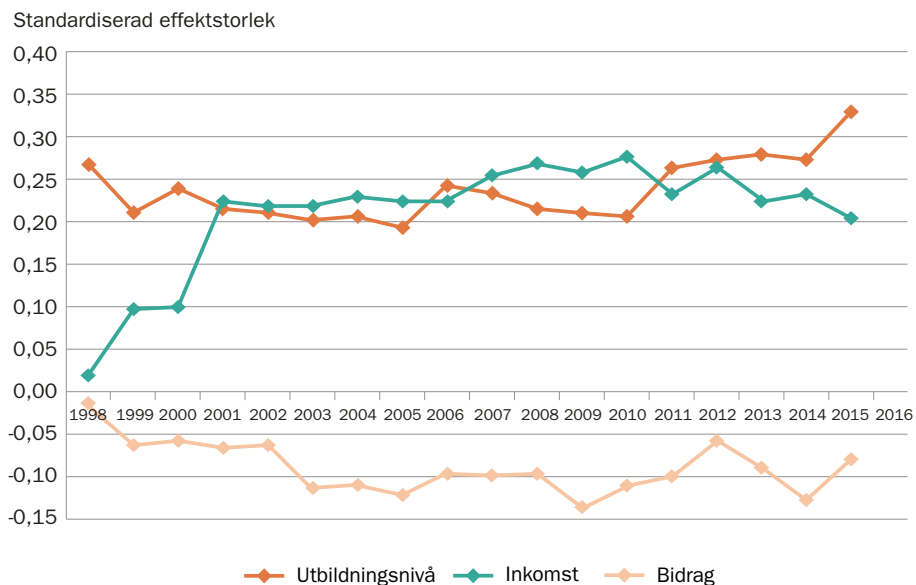
Figur B2.3a Betydelsen av socioekonomisk bakgrund för meritvärdet, uttryckt som effektstorlek, utifrån de bakgrundsvariabler som utgör det socioekonomiska indexet, Elever med svensk bakgrund



Figur B2.3b Betydelsen av socioekonomisk bakgrund för meritvärdet, uttryckt som effektstorlek, utifrån de bakgrundsvariabler som utgör det socioekonomiska indexet, Elever födda i Sverige med utländsk bakgrund.



Figur B2.3c Betydelsen av socioekonomisk bakgrund för meritvärdet, uttryckt som effektstorlek, utifrån de bakgrundsvariabler som utgör det socioekonomiska indexet, Utlandsfödda elever.



B2.2 Syskonkorrelationer

Syskonkorrelationer är ett mått på familjebakgrundens betydelse som tidigare använts av forskare.¹²³ Måttet fångar upp alla aspekter som syskon delar, som föräldrarnas utbildningsnivå (och andra aspekter av det som brukar refereras till som socioekonomisk bakgrund), grannskapseffekter, skoleffekter (i den mån syskonen går i samma skola), genetiska faktorer och uppfostran. Måttet fångar däremot inte upp saker som skillnader i hur föräldrarna behandlar (eller investerar i) sina barn eller genetiska faktorer som skiljer sig mellan syskonen.

Själva beräkningen av syskonkorrelationer bygger på variansdekomposition, enligt samma princip som vid beräkning av mellanskolvariation. Skillnaden här är att vi undersöker hur mycket av den totala variationen i exempelvis meritvärde som kan förklaras av att meritvärdet skiljer sig mellan syskonpar/syskongrupper, eller familjer.¹²⁴ En ekvivalent benämning av syskonkorrelation vore alltså ”mellanfamiljvariation” enligt samma koncept som mellanskolvariation. Vi har ändå valt att följa den konventionella benämningen. Tolkningen är fortfarande hur mycket av den totala betygsvariationen bland elever som kan förklaras av vilken familj eleverna tillhör.

I den här rapporten används 3 år långa *fönster* för att beräkna syskonkorrelationerna (men vi gör kontroller även för 5-årsfönster).¹²⁵ Det innebär att vi i strikt mening inte kan beräkna hur stor syskonkorrelationen är ett givet kalenderår, utan varje skattning av syskonkorrelationen görs för ett fönster som här består av 3 sammanhängande kalenderår och där samtliga syskonpar som kan identifieras inom detta fönster ingår i analysen. Sedan flyttas fönstret ett år framåt, så att det tidigaste året tas bort och ett senare år läggs till.¹²⁶ Men vi benämner fönstret efter det år som är i mitten av fönstret.

Detta innebär bland annat att syskonpar där det skiljer mer än 2 kalenderår inte ingår i beräkningarna. En annan konsekvens är att tvillingpar kommer att ingå i 3 olika fönster. Detta leder till en viss tröghet eller stabilitet i eventuella förändringar över tid, vilket är både en styrka och en svaghet. En ytterligare konsekvens är att det sista fönstret vi kan få jämförbara skattningar ifrån är centrerat kring 2014, även fast vi har uppgifter även för 2015.¹²⁷

Det finns för- och nackdelar med att använda syskonkorrelationer som mått på betydelsen av socioekonomisk bakgrund över tid – jämfört med de ”traditionella” måtten som vi redovisat ovan.¹²⁸ Syskonkorrelationerna är inte på samma vis beroende av förändringar över tid i bakgrundsvariabler, eftersom de inte används. Å andra sidan är syskonkorrelationerna, precis som de traditionella

123 Björklund m.fl. (2003). Böhlmark & Holmlund (2012). Holmlund m.fl. (2014).

124 Observera att det inte bara är par av syskon som ingår i analysen, utan även syskongrupper. Till exempel kan en syskongrupp utgöras av ett syskon och ett tvillingpar, så länge de är födda inom samma 3-årsfönster.

125 I likhet med Böhlmark & Holmlund (2012) s. 46 och Holmlund m.fl. (2014) så begränsar vi analysen till helsyskon, även om det, med tanke på att vi använder ”treårsfönster”, inte är många elever som har halvsyskon eller inte bor med någon av sina biologiska föräldrar, mellan 1,5–3 procent per fönster. När vi utvidgar analysen till ”femårsfönster” och fler halvsyskon kommer med, så förändras inte resultaten i någon avgörande grad.

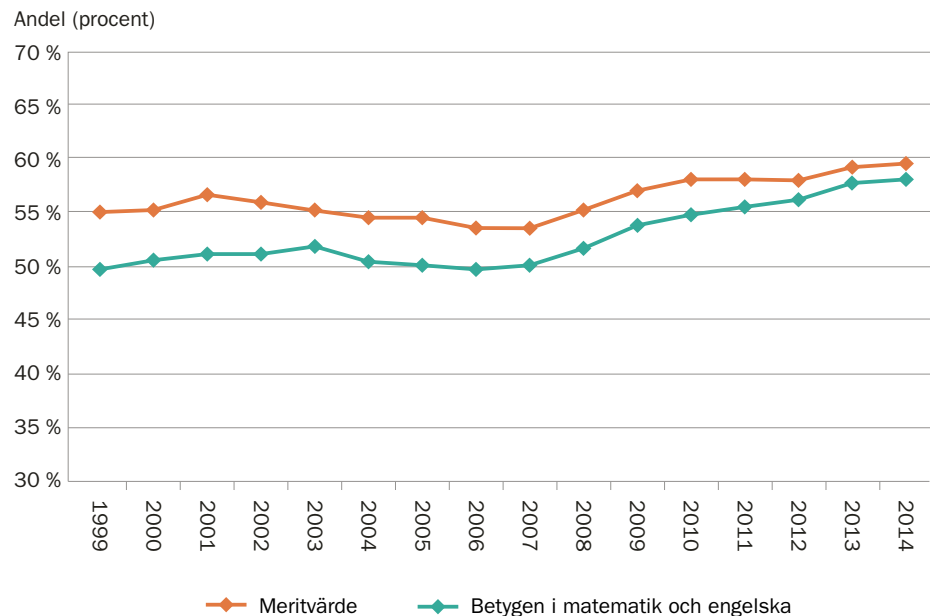
126 Om fönster 1 innefattar åren 1998, 1999 och 2000, så innefattar fönster 2 åren 1999, 2000 och 2001, osv.

127 Om vi inkluderade 2015- fönstret skulle det bara innehålla syskonpar där det skiljer högst ett kalenderår (2014 och 2015), vilket skulle leda till en överskattning jämfört med övriga fönster, eftersom syskon som är födda närmare varandra i tid tenderar att ha mer lika betyg.

128 Holmlund m.fl. (2014) s. 153 ff. Böhlmark & Holmlund (2012) s. 46 f.

måtten, känsliga för om det sker förändringar i utfallsmåttet, det vill säga betygen. En uppenbar nackdel är att måttet inte kan beräknas utifrån samtliga elever utan bara på elever med syskon inom en ganska snäv tidsram. Förändringar i korrelationerna kan också vara svårtolkade vad de i så fall beror på. Vi använder två metoder för att närma oss en tolkning av förändringar i syskonkorrelationerna. För det första gör vi en kontroll för socioekonomisk bakgrund. Då vet vi i vart fall om det är förändringar i den socioekonomiska bakgrunden (så som vi mäter den i de konventionella analyserna) som ligger bakom.¹²⁹ Den andra metoden, som oss veterligen inte gjorts tidigare, är att dela upp syskonkorrelationerna på migrationsbakgrund.

Figur B2.4 Syskonkorrelationer med avseende på meritvärdet respektive betygen i matematik- och engelska, samtliga elevgrupper.



* Syskonkorrelationen mäts med variansdekomposition. Andelen av den totala variationen i det årsvis standardiserade meritvärdet som kan förklaras av variation mellan syskonpar/syskongrupper. Observera att tidsserierna inte är justerade för bytet av betygsskala 2013. I analyserna ingår bara elever med personnummer och där det finns uppgift om föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet). Dessutom ingår bara elever som har minst ett syskon inom 2 kalenderår.

Figur B2.4 visar på en ökad syskonkorrelation för både meritvärdet och för betygen i matematik och engelska. Men ökningen mellan 2000 och 2014 är större för det senare måttet, knappt 8 procentenheter jämfört med drygt 4 procentenheter. Ökningen sker för båda måtten från och med ungefär 2007.

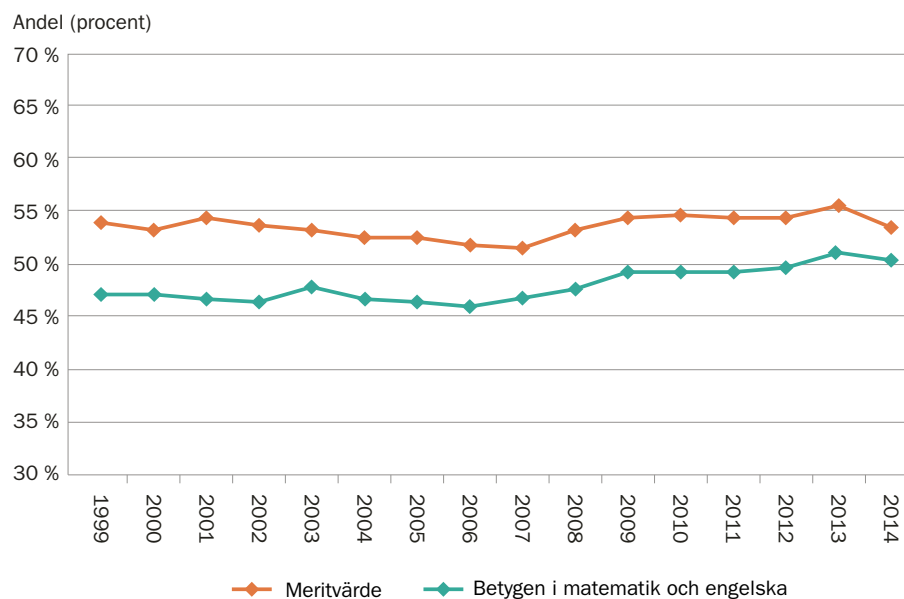
129 En del av skillnaderna i resultat mellan familjer beror på att det finns skillnader i socioekonomiska förutsättningar mellan familjer. Om vi ser en ökning i syskonkorrelationen även efter kontroll för det socioekonomiska indexet, innebär det att det skett en ökad betydelse av aspekter på den socioekonomiska bakgrunden som inte fångas upp av utbildningsnivå, inkomst eller grad av bidragstagande. Genom att kontrollera för socioekonomisk bakgrund rensar vi bort den del av aspekter på den eventuella förändringar i syskonkorrelationen över tid som beror på ökade eller minskade skillnader i socioekonomiska bakgrunden som inte fångas upp av utbildningsnivå, inkomst eller grad av bidragstagandeförutsättningar mellan syskonpar.

Sammantaget visar syskonkorrelationerna för samtliga elever ett liknade mönster som för det socioekonomiska indexet i huvudtexten (figur 2.1).

