



LUNDS UNIVERSITET
Ekonomihögskolan

Linnéskolan och Kroksbäck: Ett naturligt experiment i skolintegration

13 januari 2020

Kandidatexamen
Nationalekonomiska institutionen

Författare:
Malte Kauranen

Handledare:
Andreas Bergh

Innehåll

1	Introduktion	4
2	Teori	6
2.1	Vad påverkar elevernas skolprestationer?	6
2.2	Problemen med betyg som kunskapsmätt	8
2.3	Tidigare undersökningar om Kroksbäcksskolan och Linnéskolan	9
3	Data	9
4	Metod	13
4.1	Modelldesign	13
5	Resultat och analys	17
6	Diskussion och slutsatser	18
6.1	Vidare forskning	19

Abstract

The Swedish school system has gone from one of the least segregated and equal in outcomes in the 90's to being an average OECD country in both categories. To try and understand the relations between the two variables I study the case of Linnéskolan and Kroksbäcksskolan between 2007 and 2013. Linnéskolan is in an affluent area of Malmö where most parents have tertiary education and have a mostly Swedish background. In comparison Kroksbäcksskolan is in a poor area where parents often lack secondary and tertiary education and a majority of the students have a foreign background. During three years between 2007 and 2009 the students from Kroksbäck who were about to start the sixth grade started at Linnéskolan instead. This paper studies the grades of the students at Linnéskolan during the years these students graduate and whether changing school made these students perform better than expected from a student body of their composition. Using a difference-in-difference approach with controls for parents' education and background, no significant result on grades is found.

1 Introduktion

Skolsegregation är ett viktigt och mycket diskuterat ämne inom skolforskning. Stora delar av den tidiga litteraturen är fokuserad på USA på grund av domstolsfallet *Brown vs. Board of Education*. USA:s högsta domstol hade tidigare en doktrin att separata skolor och institutioner för svarta och vita var laglig enligt deras grundlag, men slog nu fast att skolor inte kunde vara skiljda och jämlika (Reports 1954). Det här startade en debatt som pågår än idag i USA där man fortfarande försöker klargöra hur mycket av skillnaderna i kunskapsnivån som beror på segregation som i till exempel Hanushek, Kain och Rivkin (2009).

I Sverige är det här en nyare fråga. När den första PISA-rapporten släpptes hade Sverige en för OECD väldigt låg nivå av skolsegregation (PISA 2003). I en sådan här situation är inte segregation en viktig politisk fråga och ligger långt ner bland prioriteterna i forskning. Sedan dess har situationen i Sverige förändrats. I den senaste rapporten släppt 2019 så hade Sverige en i många avseenden genomsnittlig nivå av skolsegregation (PISA 2019). Det har därför blivit mer relevant att förstå vilka effekter skolsegregation har på eleverna och vilka det är som blir segregerade.

Skolsegregationen i Sverige är inte av samma lagstiftade natur som den var i USA innan det blev förklarat olagligt. Med en segregerad skola i Sverige menar jag därmed inte en skola där vissa elever på grund av deras bakgrund inte får gå då inga sådana skolor får existera i Sverige. Böhlmark, Holmlund och Lindahl (2015) definierar segregation som när elever blir överexponerade mot elever av sin egen bakgrund. För att göra en sådan här jämförelse delar de upp eleverna i grupper baserat på olika egenskaper. Dessa egenskaper kan till exempel vara om en elev har svensk eller utländsk bakgrund. En överexponering betyder att i elevens klass så är någon sådan egenskap mer frekvent än vad den är i skolväsendet i stort. I ett integrerat skolväsende blir elever inte överexponerade av andra elever från samma bakgrund. Böhlmark, Holmlund och Lindahl (2015) finner att enligt detta mått har segregationen ökat i Sverige mellan 1993 och 2009. Ökningen av skolsegregationen i Sverige är först och främst en effekt av ökad boendesegregation men också av det fria skolvalet (Böhlmark, Holmlund och Lindahl 2015).

Sveriges skolor har ett uppdrag att ge alla elever utbildning av samma kvalitet (Auer och Sandqvist 2012). Att Sveriges elever delas upp i olika skolor baserat på bakgrund är inte nödvändigtvis ett problem. Men om skolorna har olika kvalitet och en elev med givna förutsättningar skulle få olika betyg beroende på vilken skola den går på, något som dessvärre stämmer för

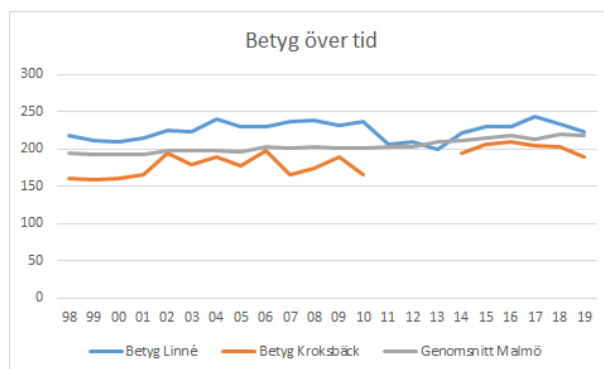
Sverige i ökande grad de senaste decennierna (Auer och Sandqvist 2012), misslyckas skolan i det här uppdraget. Ju större betydelse en skola har för vilket betyg en elev får desto mer blir skolsegregationen ett fördelningsproblem där vissa elever från tidigt i livet får institutionella fördelar som hjälper dem livet ut.

Det här betyder inte att skolsegregation sänker genomsnittsbetygen. Men det finns viss anledning att tro att den gör även det. Szulkin och Jonsson (2004) visar att höga koncentrationer av elever med utländsk härkomst ger negativa effekter på alla elever med utländsk härkomst i klassen. Skolsegregationen blir inte bara en fråga om rättvisa fördelningar utan kan också sänka den genomsnittliga kunskapen av elever som går ut grundskolan. En elevs resultat beror inte nödvändigtvis bara på vilken elev det är, vilken skola den går på men också på vilka andra elever som finns i dess klass. Klasskamraters effekt på en elevs skolresultat kallas för kamrateffekter.

För att undersöka kamrateffekter, skolkvalitet och segregation i den svenska skolan har jag som utgångspunkt ett naturligt experiment i Linnéskolan i Malmö stad som pågick under från och med skolstarten efter sommarlovet 2007 fram till sommarlovet 2010 (Wigerfelt 2010). En närliggande skola, Kroksbäcksskolan, skulle renoveras och byggas ut som förberedelse av Hylle som en ny stadsdel i Malmö och därför behövde eleverna under ett antal år gå på en annan skola. Det beslutades av närhetsskäl att de skulle gå på Linnéskolan. 2007 började därmed ungefär hälften av sjundeklassarna på Kroksbäcksskolan samt alla sjätteklassare på Linnéskolan. Det här betyder att betygen för årskurs 9 mellan 2010 och 2013 på Linnéskolan kommer från årskullar då elever har blivit blandade med varandra.

Eftersom det här är en tillfällig intervention som blandar två annars segregerade skolor kan det här experimentet möjligtvis klargöra om integration av skolor har tydligt negativa eller positiva effekter på elevernas skolresultat. Uppsatsens mål är därmed att studera om skillnader i skolkvalitet eller kamrateffekter påverkade eleverna som gick sina sista år i grundskolan i integrerade klasser.

Innan en mer komplex analys görs är det värt att titta på figur 1. Här kan man se att Linnéskolan systematiskt presterar bättre än Malmös genomsnitt bortsett från åren under experimentet. Kroksbäcksskolan har däremot lägre betyg än genomsnittet under hela perioden.



Figur 1: Genomsnittligt meritpoäng över tid för de elever som går ut årskurs 9 i Malmö stad mellan 1998 och 2019. År på y-axeln. Data hämtad från Skolverkets databas.

2 Teori

Många faktorer påverkar elevers skolresultat. Den här delen av uppsatsen gör en litteraturgenomgång och diskuterar vad som är relevant och sätter därmed vad den resterande delen av uppsatsen diskuterar i kontext. Jag diskuterar även vissa problem med vad man vanligtvis mäter, betyget, och det man är egentligen intresserad av vilket är elevernas kunskapsnivå.

2.1 Vad påverkar elevernas skolprestationer?

En av de saker som är allra bäst på att förklara hur en individ presterar i skolan är individens egenskaper och dess omgivning. Det här inkluderar till exempel elevens studiemotivation, tidigare kunskaper och föräldrarnas förståelse för utbildningsprocessen och engagemanget i den (Hattie 2008). Att mäta individens egenskaper är svårt då det sällan testas och matchas med skolresultat, men som tur är för oss som nationalekonomer så bestäms mycket av de individuella egenskaperna av individens föräldrars egenskaper som ofta sammanställs. Det finns otaliga faktorer som påverkar eleverna men några av de mest relevanta är föräldrarnas socioekonomiska status (Hattie 2008), hemmiljö (Hattie 2008) och traumatiska och stressfyllda händelser (Patel m. fl. 2016).

Socioekonomisk status är ett begrepp som i stort fångar familjens status i samhället. Det här inkluderar konkreta faktorer så som inkomst och utbildningsnivå men också upplevd status av föräldrarnas yrke. Det ter sig naturligt att ett barn med föräldrar som har högre inkomst presterar bättre i skolan. Föräldrarna med mer inkomst har mer resurser att lägga på barnen än föräldrar med lägre inkomst, allt annat lika.

Hemmiljö här är faktorer som är direkt kopplade till föräldrarnas stöd till elevens skolgång. Det här kan vara hjälp med läxor, mer engagerat deltagande i föräldramöten och mer intellektuell stimulans under uppväxten (Hattie 2008).

De sista faktorerna, traumatiska upplevelser och stress i hemmet, kan ta många olika former. I diskussioner om skolor brukar det ofta fokusera på antingen skilsmässor eller delat föräldraskap som i Hattie (2008) men det är även relevant med till exempel invandrare som tappat kontakten med stora delar av släkten eller upplever trauma före eller under flykt till landet där de sedan går i skolan (Patel m. fl. 2016).

Inga av dessa effekter är direkt kopplade till skolsegregation men ändå kritiska för att förstå eventuella effekter av skolsegregation. Eftersom segregation för det mesta är uppdelat på något som är korrelerat med de här faktorerna så kommer vissa skolor ha klasser som består av elever som presterar bättre än andra elever även om skolorna har exakt samma kvalitet. Därför måste en kontroll göras för att sammansättningarna på skolorna och att dem verkligen presterar bättre eller sämre väntat utifrån de individuella egenskaperna av eleverna på skolan.

Efter en sådan här kontroll kan skolkvalitet och kamrateffekter undersökas. Det här är två av de faktorer som kan förklara varför vissa elever med likadana förutsättningar presterar olika när de hamnar i olika skolor eller klasser.

Skolkvalitetens effekt är formellt hur stor del av elevernas kunskap som beror på var de går i skolan. Det här har flera möjliga förklaringar. Den simplaste kausala förklaringen är att skolor är olika bra på att ge elever kunskap. Det här kan vara genom att det slumpmässigt är många bra lärare på en skola. En annan möjlighet är att mekanismer som inte är uppmätta sorterar eleverna på skolor. Föräldrarna som ger sina barn bra förutsättningar kanske väljer skolor på ett sätt som är annorlunda än för andra. Oavsett vilken av de här anledningarna det är så är det en allt viktigare faktor för att förstå svenska elevers resultat i grundskolan (Auer och Sandqvist 2012).