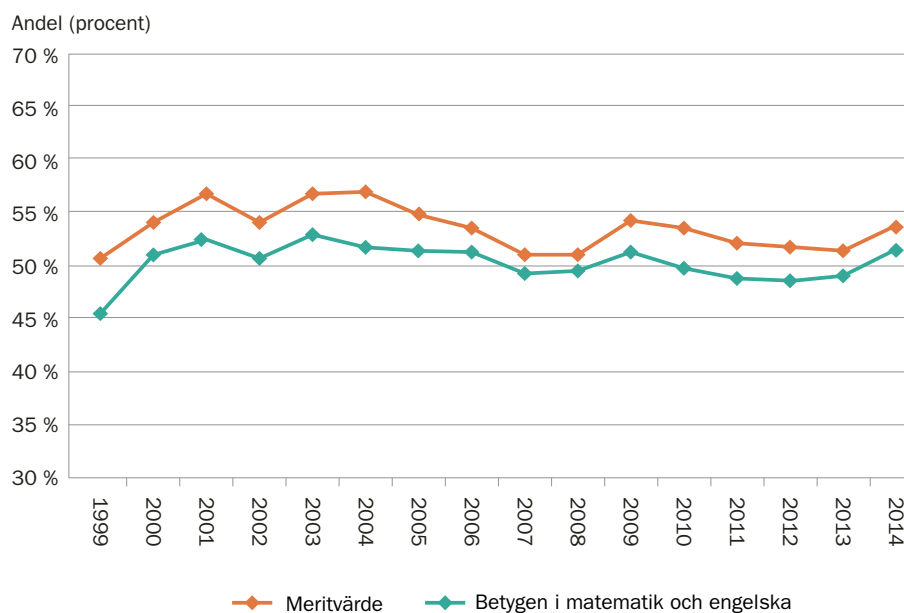
Figur B2.5a-c visar syskonkorrelationerna uppdelat på migrationsbakgrund för meritvärdet respektive betygen i matematik och engelska.

Figur B2.5a-c Syskonkorrelationer med avseende på meritvärdet och betygen i matematik och engelska, uppdelat på migrationsbakgrund.*

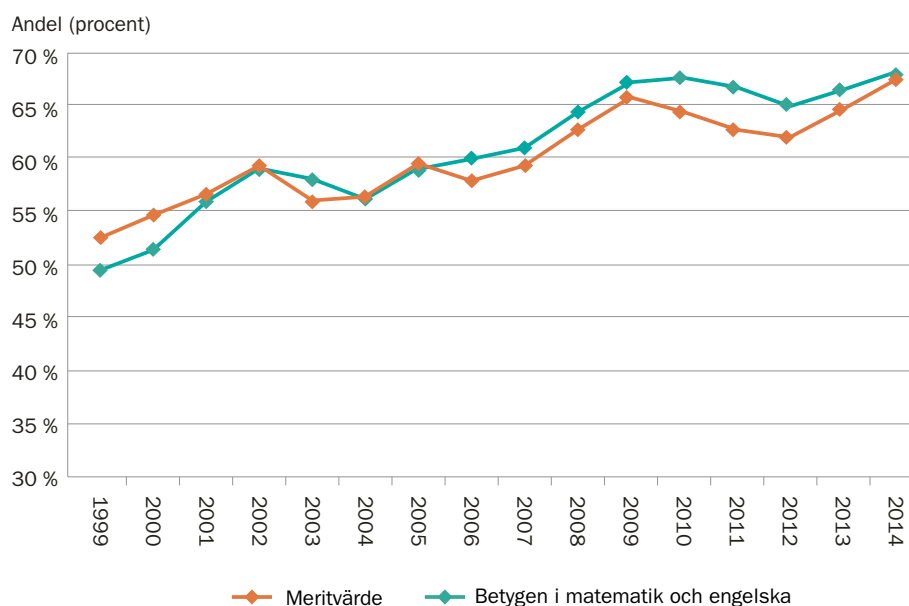
a. Elever med svensk bakgrund.



b. Elever födda i Sverige med utländsk bakgrund.



c. Elever födda utomlands.



* Syskonkorrelationen mäts med variansdekomposition. Andelen av den totala variationen i det årsvis standardiserade meritvärdet respektive betygen i matematik och engelska, som kan förklaras av variation mellan syskonpar/syskongrupper. Observera att tidsserierna inte är justerade för bytet av betygsskala 2013. I analyserna ingår bara elever med personnummer och där det finns uppgift om föräldrars utbildningsnivå, inkomst och bidragstagande (de variabler som ingår i det socioekonomiska indexet). Dessutom ingår bara elever som har minst ett syskon inom 2 kalenderår. Observera att skalan för den vertikala axeln inte börjar på noll.

Uppdelningen av syskonkorrelationerna utifrån på migrationsbakgrund baserat på elevernas meritvärden visar kraftigt ökande korrelationer för elever födda utomlands. För elever födda i Sverige med utländsk bakgrund är det en tvärtom en sjunkande tendens – åtminstone från början till mitten av 00-talet. För elever med svensk bakgrund syns en ökning av korrelationerna från ungefär 2007–2008 som avtar de senaste åren. Korrelationerna för meritvärdet sjunker något från början till mitten av 00-talet, medan det för betygen i matematik och engelska är stabilt 1999–2006.

Den samlade bilden är att det är de ökade korrelationerna för utlandsfödda elever med utländsk bakgrund och att denna grupp ökar som är den huvudsakliga förklaringen till att korrelationerna ökar för samtliga elever i figur B2.4.

I Excelfilen redovisas även figurer där syskonkorrelationerna kontrolleras för det socioekonomiska indexet. Tolkningen av tidserierna efter kontroll för socioekonomisk bakgrund är inte självklar.¹³⁰ En tolkning är att differensen i syskonkorrelation mellan den okontrollerade och den kontrollerade syskonkorrelationen utgör ett mått på hur mycket de mätbara socioekonomiska variablerna utgör av den totala syskonkorrelationen. Den generella bilden är att kontrollen för socioekonomisk bakgrund som förväntat sänker nivån för korrelationerna. Kurvorna efter kontroll följer respektive okontrollerad kurva väl.

¹³⁰ Se Björklund m.fl. (2010b) och Böhlmark & Holmlund (2012) s.22 för en mer detaljerad diskussion om vad det innebär att kontrollera syskonkorrelationer för socioekonomisk bakgrund.

B2.3 Betydelsen av socioekonomisk bakgrund enligt PISA

När det gäller frågan om hur socioekonomisk bakgrund påverkar skolresultaten har rapporten lagt större vikt vid registeruppgifter än de internationella kunskapsmätningarna. Det finns två huvudsakliga orsaker till detta. Den ena orsaken är att uppgifterna om socioekonomisk bakgrund i de internationella kunskapsmätningarna bygger på enkätuppgifter från eleverna.¹³¹ Uppgifter som samlas in via enkätfrågor till elever om till exempel föräldrarnas utbildningsnivå är behäftade med vissa mätfel, vilket inte alls i lika hög grad är fallet med registeruppgifter. Den andra orsaken är att PISA liksom andra internationella kunskapsundersökningar bygger på urval (i storleksordningen 5000 elever) och som samlas in vart tredje år (PISA). Betygs- och registeruppgifter byggs däremot på samtliga elever i en årskurs, dvs totaldata, och finns dessutom tillgängliga för varje år. Det finns således en viss urvalsosäkerhet när det gäller urvalsundersökningar och som leder till att eventuella förändringar över tid blir svårare att urskilja samt att tidserierna kan framstå som skakigare. Med registerdata finns inte detta problem.

Men det finns även fördelar med att mäta betydelsen av socioekonomisk bakgrund utifrån de internationella kunskapsmätningarna. Den mest uppenbara fördelen är att vi kan jämföra betydelsen av socioekonomisk bakgrund i Sverige med hur det ser ut i andra länder. En annan möjlig fördel med de internationella kunskapsmätningarna är att de index som används för att mäta socioekonomisk bakgrund delvis bygger på uppgifter som inte finns i register, till exempel frågor om hur många böcker som finns i hemmet eller om eleven har tillgång till en egen studieplats. För PISA kallas indexet *The PISA index of Economic, Social and Cultural Status*, ESCS.¹³² Detta skulle potentiellt kunna betyda att man fångar in andra aspekter av en elevs socioekonomiska bakgrund än med tillgängliga registeruppgifter. En ytterligare fördel med de internationella kunskapsmätningarna är att dessa kunskapstest är ett mer trovärdigt mått på elevernas kunskaper jämfört med betygen, särskilt vad gäller kunskapsutvecklingen över tid.

Men sammantaget gör vi ändå bedömningen att analysen av registerdata ger oss det mest pålitliga måttet på hur betydelsen av socioekonomisk bakgrund har utvecklats över tid. En indikation på detta är att de socioekonomiska registervariablerna förklarar betydligt mer av variationen i skolresultat jämfört med det socioekonomiska indexet i PISA. Nedan redovisas i lite mer detalj några problem som är behäftade med de internationella studierna när det gäller att mäta socioekonomisk bakgrund.¹³³

131 Brese & Mirazchiyski (2013).

132 OECD (2016) s.205.

133 Leslie och David Rutkowski har skrivit flera artiklar om osäkerheten gällande de socioekonomiska bakgrundsmåtten i PISA och TIMSS. Rutkowski & Rutkowski (2010), Rutkowski (2011), Rutkowski & Rutkowski (2013).

Mätfel i enkätsvar om föräldrars utbildningsnivå

Det finns en osäkerhet om uppgifternas tillförlitlighet eftersom det är eleverna själva som lämnat uppgifterna. Vi har i den internationella kunskapsmätningen TIMSS från och med år 2015 tillgång till kompletterande registeruppgifter om föräldrarnas utbildningsnivå för de elever som deltagit i undersökningen.

När vi jämför enkätsvaren från eleverna som går i årskurs 8 med registeruppgifterna visar det sig att en stor andel av eleverna svarar fel om sina föräldrars utbildning och över en tredjedel av eleverna svarar att de inte vet. Enbart runt hälften av eleverna har korrekt pekat ut utbildningsnivån för den förälder som har högst utbildning. En konsekvens av detta mätfel blir att utbildningsnivåns påverkan på resultaten underskattas med enkätuppgifterna. När utbildningsnivåns betydelse för TIMSS-resultaten istället mäts med registervariabeln förklaras ungefär dubbelt så mycket av resultatvariationen jämfört med om man använder uppgifterna från elevenkäten.¹³⁴ Då vi inte har tillgång till registervariabler för tidigare år för TIMSS och inte har det alls i PISA så kan vi inte avgöra hur detta mätfel påverkar analyserna över tid.¹³⁵

Vissa av enkätfrågorna i PISA har ändrats mellan olika undersökningsomgångar på ett sätt som gör det svårt att veta om detta påverkat svaren.¹³⁶

Förändringar i variablerna över tid när det gäller internationella kunskapsmätningar

Precis som med registervariabler kan det vara problematiskt om variablerna förändras allt för mycket över tid. En variabel som förändrats över tid är *antal böcker i hemmet* som brukar framhållas som en som de variabler som har starkast samband med testresultaten och ingår i index i både PISA och TIMSS.¹³⁷ Andelen elever i TIMSS som anger att de har fler än 200 böcker i hemmet har halverats mellan 1995–2015, från cirka 40 till cirka 20 procent av eleverna. Det är svårt att veta hur detta påverkar analyser över tid. Elever med få böcker i hemmet och låga testresultat tenderar enligt en studie också att underrapportera antal böcker i hemmet.¹³⁸ Det gör att sambandet mellan antal böcker i hemmet och testresultaten ser starkare ut än vad det är i verkligheten. Det är alltså tvärtom mot vad som tidigare beskrevs när det gäller föräldrars utbildning – där felaktig rapportering från eleverna bidrar till att minska sambandets styrka med testresultaten. Inget av detta behöver dock nödvändigtvis påverka familjebakgrundens betydelse för resultaten över tid. För det krävs att under- respektive överskattningen ökar eller minskar över tid och vi vet inte om så är fallet. När det gäller PISA:s ESCS-index så sker, på samma vis som för våra registervariabler, vissa förändringar över tid.¹³⁹ Vi hanterar därför detta på samma vis som registervariablerna genom att göra en årsvis standardisering av ESCS-indexet.

134 Korrelationen mellan utbildningsnivån för föräldrarna som elevernas uppgett i enkäten och elevernas matematikresultat i årskurs 8 i TIMSS 2015 är 0,26. Korrelationen ökar till 0,40 om man istället använder registeruppgifterna om föräldrarnas utbildningsnivå. Andelen förklarad variation (R²) ökar på samma vis från 0,07 till 0,16. Se även Engzell & Jonsson (2015) för en analys av bortfall i elevenkäter om föräldrarnas utbildningsnivå.

135 Registeruppgifter kommer först i samband med PISA 2018.

136 En genomgång av förändringarna av frågorna i PISA görs i Heller Sahlgren (2017) s.47.

137 Brese & Mirazchiyski (2013) s.98. För en studie som utgår ifrån antal böcker i hemmet se Gustafsson m.fl. (2013).

138 Engzell (2016b).

139 Se Excel-filen.

Men standardiseringen av ESCS-indexet har visat sig innebära ett oväntat problem.