Kamrateffekter är istället hur mycket av elevernas kunskap som beror på vem de går i skolan med. Högpresterande elever har en tendens att gynnas av att gå i skolan med många andra högpresterande elever och lågpresterande elever missgynnas av att gå med många andra lågpresterande elever (Burke och Sass 2013). Szulkin och Jonsson (2004) visar att i Sverige när elever med utländsk bakgrund blir en viss andel av klassen ger det alla elever med utländsk bakgrund negativa effekter. Det här betyder att även i Sverige har ickelinjära kamrateffekter hittats. Litteraturen är dock inte entydig. Hanushek, Kain och Rivkin (2009) finner signifikanta

linjära kamrateffekter bland svarta studenter i USA men inte bland andra. Angrist och Lang (2004) finner bara kortvariga och små effekter som inte varar hela elevens skolgång.

Det finns även en möjlighet att elevsammansättningen påverkar skolmiljön på andra mer strukturella sätt. Det här kan till exempel vara att bra lärare väljer bort skolan, föräldrar är mindre engagerade i skolans verksamhet eller att skolan blir bortprioriterad av kommunen. Men vi har goda skäl att anta att så inte är fallet för Linné under de här åren. Dessa effekter borde inte vara systematiska i sin natur eftersom alla inblandade visste att blandningen skulle vara tillfällig (Wigerfelt 2010). Det är i högsta grad möjligt att elever har valt bort Linnéskolan under de här åren men det är mer osannolikt att lärare slutar eller att det sker långvariga förändringar i skolkultur.

Vad som är intressant med det här är hur det påverkar olika elever på olika sätt och att det inte behöver vara linjärt. Tillsammans med ett segregerat skolsystem kan det här betyda att skadliga fördelningar är vanligare än nödvändigt, dock möjligtvis med andra positiva kamrateffekter. Det här gör det till en högaktuell forskningsfråga. Om beslutsfattare kan korrekt identifiera när segregation blir skadligt skulle det i teorin gå att undvika dessa fall genom att omfördela elever. Men det första steget för att göra det här är att förstå när det sker och att skilja det från andra effekter.

2.2 Problemen med betyg som kunskapsmätt

Alla elever i Sverige som går ut grundskola får ett betyg i samtliga ämnen. Det här betyget är tänkt att det ska vara ett objektiva mått på vilka mål eleven har uppnått i skolans ämnen, men enligt Gustafsson, Cliffordson och Erickson (2014) är det ett väldigt bristande mått. Från Gustafsson, Cliffordson och Erickson (2014, p 8):

Likvärdigheten i betygen uppvisar stora brister. De målrelaterade betygen lider av bristande likvärdighet mellan lärare och skolor, och i synnerhet gymnasiebetygen och vissa av grundskolebetygen är påverkade av inflation.

För en som undersöker betyg som kunskapsmätt är både betygsinflation och att betyg inte är likvärdiga på skolor problem som måste lösas. Inflationen gör det svårt att göra jämförelser över tid då samma kunskapsnivå kan ge olika betyg vid olika tider. Bristen av likvärdighet ökar felkällorna i analysen. Det är svårare att jämföra olika skolor om de gör olika systematiska fel när de bedömer elever. I nationalekonomiska termer kan det betyda heteroskedasticitet eller

autokorrelation i feltermerna. Nationella prov visar sig dela samma problem med godtyckliga bedömningar (Gustafsson, Cliffordson och Erickson 2014).

2.3 Tidigare undersökningar om Kroksbäcksskolan och Linnéskolan

Malmö stad har redan i Wigerfelt (2010) gjort en utredning om det här experimentet. Utredningen har ett genusvetenskapligt perspektiv och fokuserar för det mesta på konflikter mellan elever på skolan, skolan och föräldrar och försöker förklara dem. Den tar dock kortfattat upp gymnasiebehörigheten för de elever som fram tills sjätte klass hade gått i olika skolor Wigerfelt (2010, p 29):

Efter sommarskolans slut visade det sig att 9,9 % (siffran var 14,0 innan sommarskolan) av det totala elevantalet på Linnéskolan inte var behöriga till gymnasiet. Totalt sett har det alltså skett en nedgång från vårterminen 2009 då andelen elever som inte fått behörighet var 4,9 procent. Det finns dock en skillnad i betygsresultat mellan eleverna från Limhamn – Bunkeflo och kroksbäckseleverna som har att göra med de båda elevgruppernas olika förutsättningar. Av eleverna från Limhamn-Bunkeflo har 6,3 % inte nått behörighet till gymnasiet medan 20 % av eleverna från Kroksbäck inte uppnått behörighet. Som jämförelse kan nämnas att i den elevgrupp som gick ut nionde klass 2009 på Kroksbäcksskolan så var 35,7 % av eleverna ej behöriga till gymnasiet. Gruppen från Kroksbäck som gick utan [sic] nian 2010 har alltså uppnått bättre resultat än de nior som har gått ut motsvarande nivå på Kroksbäcksskolan under de senaste åren.

Det bristfälliga i den här analysen är att den inte gör någon analys av signifikans eller robust korregerar för olika förutsättningar för eleverna. Det nämns att de har olika förutsättningar, men inte vilka dessa förutsättningar är eller hur de skiljer sig åt mellan elevgrupperna. En mer robust analys krävs för att analysera om eleverna faktiskt presterade bättre än väntat men kan också möjligtvis klargöra varför det är så.