Det man framförallt vill uppnå med en årsvis standardisering är att kontrollera för om spridningen förändras över tid – det vill säga om spridningen i indexet ökar eller minskar. Det vi ser när vi tittar på spridningen av ESCS-index är dock ingen sådan jämn utveckling av spridningen i index. Det vi ser är istället att spridningen i index ligger på en relativt konstant nivå över tid, med undantag för ett enda år. År 2003 sker en plötslig ökning av spridningen i index och denna ökning sker inte bara för Sverige utan även för OECD som helhet. Vi har inte underlag för att avgöra om det skett någon förändring i verkligheten som gör att spridningen i index är extra stor detta år. Men det finns skäl att misstänka att det finns någon annan förklaring än att den verkliga spridningen mellan elevers familjebakgrund skulle vara särskilt stor just år 2003.

Problemet är att en årsvis standardisering av indexet förstärker effekten av att spridningen i indexet är särskilt stor 2003. Detta får konsekvenser för trenden när man använder 2003 som startår.¹⁴⁰

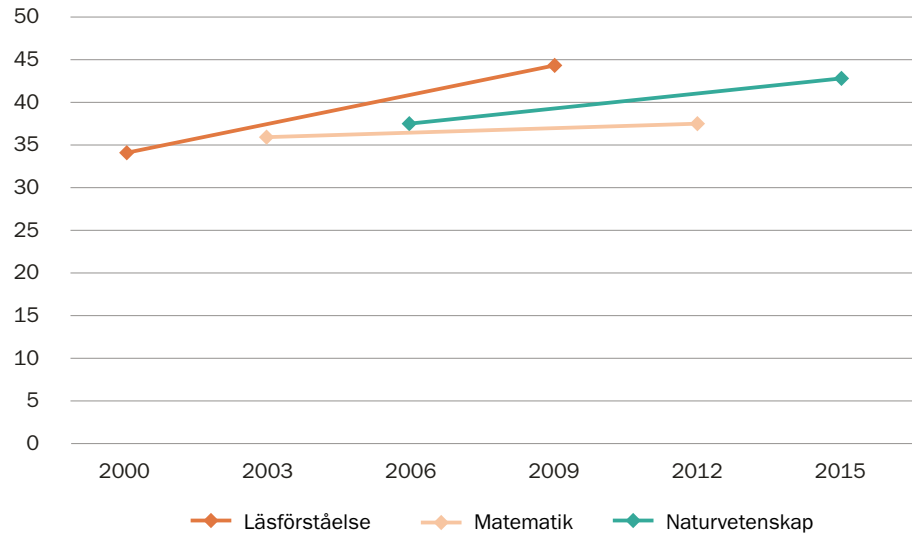
När vi redovisar betydelsen av socioekonomisk bakgrund enligt PISA nedan har vi därför valt att redovisa resultaten för såväl det icke standardiserade som det standardiserade ESCS-index (B2.6a-b).

I analysen av PISA redovisar vi enbart resultaten för de år där respektive ämne varit huvudområde – detta eftersom vi då har säkrare skattningar av elevernas resultat i respektive ämne.

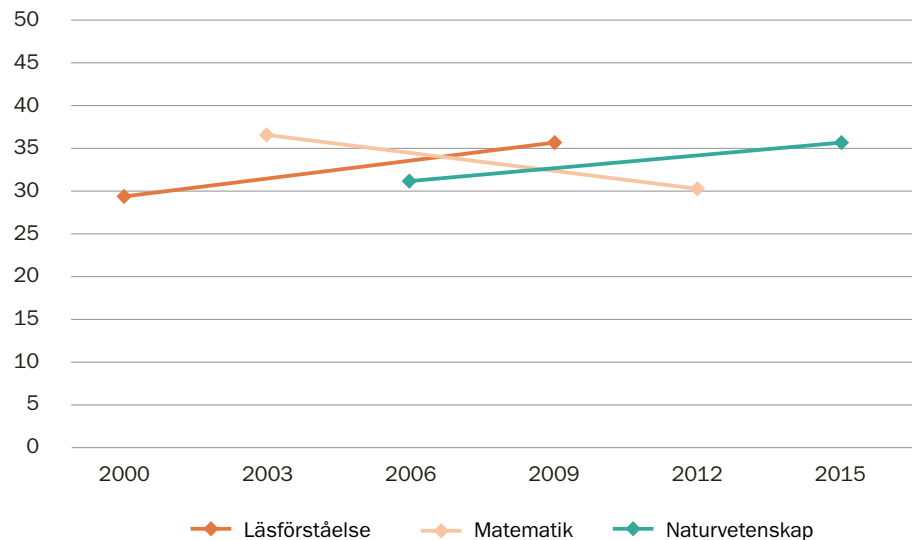
140 Allt annat lika kommer en större spridning i ESCS-indexet att leda till att den standardiserade effektstorleken (vad gäller betydelsen av ESCS för resultat) för 2003 blir större än vad den annars skulle vara.

Figur B2.6a-b Betydelsen av socioekonomisk bakgrund för resultaten i läsförståelse, matematik och naturvetenskap, utifrån standardiserat respektive icke standardiserat socioekonomiskt index (ESCS). Samtliga elevgrupper.*

a. Ej standardiserat socioekonomiskt index



b. Standardiserat socioekonomiskt index

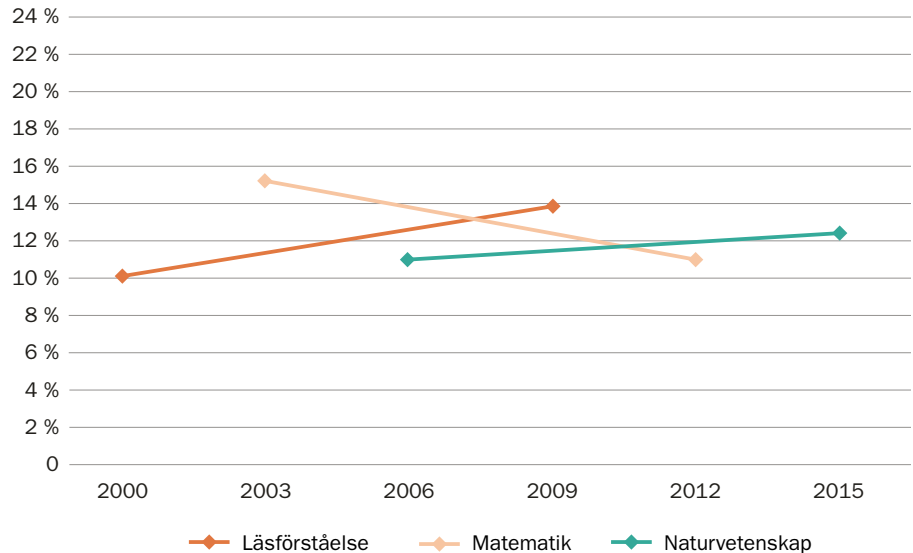


* Betydelsen av socioekonomisk bakgrund mäts som effektstorleken i enkel regression med PISA:s socioekonomiska index, ESCS, som förklarande variabel.

Figur B2.6 a och b visar att effektstorleken har ökat i läsförståelse och naturvetenskap på ett likartat vis både med ett standardiserat och ostandardiserat index. För matematik påverkas analysen däremot av om indexet standardiseras eller inte. Den ostandardiserade analysen visar ingen förändring 2003–2012 medan analysen med ett standardiserat index visar på en sjunkande betydelse. Som påpekats ovan sker en ökad spridning i indexet år 2003 vilket kan ha betydelse för utvecklingen i matematik som har startåret just 2003.

När det gäller *andel förklarad variation* (R2) av ESCS-indexet för PISA-resultaten så är inte analysen beroende av förändringar i indexet över tid på samma vis. Därför är inte index standardiserat i figur B2.6c.¹⁴¹

Figur B2.6c Betydelsen av socioekonomisk bakgrund för resultaten i läsförståelse, matematik och naturvetenskap, uttryckts som andel förklarad variation (R2), utifrån socioekonomiskt index (ESCS). Samtliga elevgrupper. *



* Betydelsen av socioekonomisk bakgrund uttryckt som andel förklarad variation i enkel regression med PISA:s socioekonomiska index, ESCS, som förklarande variabel. Observera att när betydelsen mäts med andel förklarad variation spelar det ingen roll om man använder standardiserat eller icke standardiserat ESCS.

Andelen förklarad variation (R2) utifrån ECSC-index visar en ökande förklaring i läsförståelse, ingen förändring i naturvetenskap och en sjunkande förklaring i matematik.

Sammantaget ger analyserna av PISA-resultaten inte någon samstämmig bild av utvecklingen. I *läsförståelse* och *naturvetenskap* har betydelsen av familjebakgrund ökat över tid, medan den minskat i *matematik*. Osäkerheterna är flera och det är svårt att dra några tydliga slutsatser om utvecklingen. Förutom osäkerheterna kan resultaten vara beroende på ämne och på startår och slutår.

141 I enkel regression finns ett direkt samband mellan andel förklarad variation och den standardiserade effektstorleken, därav det snarlika mönstret mellan figur B2.6b och B2.6c.

Vi har även gjort PISA-analyserna i figur B2.6a-c enbart för elever med svensk bakgrund och vi ser då att trenderna inte förändras på något avgörande vis.¹⁴² En ökad invandring har alltså inte haft någon avgörande betydelse för resultaten – men då har inte heller PISA 2015 fångat in det stora antal invandrare som kom hösten 2015.¹⁴³

B2.4 Förändrade demografiska faktorer för utlandsfödda elever

Hur en invandrad elev klarar sig i skolan hänger också samman med vilken skolgång eleven har med sig från hemlandet. En del barn kan ha bakom sig en lång period av bristande utbildning, till exempel för att de har varit på flykt eller att ursprungslandets skolsystem är relativt outvecklat. Ett sätt att beskriva detta är att dela upp invandringsländerna utifrån klassificeringen *Human Development Index* (HDI). HDI delas in i sju grupper där till exempel HDI-grupp 1 motsvarar Sverige och HDI-grupp 7 utgörs av länder med den lägsta utvecklingsnivån. Figur B2.7 visar andelen utlandsfödda årskurs 9-elever som invandrat från olika utifrån deras HDI. Ökningen de senaste åren är framför allt från länder med allra lägst HDI. I grupp 7 ingår länder som Irak, Afghanistan och Somalia, medan Syrien ingår i grupp 6.¹⁴⁴

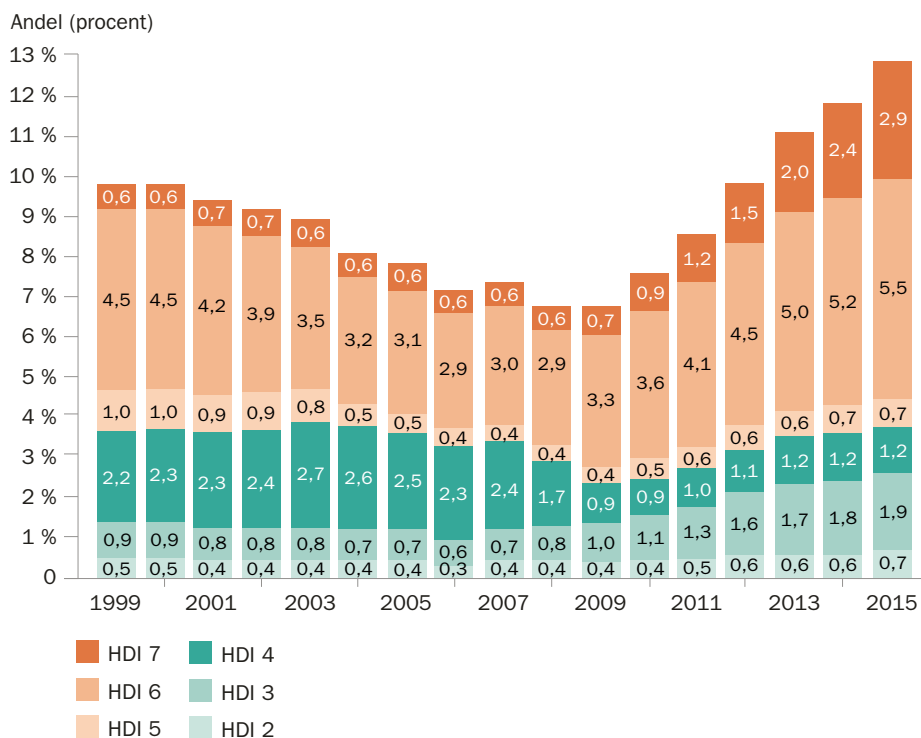
Samtidigt har den genomsnittliga utbildningsnivån hos den här elevgruppens föräldrar (HDI=7) minskat avsevärt (figur B2.8). Detta visar sig också i betygresultaten för samma grupp (figur B2.9), då det standardiserade meritvärdet för elever som tillhör HDI-grupp 7 har gått ned avsevärt under i stort sett hela 00-talet. Även elever från HDI-grupp 6 och 4 har tappat jämfört med övriga elevgrupper. Dessutom har den genomsnittliga åldern när utlandsfödda elever invandrat till Sverige ökat med mer än 2 år sedan mitten av 00-talet (figur B2.10).