Ett bidrag av den här uppsatsen blir då att analysera experimentet från ett nationalekonomiskt perspektiv för att komplettera den utvärdering Malmö stad redan har gjort.

3 Data

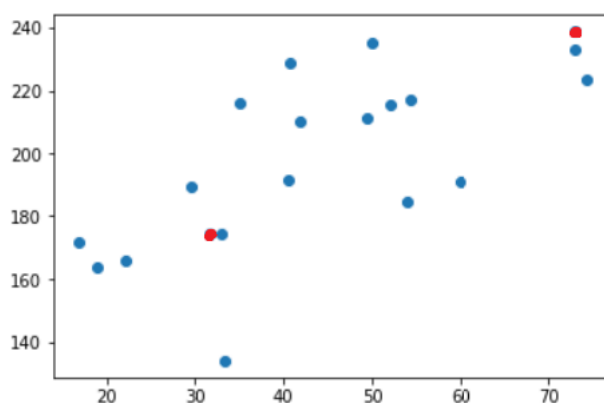
Skolverket tillhandahåller en offentlig databas som publicerar resultatet för alla Sveriges grundskolor sedan 1998. I den här databasen publiceras betygen för varje årskull på varje samt viss information om årskullen elevsammansättning. Informationen om elevsammansättningen

uppgår i flera olika kategorier där resultat även rapporteras för dessa undergruppers prestationer. Dessa undergrupper är elevernas kön, elevernas föräldrars bakgrund samt föräldrarnas utbildningsnivå. Kategorierna för kön är flickor och pojkar, kategorierna för bakgrund är svensk bakgrund, utländsk bakgrund men född i Sverige och utländsk bakgrund född utomlands och för föräldrarnas utbildningsnivå förgymnasial, gymnasial eller eftergymnasial utbildning.

Eftersom jag enbart har tillgång till den offentliga datan över skolorna med årskullarnas sammanfattade resultat och sammansättningen av eleverna på nivå av hela årskullar har jag enbart möjlighet att studera effekterna på skolnivå. Jag kan därför inte studera om en sådan här integration gynnade vissa enskilda individer i en årskull mer eller om den skadade andra individer. Det gör också att jag kommer underskatta eventuella kamrateffekter då de är större på klassnivå (Burke och Sass 2013).

Den här datan har även en viss begränsning då om en årskull på en skola inte har elever med en viss egenskap - till exempel om det inte finns tio elever som har utländsk bakgrund i en årskull på en skola - så rapporteras inte datan för den årskullen. Det här verkar vara på grund av att Skolverket inte vill att informationen ska vara alltför lättillgänglig när det handlar om få individer. Om man till exempel tittar på Apelgårdsskolan data från 2003/2004 så är 47,3% av eleverna födda utomlands och deras föräldrar av utländsk bakgrund. Man får även reda på att det här är 26 elever. Men trots det rapporteras inte resultat för elever med svenska bakgrund eller de vars föräldrar har utländsk bakgrund men eleven själv är född i Sverige. Rent logiskt måste minst en av dem måste ha minst 10 elever då dessa är de enda tre kategorierna som finns. Det här är troligtvis för att om den andra vanliga kategorin också skulle rapporteras kan man göra just det jag skulle vilja göra - avgöra den minsta gruppens storlek och egenskaper och därmed potentiellt kunna identifiera personer. Samma problem gäller för utbildningsnivå.

Det här leder till att jag får inkomplett information om många årskullar från vissa skolor. Om två av tre datapunkter hade varit rapporterade hade jag kunnat rekonstruera den tredje punkten på alla skolor utom de minsta men det visar sig i praktiken omöjligt till en följd av avsaknaden av data. På grund av det här finns det inte information om alla årskullar utan bara 253 årskullar från olika skolor att studera. Det är dessa 253 årskullar som alla uträkningar i uppsatsen är baserade på. Lyckligtvis finns den kompletta datan för Linnéskolan under åren för experimentet. Värt att notera är att datan efter 2014 är givet i ett annat format där det i de nedladdade filerna bara finns det genomsnittliga meritvärdet för 17 istället för 16 ämnen. För att hålla metoden så simpel som möjligt har inte den här datan eller datan från de platser där



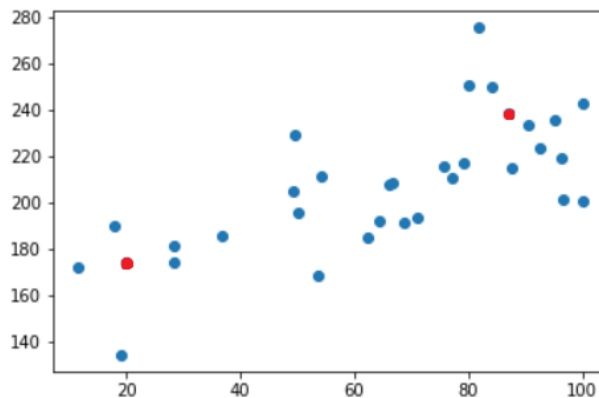
Figur 2: Betyg på y-axeln och andel av barnen där minst en av föräldrarna har eftergymnasial utbildning på x-axeln.

datan har sänkts använts i den här undersökningen.

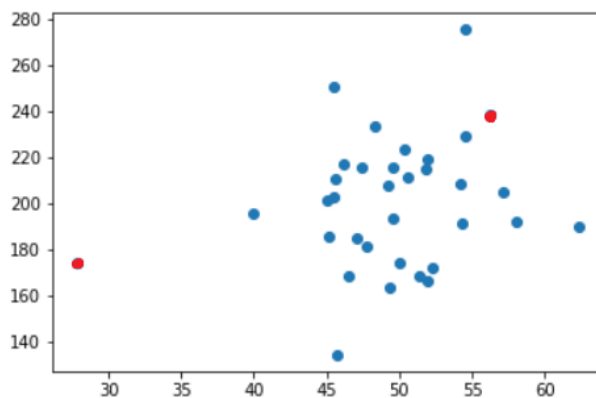
För att ge en överblick över datan och hur de givna undergrupperingarna korrelerar med utfallet, betyget, så visas skolresultaten för årskullen 07/08 från flera Malmöskolor upp. I samtliga diagram så är Kroksbäcksskolan och Linnéskolan markerade med rött, där Kroksbäck är den med genomsnittligt meritpoäng på 177. Man ska vara försiktig med att tolka den här datan alltför mycket på grund av endogenitetsproblem men man kan se tydliga trender i vissa av variablerna. Som man ser i figur 2 och figur 3 så finns ett tydligt samband i andel föräldrar med svensk bakgrund och eftergymnasial utbildning och betyget. I figur 4 med andel flickor så är det inte lika tydligt med en trend.

Det är tydligt hur det kan förklara en del av Linnéskolans nedgång 2010-2013. Under experimentet sjunker andelen elever med svensk bakgrund från 90% till 50% och andelen elever vars föräldrar har akademisk examen från uppemot 70% till 50%. Om dessa variabler bestämmer hur bra elever presterar och elevsammansättningen ändras så att de minskar så är Wigerfelt (2010) påstående att elevernas ändrade förutsättningar var den viktigaste faktorn för de sänkta betygen rimlig.

Viss deskriptiv statistik om Malmös skolor finns samlad i tabell 3. Det här sätter informationen Kroksbäcksskolan och Linnéskolan från punktdiagrammen ovan i lite mer kontext. Notera att minimivärdena på svensk bakgrund och eftergymnasial bakgrund nästan garanterat inte är de verkliga minimumen under tidsperioden på grund av saknad data.



Figur 3: Betyg på y-axeln och andel av barnen där båda föräldrarna är födda i Sverige på x-axeln.



Figur 4: Betyg på y-axeln och andel av barnen som är flickor på x-axeln.

	Genomsnitt	Standardavvikelse	Max	Min
Meritpoäng	202,8	25,6	266,8	121,3
Eftergymnasial utbildning	47,4	16,6	100,0	14,9
Svensk bakgrund	59,7	25,7	100,0	7,7
Flickor	48,7	5,9	66,7	27,9

Figur 5: Deskriptiv statistik om de årskullar i Malmö stad som det finns komplett data för mellan 1998-2019. Betyget är årskullens på skolans genomsnitt, eftergymnasial utbildning och svensk bakgrund är givet som andel av föräldrar som har det och flickor är andel av klassen. All data är tagen från Skolverket.

4 Metod

Som man ser i figur 2, 3 och 4 har elevernas uppmätta egenskaper troligen effekt på elevernas förutsättningar för att inhämta kunskap i skolan. Ramverket som presenterade i teoriavsnittet är nog för att förklara varför både eftergymnasial utbildning och svensk bakgrund har kausala effekter på elevers meritpoäng. Kön har också i tidigare studier funnits ha en väldigt svag men signifikant effekt Hattie (2008). Eftergymnasial utbildning är direkt kopplad till socioekonomisk status och därmed redan belagd som relevant.

Det här lämnar frågan om svenskhet har en effekt på skolresultat. Det finns flera faktorer som agerar här. En är återigen hemmiljön och stödet eleverna får från föräldrarna. En invandrarres föräldrar bör ha mindre erfarenhet av det svenska skolväsendet än någon som själv har gått igenom det och det är rimligt att de inte kommer ha lika bra kunskap över vad som behöver göras från deras sida för att deras barn ska få det bästa stödet hemifrån. Det är också sannolikt att de har fler traumatiska och stressfyllda upplevelser under sin barndom. Invandrare har generellt sett oftare tappat kontakt med delar av sitt familjenätverk och kan ha haft traumatiska och stressfyllda upplevelser innan och under resan till Sverige. Det finns också en möjlighet att föräldrarnas etniska bakgrund är korrelerat med andra faktorer som bestämmer socioekonomisk status som utbildningsnivå inte fullt ut fångar. Svenskhet blir då ett sätt att fånga även dessa faktorer på ett imprecist sätt.

Efter en kontroll för dessa faktorer, om alla skolor har samma kvalitet och elevsammansättning inte ger kamrateffekter så skulle en sådan här blandning inte ge någon effekt. Men det här är ett osannolikt utfall av undersökningen. Auer och Sandqvist (2012) finner att svenska skolor skiljer i kvalitet vilket innebär att elevernas fördelning på skolor spelar roll för deras utbildning. Burke och Sass (2013) finner att kamrateffekter är signifikanta i icke-linjära modeller och påverkar olika elever olika. Effekten av experimentet borde bestämmas av hur dessa tre variabler interagerar och vilken som är störst.

4.1 Modelldesign

Det första steget för att designa en modell är att vi nu mer formellt undersöker de variabler som tros påverka klassens prestationer genom att påverka elevers individuella egenskaper. Det här görs med fördel innan experimentet för att försäkra oss om att vi har en modellspecifikation utan onödiga variabler. För att göra det här används en simpel linjär modell med en konstant

	Eftergymnasial utbildning	Svensk bakgrund	Flickor
Koefficient	1,19***	0,74***	0,11
R^2	0,591	0,550	0,001

Figur 6: Koefficienter för alla variabler när de modelleras var för sig med en simpel linjär modell. Beroende variabel är genomsnittligt meritpoäng. * betyder $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$.