142 Se Excel-filen.

143 Jmf Skolverket (2016c) s.32.

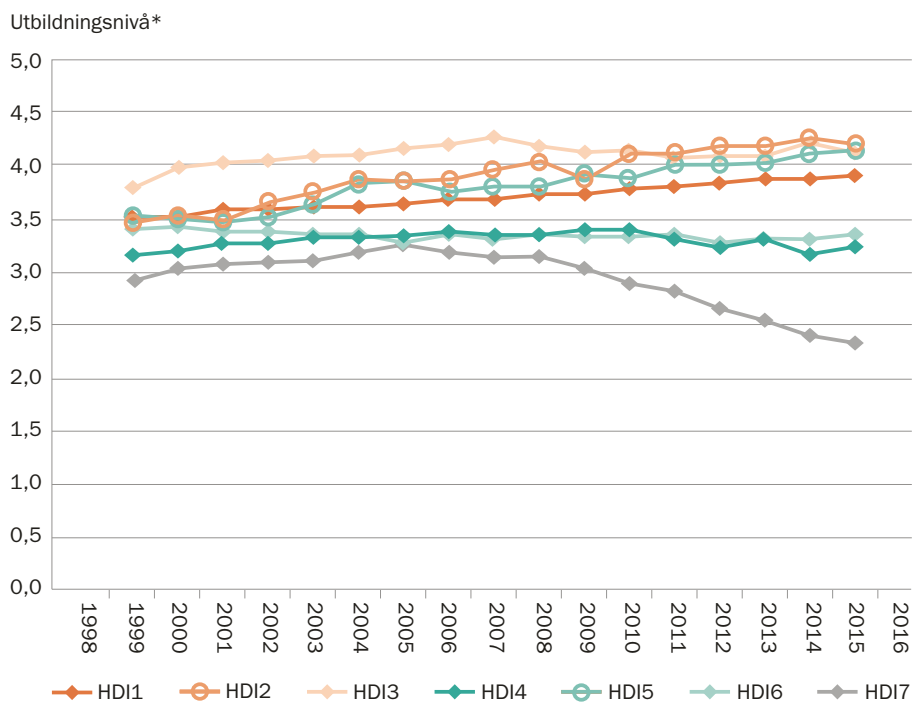
144 En beskrivning av vilka länder som ingår i olika HDI-grupper görs i Flood & Ruist (2015) appendix III.

Figur B2.7 Utlandsfödda elevers ursprungsland, grupperat efter landets HDI (Human development index).*



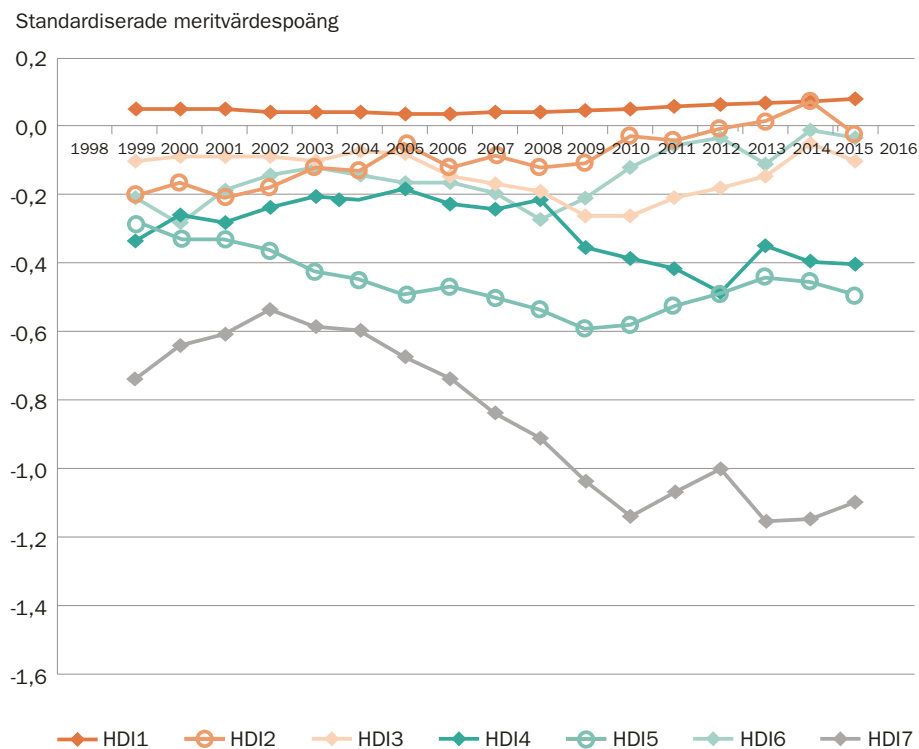
* HDI-grupp 1 motsvarar Sverige och HDI-grupp 7 utgörs av länder med den lägsta utvecklingsnivån. För en mer detaljerad beskrivning av vilka länder som ingår i olika HDI-grupper, se Flood & Ruist (2015) appendix III.

Figur B2.8 Genomsnittlig utbildningsnivå hos elevernas föräldrar, uppdelat efter HDI-grupp.

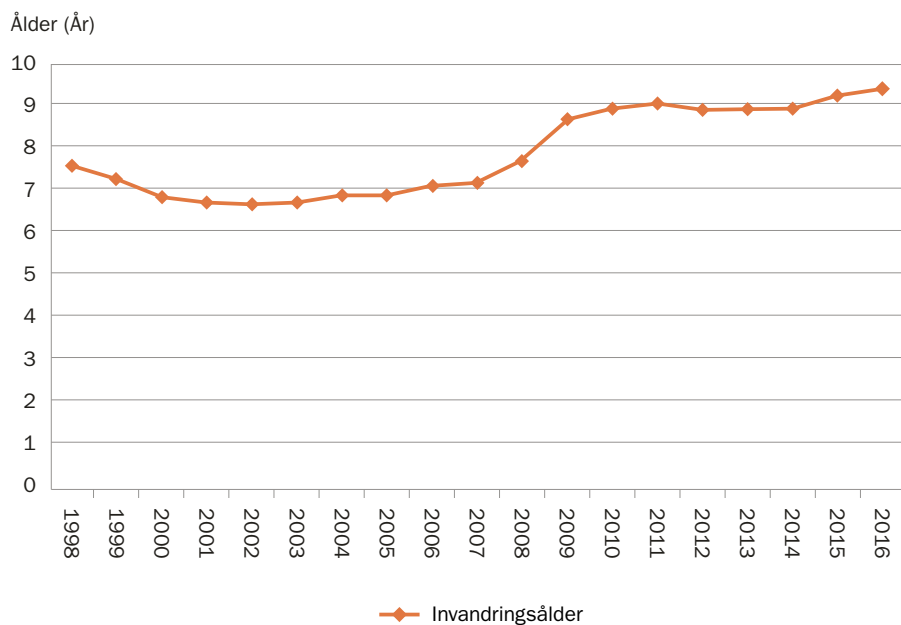


* Utbildningsnivå efter [edu13] som bygger på föräldrarnas genomsnittliga utbildningsnivå. Se bilaga B1 för mer information.

Figur B2.9 Genomsnittligt meritvärde (standardiserat), uppdelat efter HDI-grupp.



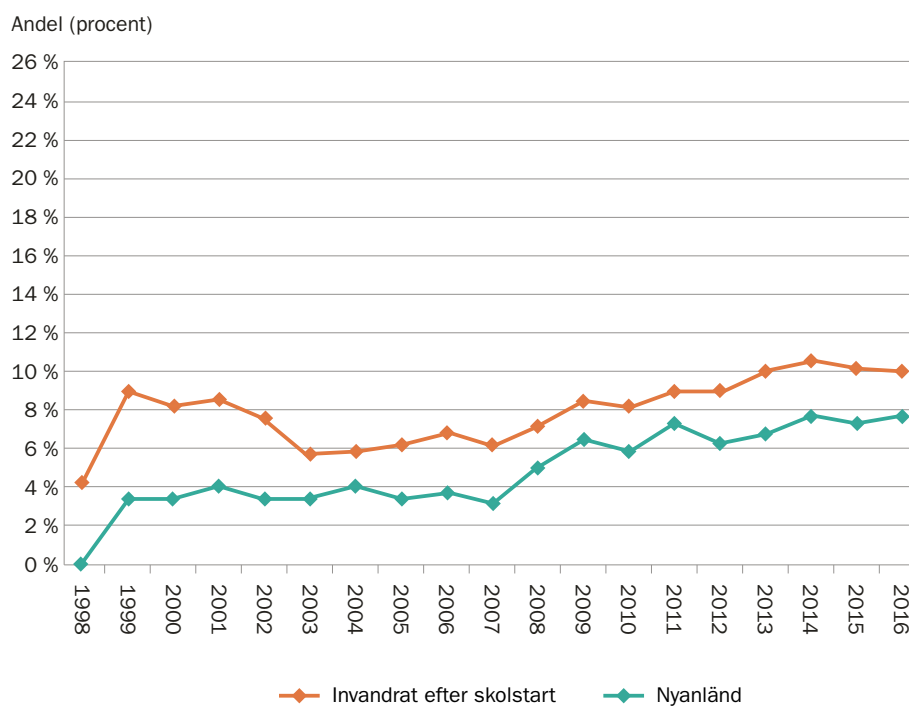
Figur B2.10 Genomsnittlig invandringsålder för elever födda utomlands.



Bilaga B3

B3.1 Skolsegregation utifrån nyanlända respektive elever som invandrat efter skolstart

Figur B3.1 Skolsegregation utifrån nyanlända elever respektive elever som invandrat efter skolstart.*



* Data är viktat efter skolstorlek. Med nyanländ menas att elever invandrat under de senaste 4 åren i förhållande till avgångsåret. Med elever som invandrat efter skolstart menas elever som invandrat efter 7 års ålder.

Bilaga B4

B4.1 Variation i resultat mellan klasser och mellan kommuner

Variation mellan klasser

På samma sätt som att det finns skillnader i skolresultat mellan skolor kan det finnas skillnader i betygsresultat inom skolor mellan klasser. Mellanklassvariationens utveckling över tid kan ge en indikation på i vilken grad elever nivågrupperas på skolorna, eller i vilken utsträckning eleverna fördelas på klasser efter hur de bor. Det finns inga statliga krav på indelningen i klasser.¹⁴⁵ Rektor avgör själv vilka principer som ska gälla.¹⁴⁶ Det finns inte heller någon garanti för att det existerar permanenta klasser/undervisningsgrupper, där eleverna går tillsammans i ett ämne eller i de flesta ämnen under flera år innan de går ut årskurs 9.

Uppgifterna om klasstillhörighet måste tolkas varsamt, eftersom det är svårt att veta i vilken utsträckning eleverna undervisas i de klasser de formellt tillhört enligt klassbeteckningen.¹⁴⁷ Uppgifter om klass finns också bara tillgängliga sedan 2004.¹⁴⁸

Figur B4.1 visar utvecklingen av mellanklassvariationen i det standardiserade meritvärdet.

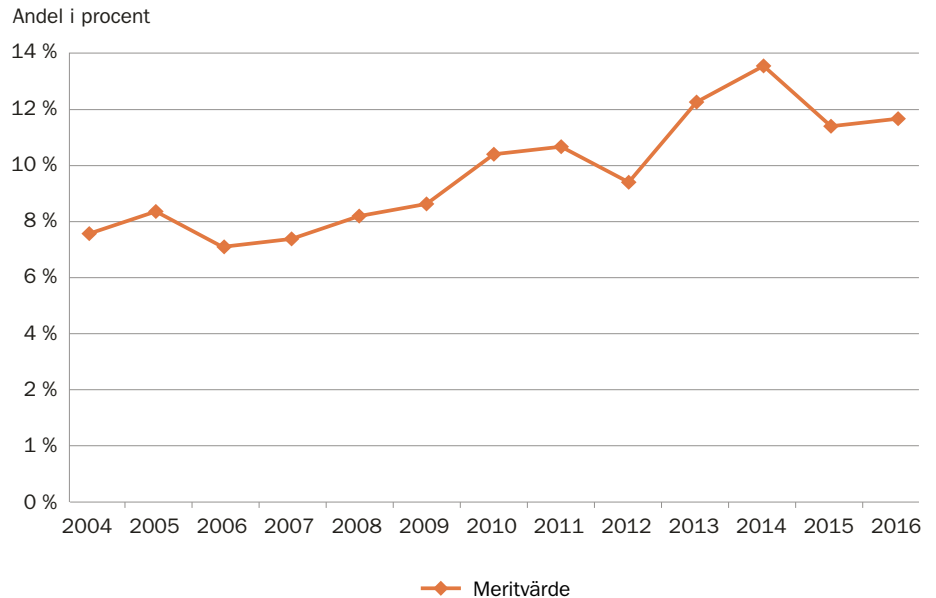
145 Skolverket (2012) kap. 3.2.

146 Det finns inte någon explicit bestämmelse om att rektor beslutar om indelningen i klasser/undervisningsgrupper. Det följer indirekt av bestämmelsen i skollagen 2 kap. 10 § om att rektor "beslutar om sin enhets inre organisation". Se författningskommentaren SOU 2009/10:165 s. 648.

147 Uppgiften om "klass" kommer från insamlingen av årskurs 9-betygen. Vi har även tillgång till uppgifter om "undervisningsgrupp" genom insamlingen av de nationella proven. Det är en hög överensstämmelse mellan elevernas "klass" och "undervisningsgrupp", vilket är en indikation på att uppgifterna är relevanta att analysera. Uppgifterna om klass/undervisningsgrupp är av olika skäl ganska osäkra men bortfallet för klasstillhörighet är betydligt mindre och stabilare jämfört med uppgift om undervisningsgrupp i matematik och engelska sett över hela tidsperioden. Se Excelfilen. Vi har exkluderat alla elever/skolor där klasstorleken för någon klass överskrider 40 elever (orealistisk klasstorlek), vilket innebär att cirka 9 procent av alla elever exkluderats på grund av denna restriktion. Ytterligare 0,1 procent av eleverna har exkluderats utifrån restriktionen att antalet klasser per skola är 10 eller fler, samtidigt som antalet elever per klass understiger 5 elever.

148 Holmlund m.fl. (2014) s. 134 har tillgång till klassuppgifter längre tillbaka i tiden och kan visa på en ökning av mellanklassvariationen som startar långt tidigare.

Figur B4.1 Mellanklassvariation i meritvärde. Uttryckt som andel av variationen i meritvärde inom skolor som kan förklaras av variation i meritvärde mellan klasser. Viktat på skolnivå.*



* Mellanklassvariationen är skattad med variansdekomposition i en hierarkisk 3-nivåmodell (programvara: hlm). Data är viktat på skolnivå, men inte på klassnivå.

Figur B4.1 visar en tilltagande mellanklassvariation i meritvärdet över tidsperioden 2004 till 2016. Ökningen ligger på ungefär 4 procentenheter, från cirka 8 procentenheter 2004 till 12 procentenheter 2016. Tolkningen är att eleverna tenderar att ha blivit mer resultatmässigt segregerade över tidsperioden.

Det kan alltså ha skett en ökad indelning av klasser efter elevernas förutsättningar. En alternativ förklaring är att betygssättningen blivit mindre likvärdig mellan olika lärare. Men vi kan inte se att lärares betygssättning i förhållande till de nationella proven blivit vare sig mer eller mindre likvärdig över tid – möjligen då att likvärdigheten i betygssättning förämrats något efter införandet av den nya betygsskalan 2013.¹⁴⁹ En ytterligare förklaring är att variationen i undervisningskvaliteten har ökat mellan olika lärare på samma skola, men det finns inga empiriska belägg för att så skulle ha skett.

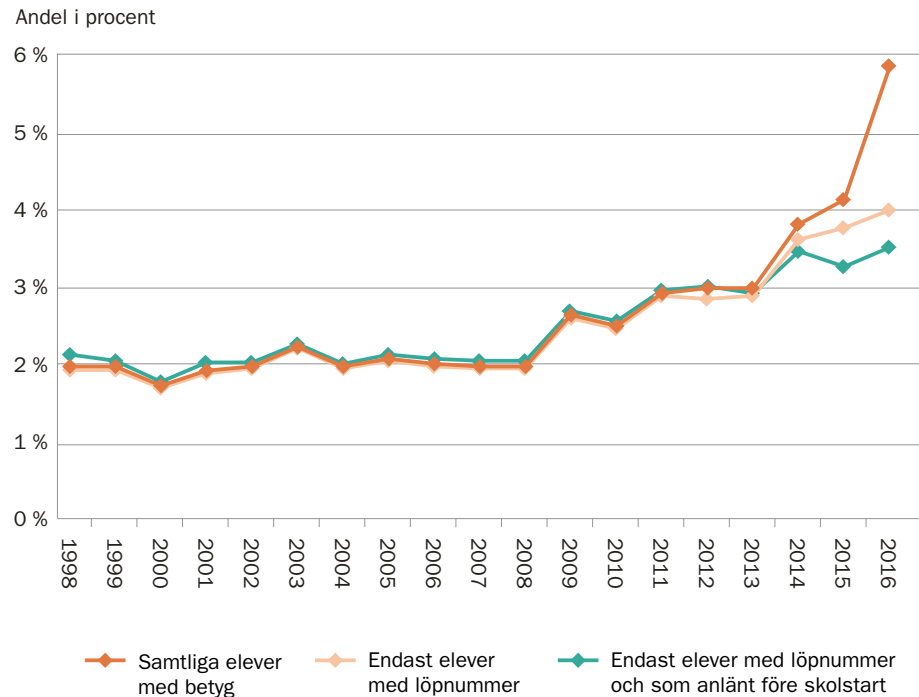
Mellankommunvariation

Figur B4.2 visar hur stor andel av den totala variationen i meritvärde som kan förklaras av skillnader i kommuners genomsnittliga meritvärde, enligt samma princip som när mellanskolvariationen och mellanklassvariationen analyserades.

Eftersom det skett en kraftig ökning av nyinvandrade elever under senare år och i synnerhet hösten 2015, har vi delat upp analysen i tre grupper: samtliga elever, endast elever med ett registrerat löpnummer samt en grupp med enbart elever med löpnummer och som dessutom invandrat till Sverige före skolstart.

149 Skolverket (2016b).

Figur B4.2 Mellankommunvariation i betygsresultat. Uttryckt som andel av den totala variationen i meritvärde som kan förklaras av variation i kommuners genomsnittliga meritvärde.*



* Mellankommunvariationen är skattad med variansdekomposition i en hierarkisk 2-nivåmodell (programvara: spss). Data är inte viktad.

Figur B4.2 visar att variationen i betygsresultat mellan kommuner är relativt låg. Under större delen av 00-talet låg den stabilt på cirka 2 procent, dvs. bara 2 procent av den totala variationen i elevers meritvärde kunde förklaras av skillnader i kommunernas genomsnittliga meritvärde. Efter 2008 bryts detta mönster och resultatvariationen mellan kommuner börjar att öka, och ligger 2013 på cirka 3 procent. Visserligen är ökningen bara en procentenhet, men med tanke på att ökningen skett från en mycket låg nivå är den ändå anmärkningsvärd.

Fram till och med 2013 ser både nivån och utvecklingen identisk ut, oavsett vilken delmängd av elever som analyseras. Därefter visar mellankommunvariation för alla elever med betyg en kraftig uppgång till nästan 6 procent 2016. Även när endast elever med betyg och löpnummer analyseras ökar mellankommunvariationen, men ökningen är betydligt svagare. Förklaringen till denna skillnad är att de elever som saknar löpnummer i registren till övervägande del nyligen har anlänt till Sverige och ännu inte fått uppehållstillstånd, och därmed ännu inte fått personnummer. När vi exkluderar dessa elever blir ökningen inte lika extrem. När vi dessutom exkluderar alla elever som anlänt till Sverige efter skolstart blir ökningen ännu svagare.

Vad figur B4.2 visar är alltså att många av de elever som invandrat till Sverige under senare år och speciellt de allra senaste åren inte fördelas jämnt på olika kommuner, utan koncentreras till vissa kommuner. Eftersom nyanlända elever tenderar att prestera betydligt lägre skolresultat, leder det till en större varia-

tion i mellankommunvariationen. En rimlig tolkning är att detta synliggör hur ojämnt nyanlända fördelas till olika kommuner.¹⁵⁰

B4.2 Skolnivåeffekter

I avsnitt 4.3 redovisades resultatskillnader mellan elever med samma socioekonomiska bakgrund men som går på skolor med olika socioekonomisk sammansättning. Analysen var förenklad i syfte att på ett mer pedagogiskt sätt illustrera dessa så kallade *skolnivåeffekter*. I detta avsnitt görs en mer formell redovisning av skolnivåeffekter och enligt samma koncept som i Skolverkets likvärdighetsrapport från 2012.¹⁵¹

I avsnitt 4.3 kategoriserades elever och skolor efter sin socioekonomiska bakgrund respektive socioekonomiska sammansättning. Här görs analysen utifrån de underliggande kontinuerliga variablerna, dvs. elevens egen individuella socioekonomiska bakgrund enligt det socioekonomiska indexet samt skolornas genomsnittliga socioekonomiska bakgrund baserat på samma index.

Skolnivåeffekten skattas formellt i en 2-nivåmodell:

Elevnivå:

$$[\text{meritv}] = b_0 + b_1[\text{sex}] + b_2[\text{migswe}] + b_3[\text{migfor}] + b_4[\text{elev_tidig}] + b_5[\text{elev_sen}] + b_6[\text{socioindex}] + r$$

Skolnivå:

$$b_0 = g_0 + g_1[\text{socioindex_skola}] + u$$

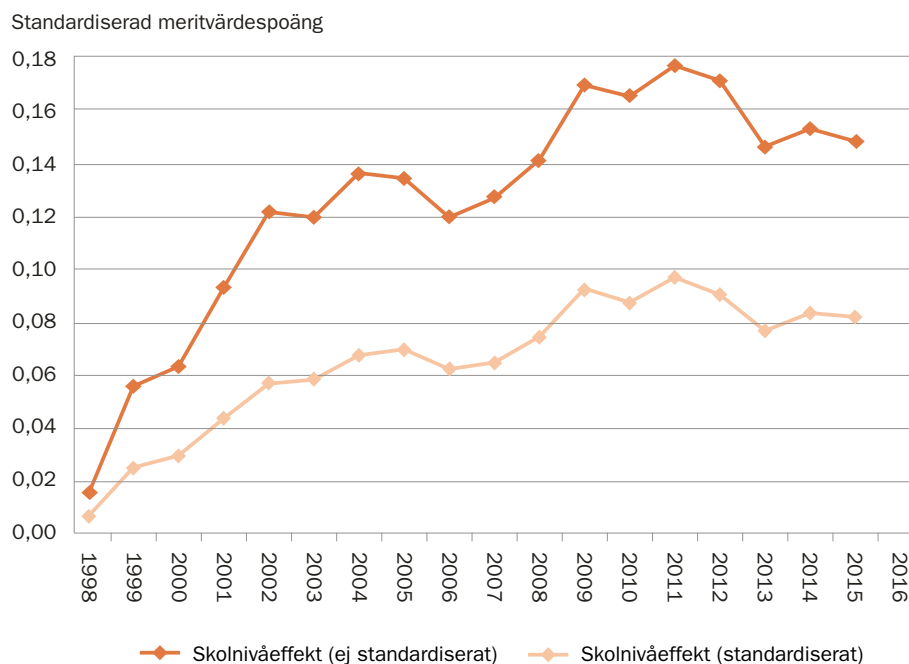
Där skolnivåeffekten utgörs av den skattade regressionskoefficienten g_1 .

Figur B4.3 visar både den standardiserade och den icke standardiserade skolnivåeffekten.

150 SOU 2016:9. Skolverket (2017c).

151 Skolverket (2012) kap 6.2.

Figur B4.3 Skolnivåeffekt på det standardiserade meritvärdet utifrån skolans socioekonomiska sammansättning.*



* Utfallsvariabel är det standardiserade meritvärdet.

Tolkning av den icke standardiserade skolnivåeffekten i figuren:

När vi jämför två elevers resultat där båda har identisk bakgrund utifrån samtliga variabler som ingår på elevnivå i modellen (se sidan innan), men som går på två olika skolor och där skolornas socioekonomiska sammansättning skiljer sig med en enhet i det socioekonomiska indexet, är den förväntade resultatskillnaden knappt 2 procent av en standardavvikelse i meritvärdet år 1998. År 2015 är samma förväntade skillnad drygt 14 procent av en standardavvikelse i meritvärdet.

Tolkning av den standardiserade skolnivåeffekten i figuren:

När vi jämför två elevers resultat där båda har identisk bakgrund utifrån samtliga variabler som ingår på elevnivå i modellen (se sidan innan), men som går på två olika skolor och där skolornas socioekonomiska sammansättning skiljer sig med en standardavvikelse i skolornas genomsnittliga socioekonomiska index, är den förväntade resultatskillnaden ungefär 1 procent av en standardavvikelse i meritvärdet år 1998. År 2015 är samma förväntade skillnad 9 procent av en standardavvikelse i meritvärdet.

Eftersom vi i båda fallen använder det standardiserade meritvärdet som utfallsvariabel, utgörs skillnaden mellan de båda kurvorna i praktiken av vilken förändring i x-led (socio) vi väljer att mäta förändringen i y-led (resultat) över. I den icke standardiserade tidserien är skillnaden mellan skolorna en hel enhet i det socioekonomiska indexet. För den standardiserade tidserien utgörs den i stället av en standardavvikelse i skolornas genomsnittliga socioekonomiska index, vilket är betydligt mindre. Därför blir också skillnaden i resultat mindre. En annan skillnad är att för den icke standardiserade tidserien hålls enheten i x-led konstant, medan för den standardiserade tidserien varierar den med hur spridningen i skolornas socioekonomiska sammansättning varierar över tid.