	Modell 2
Eftergymnasial utbildning	0,75***
Svensk bakgrund	0,39***
R^2	0,667

Figur 7: Koefficienterna när andel föräldrar med eftergymnasial utbildning och svensk bakgrund tas med i samma modell. Beroende variabel är genomsnittligt meritpoäng. * betyder $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$.

term, den förklarande variabeln med en koefficient samt en felterm. Formellt är specifikation den följande:

$$y_i = \alpha + \beta x_i + \epsilon_i \quad (1)$$

Där α är en konstant och β är koefficienten för den förklarande variabeln i de tre modellerna med de tre värdena tillgängliga från Skolverket. Resultatet för de här tre regressionerna kan ses i figur 4.1.

Då det finns skäl att tro att eftergymnasial utbildning samt svensk bakgrund hos föräldrarna delvis ger effekter då de variablerna kan fånga upp samma skillnader i socioekonomisk status testas även de i konjunktion. Modellen som testas är den i ekvation 2, där x_{1i} är andelen barn med föräldrar som har svensk bakgrund i en viss klass och x_{2i} är andelen barn där minst en förälder har eftergymnasial utbildning. Resultat kan ses i tabell 7. $R^2 = 0.667$ ger oss också att den här modellen har större förklaringsstyrka än våra tidigare delade modeller.

$$y_i = \alpha + \beta x_{1i} + \gamma x_{2i} + \epsilon_i \quad (2)$$

Det här tyder på att de delvis fångar upp samma effekter men även att de har olika sor-

ters information som båda är relevanta. Troligtvis beror det här på avsaknaden av flera andra okända variabler. Ett stort antal variabler påverkar både en elevs kognitiva utveckling samt dess möjlighet att använda dem effektivt i ett specifikt skolväsen och de två variabler jag har för det är inte fullgoda för att estimeras alla dessa effekter, men då de har en väldigt hög förklaringsgrad för betygen är det klart att de fångar upp en stor del av den individuella variationen.

För att modellera kamrateffekten av experimentet föreslår jag att en difference-in-difference metod används. Experimentet kan tolkas som en behandling av en skola och kan därför modelleras så. Sammanblandningen av elever valdes inte heller nödvändigtvis eftersom det var ett medvetet försök att integrera Malmös skolor heller utan främst från närhet samt kapacitet (Wigerfelt 2010). Det betyder att tilldelningen av behandlingen inte är korrelerad med någon upplevd skolkvalitet eller annat typ av behov av behandling, vilket gör att ett möjligt problem med endogenitet är uteslutet (Besley och Case 2000). Det stora antagandet i difference-in-difference modeller med parallella trender har vi också god anledning att tro det stämmer. Jag påminner om figur 1 och Linnéskolans och Malmö stads trender i den. Linnéskolan har i grafen en nästan ständig nivå en bit över genomsnittet i Malmö stad, bortsett från de åren då Linnéskolan är behandlad.

Difference-in-difference har också en annan kvalitet som är eftertraktansvärt: Utelämnning av variabler ger inte ett bias till estimatören om de är konstanta under tiden eller ändras lika mycket för alla skolor. Till exempel om betygsinflationen är densamma för alla skolor eller om en ny läroplan påverkar alla skolor. Vi har liten anledning att tro att saker som är utelämnade har en stor effekt på resultatet och ändras under tiden.

För modellen vi använder för difference-in-difference så utelämnar vi nu kön då det visar sig ha en minimal effekt på vår datamängd. Vår modell för difference-in-difference blir då den i ekvation 3.

$$y_i = \alpha + \beta x_{1i} + \gamma x_{2i} + \zeta \delta_i + \eta T_i + \theta \delta_i T_i + \epsilon_i \quad (3)$$

α , β och γ har samma betydelse som i ekvation 2. δ_i är 1 för den behandlade skolan, det vill säga Linnéskolan, och 0 för alla andra skolor. T_i är 0 under alla år då experimentet inte pågår men 1 under de fyra relevanta åren. Produkten $\delta_i T_i$ är därmed lika med 1 för Linnéskolan under experimentåren och därmed hur experimentets effekt modelleras.

Anledningen till att jag lägger in mer än bara en variabel för Linnéskolan under de behandlade åren är för att man försöker isolera behandlingseffekten med andra effekter som kan

korrelera med den. Tänk ett scenario där man bara använder en dummyvariabel som är ett underbehandlingsåren för enbart Linnéskolan. Om Linnéskolan är av synnerligen hög kvalitet så skulle en sådan här metod alltid överskatta effekten då skolkvaliteten är korrelerad med vilken skola som blir behandlad. Om det finns en allmän tidstrend i Malmö stad skulle den också delvis fångas upp i den ensamma dummyvariabeln då behandlingen är enbart i den senare delen av tidsskedet. Vice versa om Linnéskolan skulle vara dålig eller Malmö stad skulle ha en nedåtgående trend. De två extra dummyvariablerna garanterar att så inte är fallet då dessa två skeenden bättre förklaras av andra delar av modellen.

Vi kan alltså med den här modellkonstruktionen modellera både Linnéskolans kvalitet och om det sker statistiskt signifikanta förändringar för elevernas betyg under åren av experimentet. Men det finns en tredje väg som Kroksbäckselevernans resultat kan förbättras genom. Om Kroksbäcksskolan innan experimentet var en signifikant sämre skola än Linnéskolan så kommer en flytt till Linnéskolan förbättra deras resultat. För att göra en fullgod undersökning över effekten av experimentet måste även det undersökas. Det här görs genom att vi delvis återanvänder ekvation 2 men lägger till en dummyvariabel. Dummyvariabeln är 1 för årskullar från Kroksbäcksskolan och 0 för alla andra datapunkter. Med δ_i denna gång som dummyvariabel för Kroksbäcksskolan får vi modellen i ekvation 4.

$$y_i = \alpha + \beta x_{1i} + \gamma x_{2i} + \zeta \delta_i \quad (4)$$

För att visa vikten av att ta med både individuella förmågor, skol och tidseffekter så kommer ytterligare några regressioner göras än den i ekvation 3. Anledningen är att visa att man kan få ut mycket av den här datan och att Malmö stad och andra svenska kommuner kan lära sig mycket om den information som finns tillgänglig om deras skolor men att man måste vara noggrann i sin analys.