Oavsett vilken tidserie man tittar på i figur B4.3, framgår att skolnivåeffekten har ökat över tid. Men samtidigt framgår att den har avtagit något sedan 2012. I rapportens huvudtext såg vi i stället att den fortsatte öka för den svagaste gruppen över hela tidsperioden. Det som skiljer är först och främst att det som illustreras i figur 4.6 i huvudtexten utgörs av den samlade betydelsen av skolans socioekonomiska sammansättning (effekt per enhet) \times (avstånd i socioekonomisk sammansättning mellan A- och C-skolor). Den mer formellt korrekta skolnivåeffekten i figur B4.3 visar bara effektstorleken.

Detta innebär att även om skolnivåeffekten avtagit något under de senaste åren, så har skolegregationen fortsatt att öka. Det innebär att avståndet mellan det som vi benämnde A- och C-skolor i huvudtexten har ökat när man utgår från genomsnittlig socioekonomisk sammansättning. Här visas alltså den statistiskt mest korrekt beräknade skolnivåeffekten, medan huvudtexten visar hur denna skolnivåeffekt slår igenom i ”verkligheten” genom att illustrera effekten med meritvärdespoäng för olika socioekonomiska grupperingar av elever och skolor.

Om det inte fanns någon skillnad i socioekonomisk sammansättning mellan skolorna spelar det ingen roll hur stor den riktiga skolnivåeffekten (som framgår av figur B4.3) är, eftersom alla elever går i skolor med identisk sammansättning. Skolnivåeffekten får bara betydelse när det finns skillnader i sammansättning mellan skolorna. Ju större skillnader mellan skolorna är, desto större blir betydelsen av vilken skola en elev går på, utifrån det förväntade meritvärdet.

Vad ligger bakom skolnivåeffekterna?

Det är metodologiskt svårt att närmare precisera vad som exakt ligger bakom skolnivåeffekten. Det krävs bättre data och möjlighet att genomföra kliniska experiment genom att allokera elever slumpmässigt till olika skolor och sedan mäta deras resultat över tid. Vi har inte möjlighet att göra detta i verkligheten. Däremot kan vi precisera vilka orsaker som teoretiskt kan ligga bakom.

Sortering av elever efter bakgrundsvariabler som vi inte kontrollerat för i modellen

Om till exempel mer motiverade elever väljer att gå på skolor där sammansättningen är mer fördelaktig, presterar dessa elever i genomsnitt bättre resultat än vad våra variabler predicerar. Om hela skolnivåeffekten skulle bero på sortering efter sådana bakgrundsfaktorer, skulle det inte göra någon skillnad för en enskild elev om den blev flyttad till en skola med mer gynnsam socioekonomisk sammansättning. Eleven skulle prestera precis samma resultat som i den gamla skolan med sämre sammansättning.

Kamrateffekter

Om en elevs resultat påverkas positivt av hur väl övriga skolkamrater presterar kommer detta att till viss del fångas upp av skolnivåeffekterna så som vi mäter dem. Eftersom elever med en mer gynnsam socioekonomisk bakgrund tenderar att prestera bättre resultat, kommer den genomsnittliga resultatnivån på en skola med en mer gynnsam sammansättning också att vara högre. I så fall kommer en elev att gynnas eller missgynnas beroende på om hen går på en skola med en mer eller mindre gynnsam sammansättning. Ju större skillnaderna är (högre skolegregation), desto större betydelse har det vilken skola man går på.

Utbildningskvalitet

Om skolkvaliteten skiljer sig mellan skolor på ett sätt som samvarierar med socioekonomisk sammansättning, kommer detta att fångas upp av skolnivåeffekten så som den mäts. Det kan till exempel vara så att duktiga och erfarna lärare tenderar att söka sig till skolor med en mer gynnsam socioekonomisk sammansättning.¹⁵² I så fall kommer en elev som byter till en skola med mer gynnsam sammansättning att prestera bättre skolresultat, eftersom eleven får tillgång till bättre lärare.

Läraryväntningar

Det finns en hel del forskning som tyder på att lärares förväntningar på eleverna påverkar elevernas resultat.¹⁵³ Vidare finns forskning som visar på en tendens att lärare sätter en förväntansnivå på alla elever i en klass.¹⁵⁴ Utifrån dessa resultat har forskare diskuterat risken med att ökande prestationsskillnader mellan skolor skulle förstärkas av läraryväntningar – att lärare sänker förväntningar på en klass om den i genomsnitt är mer lågpresterande och att elever påverkas positivt av lärares högre förväntningar.¹⁵⁵ I så fall kommer detta att fångas upp av skolnivåeffekten så som vi mäter den. Denna effekt skulle kunna kategoriseras antingen under kamrateffekter eller som en del av utbildningskvaliteten, men man skulle också kunna se den som en separat effekt. Oavsett så säger denna hypotes att en elev som byter till en skola med en mer gynnsam sammansättning kommer att prestera bättre resultat när läraryväntningarna höjs.

Om skolnivåeffekten helt och hållet utgörs av elevsortering i variabler som vi inte har med i modellen, finns inget likvärdighetproblem med en tilltagande skolnivåeffekt. Om däremot övriga tre effekter ingår till någon del, kommer en ökad skolnivåeffekt att ha negativa effekter på likvärdigheten, allt annat lika. Om dessutom skolsegregationen ökar i ett sådant system, kommer det samlade genomslaget på likvärdigheten att öka ytterligare.

152 Karbownik & Martinsson (2014). Hansson & Gustafsson (2016). Holmlund (2016). Skolverket (2014) s.21 ff. Skolverket (2016c) s.43.

153 Hattie (2009) s. 121 ff.

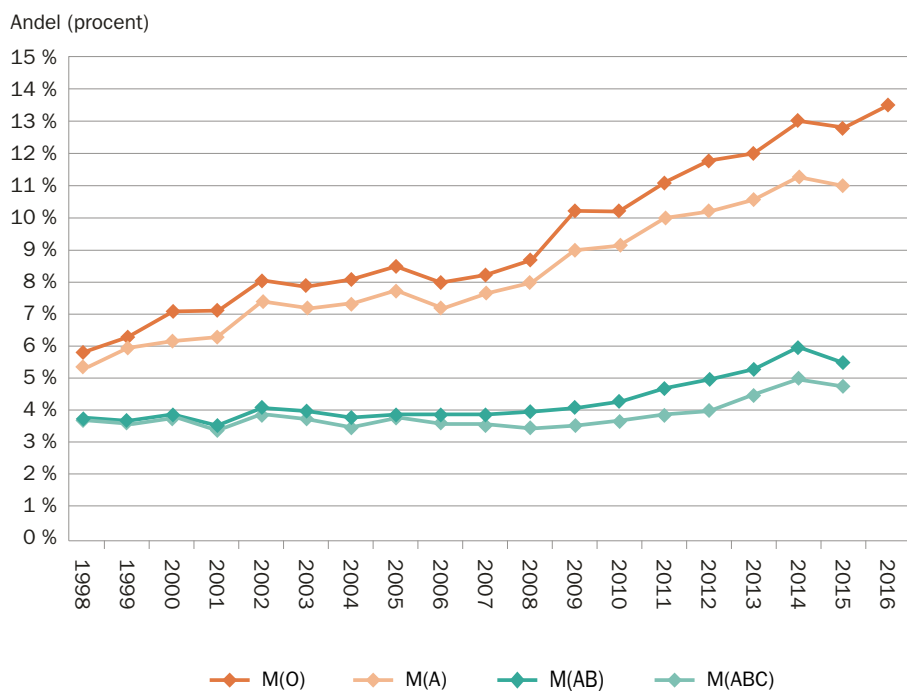
154 Rubie-Davis (2010).

155 Skolverket (2012) s. 79, Böhlmark & Holmlund (2012) s. 39, Holmlund m.fl. (2014) s. 28.

B4.3 Hur mycket minskar mellanskolvariationen när skolnivåeffekter utifrån socioekonomisk bakgrund läggs till modellen?

När skolornas genomsnittliga socioekonomiska sammansättning läggs till modellen som framgår av figur 4.3 i rapportens huvudtext, sjunker mellanskolvariationen ytterligare något (figur B4.4). Det innebär att en del av de skillnader som finns i resultat mellan skolor kan förklaras med skolnivåeffekter med avseende på socioekonomisk bakgrund, utöver det som kan förklaras av att elevernas egna individuella bakgrund skiljer sig. Samtidigt visar figuren att mellanskolvariationen har ökat under senare år, även efter kontroll för skolnivåeffekter som beror på skolornas socioekonomiska sammansättning. Fram till 2010 ökade förklaringsgraden från 85 till 100 procent av all mellanskolvariation. Under 2010–2015 ökade förklaringsgraden från 57 till 63 procent. Fortfarande finns alltså en stor del oförklarad ökad mellanskolvariation under de senare åren, även om man lägger till skolnivåeffekterna.

Figur B4.4 Återstående mellanskolvariation efter att hänsyn tagits till elevers bakgrund, 1998–2015. M(O): Mellanskolvariation utan kontroll för elevers bakgrund. M(A): Mellanskolvariation efter kontroll för elevers kön och migrationsbakgrund. M(AB): Mellanskolvariation efter kontroll för elevers kön, migrationsbakgrund och socioekonomiska bakgrund. M(ABC): Mellanskolvariation efter kontroll för samtliga variabler i M(AB) och skolans genomsnittliga socioekonomiska sammansättning.*



* I analysen ingår bara de elever som har personnummer och där vi har registeruppgifter om både migrationsbakgrund och föräldrarnas utbildningsnivå, inkomst och grad av bidragstagande. I modell M(A) ingår förutom kön och migrationsbakgrund även variabler som indikerar om eleven är yngre respektive äldre än typåldern för respektive årskull. I modell M(AB) baseras den socioekonomiska bakgrunden på det socioekonomiska indexet, som i sin tur baseras på föräldrarnas utbildningsnivå, inkomst och grad av bidragstagande. Tidserien för M(A), M(AB) respektive M(ABC) är begränsad till 2015, eftersom variablerna inkomst och bidragstagande bara är tillgängliga fram till 2015. Data viktat för skolstorlek.

Bilaga B5

B5.1 Variation i resultat mellan elever

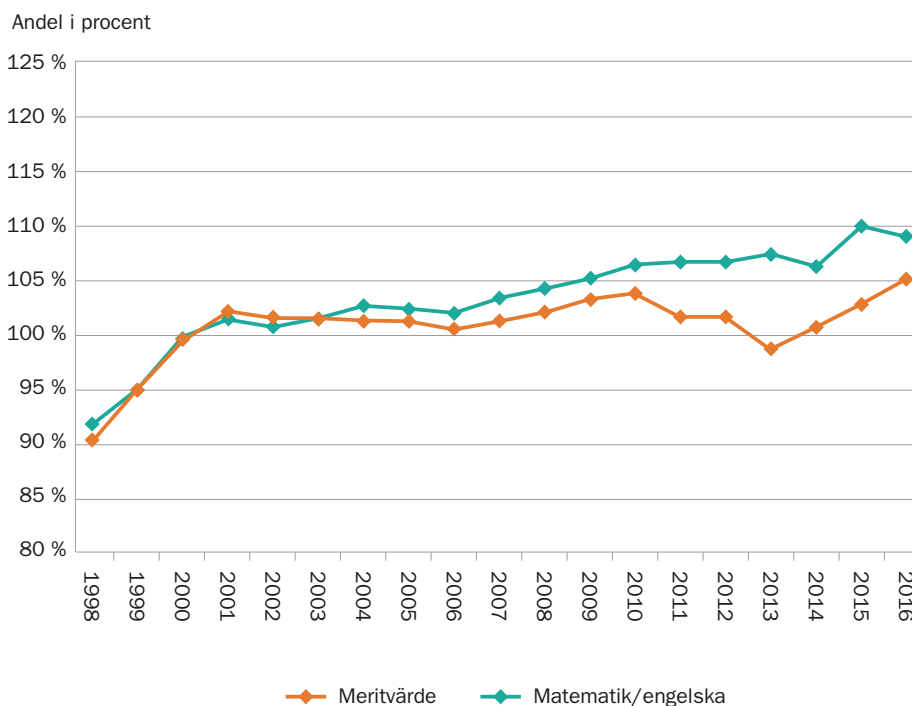
I detta avsnitt undersöker vi hur stora resultatskillnaderna är mellan elever totalt sett, oavsett deras socioekonomiska bakgrund och oavsett vilka skolor de går på.

Om betydelsen av den socioekonomiska bakgrunden ökar (som kapitel 2 visar), bör vi då inte se en ökad spridning i elevers betygsresultat, allt annat lika? Om vi utgår från en konstant spridning i den socioekonomiska bakgrunden i kombination med att betydelsen av den socioekonomiska bakgrunden ökar, är det rimligt att förvänta sig att resultatskillnaderna mellan elever ökar, dvs. det blir större skillnader mellan de högst respektive de lägst presterande eleverna.

På samma sätt kan man argumentera för att om mellanskolvariationen ökar, antingen på grund av ökade kvalitetsskillnader eller som en konsekvens av ökad skolsegregation i kombination med kamrateffekter, bör vi också förvänta oss att resultatskillnaderna mellan elever ökar.

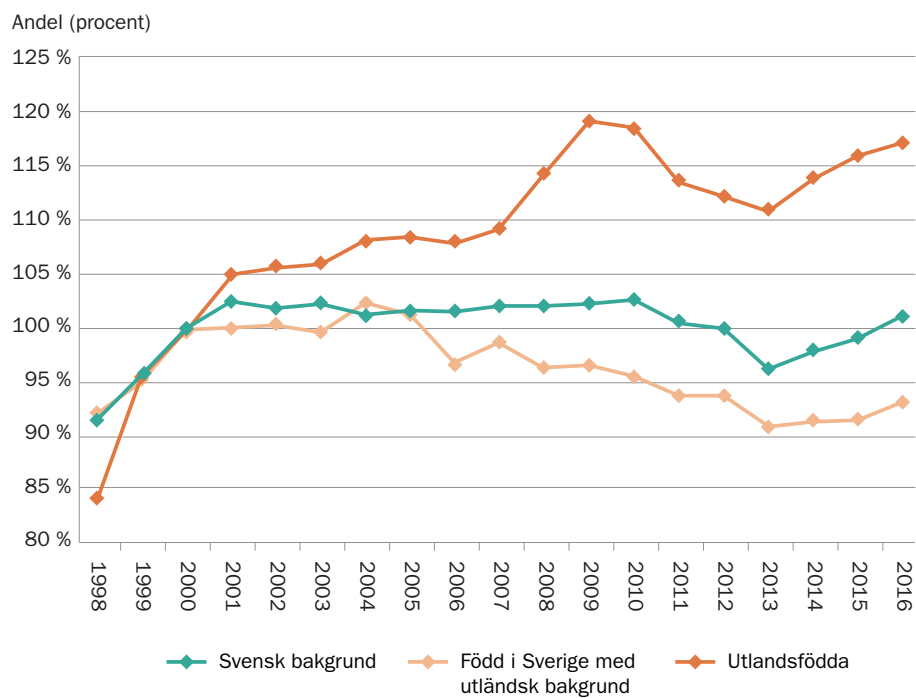
Figur B5.1 visar den totala variationen, mätt som standardavvikelsen, för elevernas meritvärde respektive betygen i engelska och matematik. Tidserierna är justerade för övergången till den nya betygsskalan 2013 och är uttryckta så att den totala variationen/spridningen år 2000 är lika med 100 procent. Figur B5.2a–b visar motsvarande spridning, nu uppdelat efter elevernas migrationsbakgrund.

Figur B5.1 Total variation (standardavvikelse) i meritvärde respektive betygen i matematik och engelska, samtliga elevgrupper.*



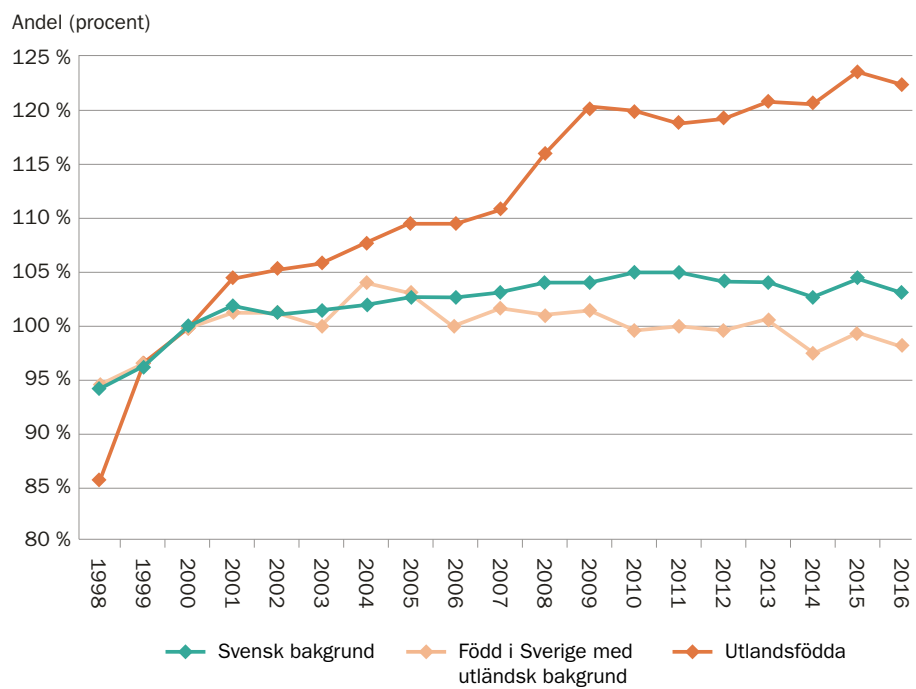
* Standardavvikelse är standardiserade med år 2000 som basår (= 100 %). Standardavvikelse är justerade för bytet av betygsskala 2013.

Figur B5.2a Total variation (standardavvikelse) i meritvärde uppdelat på migrationsbakgrund.



* Standardavvikelse är standardiserade med år 2000 som basår (= 100 %). Standardavvikelse är justerade för bytet av betygsskala 2013.

Figur B5.2b Total variation (standardavvikelse) i betygen i matematik och engelska uppdelat på migrationsbakgrund.



* Standardavvikelse är standardiserade med år 2000 som basår (= 100 %). Standardavvikelse är justerade för bytet av betygsskala 2013.

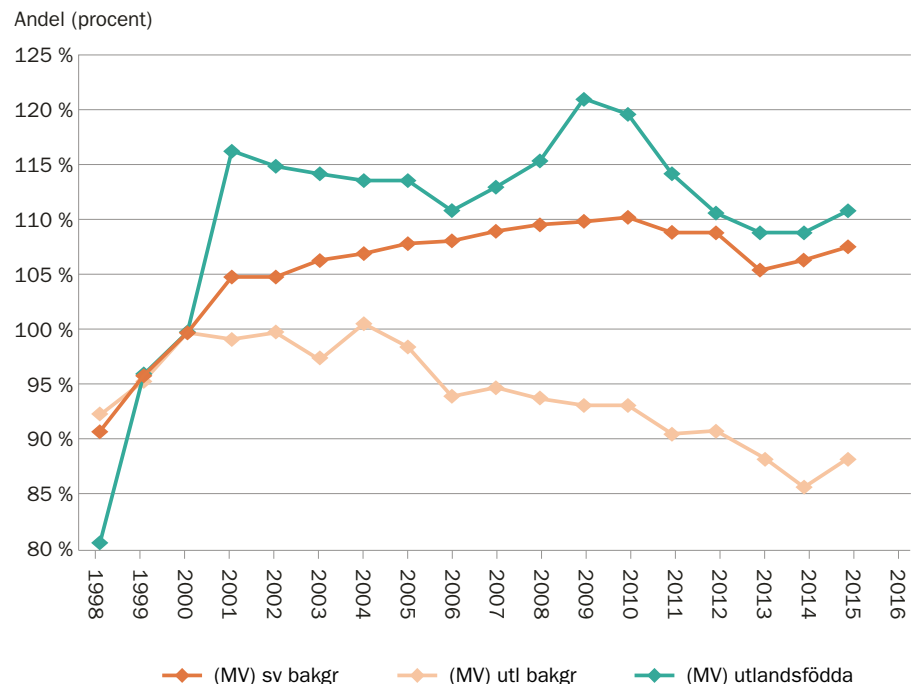
I inledningen av avsnittet nämndes att vi kunde förvänta oss att den totala variationen i elevers resultat borde öka, eftersom den socioekonomiska bakgrunden får större betydelse för elevers resultat. Figur B5.1 och B5.2a–b bekräftar bara delvis den bilden.

När det gäller samtliga elever ser vi mycket riktigt en ökad total variation i elevresultat. När vi delar upp analysen på olika migrationsgrupper stämmer bilden även för utlandsfödda elever som visar en betydande ökning i resultatvariation. Detta är i linje med att betydelsen av socioekonomisk bakgrund också ökat betydligt för den gruppen. När det gäller gruppen elever som är födda i Sverige med utländsk bakgrund såg vi en svag ökning i betydelse, men figur B5.2b visar på en minskad resultatvariation. Även för gruppen elever med svensk bakgrund har betydelsen av socioekonomisk bakgrund ökat något, men vi ser ingen förändring i resultatvariationen för meritvärdet och bara en svag ökning av matematik- och engelskabetyget. Hur ska detta förstås?

Ovanstående resonemang gäller som redan påpekats under förutsättningen att spridningen i den socioekonomiska bakgrunden är oförändrad. Men det har skett förändringar i spridningen i detta mått. För elever med svensk bakgrund har spridningen i den socioekonomiska bakgrunden minskat. Om betydelsen av den socioekonomiska bakgrunden för resultaten ökar samtidigt som spridningen i den socioekonomiska bakgrunden minskar, är det inte säkert att den totala variationen/spridningen i betygsresultat ökar.

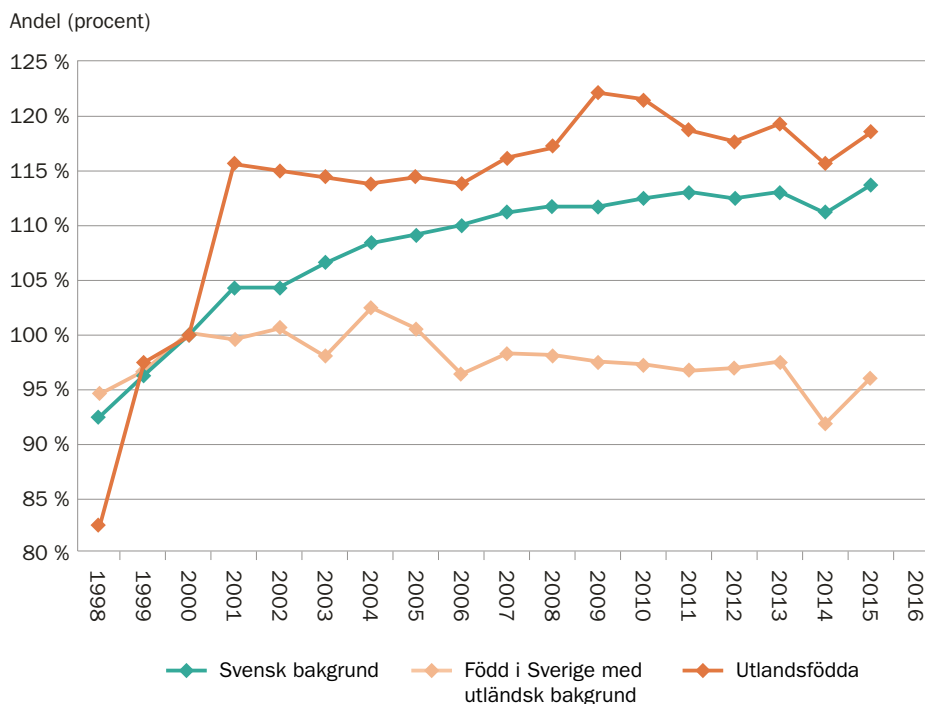
Figur B5.3a–b visar den totala resultatvariationen i meritvärde respektive betygen i matematik och engelska efter att hänsyn tagits till att spridningen i det socioekonomiska indexet har förändrats för respektive migrationsgrupp.

Figur B5.3a Total variation (standardavvikelse) i meritvärde uppdelat på migrationsbakgrund, efter kontroll för förändrad utbildningsbakgrund över tid inom respektive migrationsgrupp.



* Standardavvikelseerna är standardiserade med år 2000 som basår (= 100 %). Standardavvikelseerna är justerade för bytet av betygsskala 2013.

Figur B5.3b Total variation (standardavvikelse) i betygen i matematik och engelska uppdelat på migrationsbakgrund, efter kontroll för förändrad utbildningsbakgrund över tid inom respektive migrationsgrupp.



* Standardavvikelseerna är standardiserade med år 2000 som basår (= 100 %). Standardavvikelseerna är justerade för bytet av betygsskala 2013.

Om vi jämför figur B5.3a med figur B5.2a för respektive migrationsgrupp, ser vi nu att ökningen i den totala variationen för gruppen utlandsfödda är något mindre när hänsyn tas till att sammansättningen med avseende på utbildningsnivå har förändrats över tid, speciellt för de sista åren. För elever med svensk bakgrund blir i stället resultatvariationen nu större (+8 %), när hänsyn tas till att fördelningen med avseende på föräldrarnas socioekonomiska bakgrund blivit mer sammanpressad.¹⁵⁶ För elever födda i Sverige med utländsk bakgrund ser vi fortfarande en minskad resultatspridning.