Den första alternativa modellen i ekvation 5 tar inte elevernas förutsättningar i åtanke. Den har dock med de tre dummyvariablerna från difference-in-difference modellen. Det här är en liknande naiv syn som när Wigerfelt (2010) först presenterar Linnéskolans försämrade gymnasiebehörighet 2010.

$$y_i = \alpha + \zeta \delta_i + \eta T_i + \theta \delta_i T_i + \epsilon_i \quad (5)$$

Den andra modellen i ekvation 6 tar med föräldrars utbildningsnivå som kontrollvariabel, något som inte fullt ut men delvis kontrollerar för elevernas bakgrund.

	Modell 4	Modell 5	Modell 6	Modell 7	Modell 3
Skolkvalitet Linné		31,12***	-1,54		0,19
Tidsdummy		14,23***	2,08		6,18***
Behandlingsdummy		-34,88**	-1,73	2,29	-1,86
Eftergymnasial utbildning	0,75***		1,17***	0,75***	0,68***
Svensk bakgrund	0,40***			0,39***	0,42***
Skolkvalitet Kroksbäck	-1,15				
R^2	0,668	0,101	0,593	0,667	0,677

Figur 8: Koefficienter, signifikans och R^2 för modellerna beskriva i modellavsnittet. * betyder $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$.

$$y_i = \alpha + \beta x_{1i} + \zeta \delta_i + \eta T_i + \theta \delta_i T_i + \epsilon_i \quad (6)$$

Den tredje modellen använder en simplare, ej rigorös difference-in-difference modell där det bara finns en dummy-variabel som är 1 för skolan när den är under behandling och annars alltid 0. Den tar dock elevernas förutsättningar i åtanke. Den kan ses i ekvation 7.

$$y_i = \alpha + \beta x_{1i} + \gamma x_{2i} + \theta \delta_i T_i + \epsilon_i \quad (7)$$

5 Resultat och analys

Alla koefficienter, deras standardfel och R^2 i alla 5 regressioner kan ses i figur 5. Det mest slående med tabellen är att med någon sorts kontroll alls för elevernas förutsättningar så är effekten av experimentet inte signifikant.

Som väntat är modell 5 dålig på att förklara hur elever kommer prestera i skolan. Utan någon sorts information om hur årskullarna är sammansatta är det svårt att förstå hur de kommer att prestera. Experimentet får en starkt negativ signifikant effekt. Detta är naturligt då en stor förklaring till årskullarnas betyg, elevernas förutsättningar, inte finns med i modellen men är starkt korrelerad till tiden för experimentet. Det räcker att man kontrollerar för föräldrarnas utbildningsnivå som i ekvation 6 för att all signifikans för experimentet, Linnéskolan och tidsperioden ska försvinna.

För att förstå vad som händer i regressionen för modell 7 är det nog lättast att börja med modell 3. Kontrollerna för elevernas bakgrund samt tidsdummys är signifikanta. Att tidsdummys är signifikant och positiv tyder på att Gustafsson, Cliffordson och Erickson (2014) har rätt om betygsinflation i den svenska skolan. Den här tidseffekten blir troligen fångad av den ensamma dummyvariabeln i modell 7, koefficienten för andelen föräldrar med eftergymnasial utbildning och koefficienten för svensk bakgrund. Det här är för att andelen föräldrar med svensk bakgrund är negativt korrelerat med tiden och att föräldrar med eftergymnasial utbildning samt tiden för experimentet är positivt korrelerat med tiden. Enligt Skolverkets statistik så hade läsåret 2018/2019 57,1% av eleverna föräldrar med eftergymnasial utbildning och 48,5% med svensk bakgrund i Malmö stad. 1998/1999 var det 38,6% respektive 61,9%.

För den mest robusta regressionen är effekten av experimentet liten och inte signifikant. Som beskrivet i teoridelen så är storleken på resultatet i mycket beroende på hur stora de positiva och negativa kamrateffekterna är. Ickelinjära negativa kamrateffekter för delar av klassen kan ha tagits ut av positiva för andra delar av klassen eller så är icke linjära kamrateffekter försumbara för de här eleverna.

Den andra relevanta information vi får från modell 3 är att skolkvaliteten för Linnéskolan inte är signifikant. Inte heller Kroksbäcksskolan i modell 4. Det tyder alltså på att det mesta av förändringen i betyg i modellen kan förklaras av att färre elever hade svensk bakgrund eller föräldrar med eftergymnasial utbildning under åren av experimentet. Båda skolorna har dock små urval: Det finns bara 6 årskullar från Kroksbäcksskolan som resultatet baseras på och 14 för Linnéskolan.

6 Diskussion och slutsatser

Vad betyder det här för Wigerfelt (2010) analys av gymnasiebehörighet? Det finns ingenting i det här resultatet som tyder på att eleverna under experimentet presterade annorlunda än förväntat eller att flytten till Linnéskolan i sig förbättrade Kroksbäcksskolans elevers genomsnittliga meritpoäng. Man kan däremot inte svara på det med säkerhet då jag inte kan följa eleverna som började på Kroksbäck och jämföra dem som separat grupp från de som gick på Linné från årskurs ett. På grund av de stora begränsningarna i datan är det svårt att dra vidare slutsatser då modellval måste formos utifrån suboptimala förhållanden och datan som existerar.