Den viktiga poängen här är att den bild vi såg med en ökad betydelse av den socioekonomiska bakgrunden också får genomslag i en ökad total resultatvariation för både utlandsfödda och elever med svensk bakgrund, när vi tar hänsyn till förändringar i fördelningen av den socioekonomiska bakgrunden (spridningen i det socioekonomiska indexet).

När det gäller elever födda i Sverige med utländsk bakgrund är det svårare att förklara varför resultatvariationen minskar, även om betydelsen av socioekonomisk bakgrund för denna grupp visade på en mer osäker utveckling över tid. Det är också svårt att förklara varför spridningen i resultat temporärt går ned för både utlandsfödda och elever med svensk bakgrund mellan 2010 och 2013.

¹⁵⁶ Observera att denna siffra ska jämföras med -1 procent, eftersom då det är 2015 som jämförs. Det finns då inga data för 2016.

Observera dock att dessa justerade tidserier inte gör anspråk på att visa hur den totala spridningen faktiskt ser ut, utan bara illustrerar hur den totala variationen kan förväntas ha sett ut om spridningen i den socioekonomiska bakgrunden (indexet) varit konstant.

Referenser

- Adermon, Adrian, Mikael Lindahl & Mårten Palme (2016). *Den utvidgade familjens betydelse för barns utbildning*. IFAU rapport 2016:16.
- Andersson, Eva, Bo Malmberg & John Östh (2012). "Travel-to-school distances in Sweden 2000–2006: changing school geography with equality implications", *Journal of Transport Geography*, 23:35–43.
- Bergsten, Zara & Emma Holmqvist (2013). "Possibilities of building a mixed city – evidence from Swedish cities". *International Journal of Housing Policy*, 13:3, 288–311.
- Björklund, Anders, Mikael Lindahl & Krister Sund (2003). "Family background and school performance during a turbulent era of school reforms." *Swedish Economic Policy Review*. 10 (2003) 111–136.
- Björklund A, Fredriksson P, Gustafsson J-E, Öckert B (2010a). *Den svenska utbildningspolitikens arbetsmarknadseffekter – vad säger forskningen?*. Rapport 2010:13, IFAU.
- Björklund, Anders, Lena Lindahl, & Matthew J. Lindquist (2010b) "What More Than Parental Income, Education and Occupation? An Exploration of What Swedish Siblings Get from Their Parents," *The B.E. Journal of Economic Analysis & Policy*: Vol. 10: Iss. 1 (Contributions), Article 102.
- Brese, Falk & Plamen Mirazchiyski (2013). "Measuring Students' Family Background in Large-Scale International Education Studies." *Issues and Methodologies in Large-Scale Assessments*. Special Issue 2, IERI Monograph Series.
- Böhlmark, Anders & Helena Holmlund (2011). *20 år med förändringar i skolan: Vad har hänt med likvärdigheten?*, SNS.
- Böhlmark, Anders & Helena Holmlund (2012). *Lika möjligheter? Familjebakgrund och skolprestationer 1988–2010*, IFAU, Rapport 2012:14.
- Böhlmark, Anders, Helena Holmlund & Mikael Lindahl (2015). *Skolsegregation och skolval*. IFAU-rapport 2015:5.
- Böhlmark, Anders, Helena Holmlund & Mikael Lindahl (2016). "Parental choice, neighbourhood segregation or cream skimming? An analysis of school segregation after a generalized choice reform." *Journal of Population Economics*, 2016, vol. 29, issue 4, No 7, 1155–1190.
- Chetty, Raj, Nathaniel Hendren & Lawrence F. Katz (2016). "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 2016, 106(4): 855–902.
- Engzell, Per (2016a). *Intergenerational Persistence and Ethnic Disparities in Education*. Avhandling, sociologiska institutionen, Stockholms universitet.
- Engzell, Per (2016b). What do Books in the Home Proxy for? A Cautionary Tale. SOFI, working paper 1/2016.

- Engzell, Per, & Jan O. Jonsson (2015). Estimating Social and Ethnic Inequality in School Surveys: Biases from Child Misreporting and Parent Nonresponse. *European Sociological Review*, vol. 31, iss. 3.
- Enkvist, Victoria, Lotta Lerwall, Olle Lundin & Gustaf Wall (2017). ”En likvärdig utbildning”, i Scheutz, Sverker red. (2017). *Likvärdig utbildning*, Iustus förslag 2017.
- Figlio, David N. & Umut Özek (2017). *Unwelcome guests? The Effects of Refugees on the Educational Outcomes of Incumbent Students*. National Bureau of Economic Research. NBER working paper 23661.
- Flood, Lennart & Joakim Ruist (2015). ”Migration, en åldrande befolkning och offentliga finanser” Bilaga 6 i *Långtidsutredningen 2015*, SOU 2015:95.
- Fredriksson, Peter & Jonas Vlachos (2011). *Reformer och resultat: Kommer regeringens utbildningsreformer att ha någon betydelse?* Finanspolitiska rådet. Studier i finanspolitik 2011/3.
- Förordning om läroplan för grundskolan, förskoleklassen och fritidshemmet (SKOLFS 2010:37)
- Grönqvist, Hans & Susan Niknami (2017a). *Ankomst och härkomst – en ESO-rapport om skolresultat och bakgrund*. ESO-rapport 2017:3.
- Grönqvist, Hans & Susan Niknami, (2017b). ”The School Achievements of Refugee Children: Lessons from Sweden”. *Nordic Economic Policy Review*, 2017:520, pp. 159–185.
- Gustafsson, Jan-Eric, Kajsa Yang Hansen & Monica Rosén (2013). ”Effects of home background on student achievement in reading, mathematics, and science at the fourth grade.” In M. O. Martin & I. V. S., Mullis (red.), *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade—implications for early learning*. Boston College, USA, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Gustafsson, Jan-Eric, Christina Cliffordsson & Gudrun Erickson (2014). *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. SNS förlag.
- Gustafsson, Jan-Eric & Kajsa Yang Hansen (2017). *Changes in the Impact of Family Education on Student Educational Achievement in Sweden 1988–2014*, *Scandinavian Journal of Educational Research*.
- Gustafsson, Jan-Eric, Sverker Sörlin & Jonas Vlachos (2016). *Policyidéer för svensk skola*. SNS förlag.
- Hansson, Åse & Jan-Eric Gustafsson (2016). *Pedagogisk segregation: Lärar-kompetens i den svenska grundskolan ur ett likvärdighetsperspektiv*. Pedagogisk forskning i Sverige årg 21 nr 1–12 2016.
- Hattie, John (2009). *Visible Learning*. Routledge, London and New York.
- Heller Sahlgren, Gabriel (2017). *Likvärdigheten i PISA – förändringar och förklaringar*. Svenskt näringsliv.

- Holmlund, Helena, Josefin Häggblom, Erica Lindahl, Sara Martinson, Anna Sjögren, Ulrika Vikman & Björn Öckert (2014). *Decentralisering, skolval och fristående skolor: resultat och likvärdighet i svensk skola*. IFAU, Rapport 2014:25.
- Holmlund, Helena (2016). *Education and equality of opportunity: what have we learned from educational reforms?* IFAU working paper 2016:5.
- Hällsten, Martin & Fabian T. Pfeffer (2017a). Familjens förmögenhet och barnbarnens skolprestationer. IFAU-rapport 2017:3.
- Hällsten, Martin & Fabian T. Pfeffer (2017b). "Grand Advantage: Family Wealth and Grandchildren's Educational Achievement in Sweden". *American Sociological Review* 2017, Vol. 82(2) 328–360.
- Karbownik, Krzysztof & Sara Martinson (2014). Svenska högstadie- och gymnasielärares rörlighet på arbetsmarknaden. IFAU, Sammanfattning av Rapport 2014:11.
- Malmberg, Bo, Eva Andersson & Zara Bergsten (2013). "School choice motives: the effects of class and residential context". *Annals of the Association of American Geographers*. Januari 2013.
- Malmberg, Bo, Michael M. Nielsen, Eva Andersson & Karen Haandrikman (2016). *Residential Segregation of European and Non-European Migrants in Sweden 1990–2012*. ResSegr Working Paper 2016:1, Stockholm University, Department of Human Geography.
- OECD (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*, OECD Publishing, Paris.
- OECD (2017), *Educational Opportunity for All: Overcoming Inequality throughout the Life Course*, OECD Publishing, Paris.
- Parker, Philip D, John Jerrim, Ingrid Schoon & Herbert W. Marsh (2016). *A Multination Study of Socioeconomic Inequality in Expectations for Progression to Higher Education: The Role of Between-School Tracking and Ability Stratification*. *American Educational Research Journal*, February 2016, No 1. Pp. 6–32.
- Proposition 2009/10:165. Den nya skollagen - för kunskap, valfrihet och trygghet.
- Rubie-Davies, Christine M. (2010). *Teacher expectations and perceptions of student attributes: Is there a relationship?* *British Journal of Educational Psychology* (2010), 80, 121–135.
- Rutkowski, Leslie (2011). "The Impact of Missing Background Data on Subpopulation", *Estimation Journal of Educational Measurement*. Fall 2011, Vol. 48, No. 3, pp. 293–312.
- Rutkowski, Leslie & David Rutkowski (2010). "Getting it 'better': the importance of improving background questionnaires in international large-scale assessment", *Journal of Curriculum Studies*, 42:3, 411–430
- Rutkowski, David & Leslie Rutkowski (2013). "Measuring Socioeconomic Background in PISA: one size might not fit all", *Research in Comparative and International Education* Volume 8 Number 3 2013.

SCB (2017). *Inkomstrappport 2015: individer och hushåll*. Ekonomisk välfärdsstatistik 2017:1. Statistiska centralbyrån.

Schulz, W. m.fl. (2017). *Becoming citizens in a changing world. IEA International Civic and Citizenship Education Study (ICCS) 2016 International Report*. Amsterdam: IEA.

Sirin, Selcuk R (2005). "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research." *Review of Educational Research* Fall 2005, Vol. 75, No. 3, pp. 417–453.

Skolverket (2006). *Vad händer med likvärdigheten i svensk skola? – En kvantitativ analys av variation och likvärdighet över tid*. Rapport 275.

Skolverket (2009). *Vad påverkar resultaten i svensk grundskola? – Kunskapsöversikt om betydelsen av olika faktorer*.

Skolverket (2012). *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid*. Rapport 373.

Skolverket (2014). *TALIS 2013 En studie av undervisnings- och lärmiljöer i årskurs 7–9*. Rapport 408.

Skolverket (2015). *Skolverkets lägesbedömning 2015*. Rapport 421.

Skolverket (2016a). *Invandringens betydelse för skolresultaten*. Skolverkets aktuella analyser 2016.

Skolverket (2016b). *Utvärdering av den nya betygsskalan samt kunskapskravens utformning*.

Skolverket (2016c). *PISA 2015. 15-åringars kunskaper i naturvetenskap, läsförståelse och matematik*. Rapport 450.

Skolverket (2016d). *TIMSS 2015. Svenska grundskoleelevers kunskaper i matematik och naturvetenskap i ett internationellt perspektiv*. Rapport 448.

Skolverket (2017a). *Skolverkets lägesbedömning 2017*. Rapport 455.

Skolverket (2017b). *PIRLS 2016. Läsförståelse hos svenska elever i årskurs 4 i ett internationellt perspektiv*. Rapport 463.

Skolverket (2017c). *Elever och skolenheter i grundskolan läsåret 2016/17*. PM 23 mars 2017, Dnr 2016:1320.

SOU 2016:9: *Plats för nyanlända i fler skolor*.

SOU 2016:66: *Det stämmer! Ökad transparens och mer lika villkor*.

SOU 2017:35: *Samling för skolan. Nationell strategi för kunskap och likvärdighet*.

Yang Hansen, Kajsa & Jan-Eric Gustafsson (2016). *Causes of educational segregation in Sweden – school choice or residential segregation*, Educational Research and Evaluation.

Östh, John, Eva Andersson & Bo Malmberg (2013). "School Choice and Increasing Performance Difference: A Counterfactual Approach". *Urban Studies*, 50(2), pp. 407–425.

Det är sedan tidigare känt att elevernas familje-
bakgrund har stor betydelse för deras skolresultat,
men hur har betydelsen utvecklats över tid? I vilken
utsträckning klarar det svenska skolsystemet att kom-
pensera för elevers olika förutsättningar att nå målen
i slutet av grundskolan? Det är frågor som undersöks
i denna rapport.

Rapporten undersöker även utvecklingen av skillnader
mellan skolor, både skillnader i resultat och elev-
sammansättning. Har skolorna blivit mer segregerade
utifrån elevernas bakgrund? Har resultatskillnaderna
mellan skolor fortsatt att öka? Har skolans samman-
sättning blivit mer eller mindre viktigt för en elevs
betygsresultat?

Detta är tredje gången Skolverket tar ett samlat grepp
om dessa frågor. Tidigare rapporter publicerades 2006
och 2012. Rapporten är ett underlag för diskussion
och åtgärder när det gäller likvärdigheten i grundskolan.

Skolverket

www.skolverket.se