Den största begränsningen i modellen är att om kamrateffekterna är linjära kommer inte

experimentet heller ge någon effekt i min modell. Detta då jag kontrollerar för kunskaper med genom att lägga in årskullens andel med vissa egenskaper och skattar en linjär effekt från det här. Om andelen med föräldrar som har eftergymnasial utbildning linjärt ger kamrateffekter till alla elever i klassen skulle det modelleras på samma sätt. Modellen kommer alltså uppskatta effekter från individuella förmågor och linjära kamrateffekter i samma variabel. Detta är ett problem som endast kan lösas med data på individnivå.

En annan följd av det här blir att ifall kamrateffekterna är olinjära har de fångats upp ej väntevärdesriktigt av klassammansättningen i modell 3. Detta som en följd av att det enda modellen kan anpassa olinjära kamrateffekter efter på andra skolor än Linnéskolan är den linjära koefficienten. De olinjära effekter på Linnéskolan som inte förklaras väntevärdesriktigt av den linjära koefficienten blir experimentets resultat. Man ska alltså inte dra alltför starka slutsatser om ickelinjära effekter heller.

Ytterligare ett problem med undersökningen är att data inte rapporteras för kategorier av elever där färre än tio personer ingår. I skolor där det nästan bara går elever med utländsk bakgrund eller med låg utbildningsnivå så finns det ingen information om hur många föräldrar som är svenska eller högskoleutbildade. Det här gör att om kamrateffekterna bara blir signifikanta vid höga koncentrationer, vilket Szulkin och Jonsson (2004) hävdar, kan modellen inte anpassa trenden efter de mest segregerade fallen. Men det är också i de fallen som den matematiska modellen får mest information om att kamrateffekter sänker kunskapsnivån. Om man skulle skatta en ickelinjär modell på datan så kommer den inte vara väntevärdesriktig om ickelinjära kamrateffekter existerar.

En annan felkälla som kommer från valet av difference-in-difference metod är att betygen riskerar att vara heteroskedastiska eller autokorrelerade. Men resultatet av sådana här felkällor är att feltermen underskattas (Bertrand, Duflo och Mullainathan 2004). Då ingen signifikans hittades gör det här bara den verkliga betydelsen av experimentet ännu mindre.

6.1 Vidare forskning

Den mest uppenbara vidareutvecklingen hade varit att göra om alla regressioner med data på individnivå och anpassa modellen efter det och att använda fullständig information om Malmös skolor under tidsperioden. Det hade då varit möjligt att utesluta bias på grund av bristande data och man hade kunnat särskilja effekterna från elevernas egna förmågor från eventuella kamrateffekter. Båda dessa tillsammans gör en väldigt robust analys av experimentet möjligt

och hade varit det komplement som Wigerfelt (2010) förtjänar.

En annan intressant vidareutveckling är att följa eleverna längre fram i livet. Eleverna är nu gamla nog för att ha gått ut gymnasiet och kanske börja jobba eller läsa på högskola. Att göra en långsiktig studie om hur experimentet har påverkat dem kan säga mer om resultatet än att bara studera deras meritpoäng.

Referenser

- Angrist, Joshua D och Kevin Lang (2004). "Does school integration generate peer effects? Evidence from Boston's Metco Program". I: *American Economic Review* 94.5, s. 1613–1634.
- Auer, Anders och Jonas Sandqvist (2012). *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid*.
- Bertrand, Marianne, Esther Duflo och Sendhil Mullainathan (2004). "How much should we trust differences-in-differences estimates?" I: *The Quarterly journal of economics* 119.1, s. 249–275.
- Besley, Timothy och Anne Case (febr. 2000). "Unnatural Experiments? Estimating The Incidence of Endogenous Policies". I: *Economic Journal* 110, s. 672–94. DOI: 10.1111/1468-0297.00578.
- Böhlmark, Anders, Helena Holmlund och Mikael Lindahl (2015). *Skolsegregation och skolval*. Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU).
- Burke, Mary A. och Tim R. Sass (2013). "Classroom Peer Effects and Student Achievement". I: *Journal of Labor Economics* 31.1, s. 51–82. DOI: 10.1086/666653. eprint: <https://doi.org/10.1086/666653>.
- Gustafsson, Jan-Eric, Christina Cliffordson och Gudrun Erickson (2014). *Likvärdig kunskapsbedömning i och av den svenska skolan: Problem och möjligheter*. SNS förlag.
- Hanushek, Eric A, John F Kain och Steven G Rivkin (2009). "New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement". I: *Journal of labor economics* 27.3, s. 349–383.
- Hattie, John (2008). *Visible Learning: A Synthesis of Over 800 Meta-analyses Relating to Achievement*. Routledge. ISBN: 9780203887332.

- Patel, Sita G m. fl. (2016). "Newcomer immigrant adolescents: A mixed-methods examination of family stressors and school outcomes." I: *School psychology quarterly* 31.2, s. 163.
- PISA (2003). *Literacy Skills for the World of Tomorrow - Further results from PISA 2000*. OECD.
- (2019). *PISA 2018 Results (Volume II): Where All Students Can Succeed*. OECD.
- Reports, United States (1954). "Brown v. Board of Education". I: *Appeal from the United States*.
- Szulkin, Ryszard och Jan O. Jonsson (2004). "Ethnic Segregation and Educational Outcomes in Swedish Comprehensive Schools: A Multilevel Analysis." I: *Conference Papers – American Sociological Association*, s. 1–20.
- Wigerfelt, Berit (2010). "Kroksbäck möter Linné : en utvärdering av integration mellan två skolor i Malmö". I: FoU Malmö-utbildning, Avdelning barn och ungdom, Malmö stad.