

Effekter av skolledarskap i Sverige

En utvärdering av
Rektorsprogrammet 2009–2021

Rebecka Persson

Ema Demir

Karl Wennberg

Effekter av skolledarskap i Sverige

En utvärdering av Rektorsprogrammet 2009–2021^a

av

Rebecka Persson^b, Ema Demir^c och Karl Wennberg^d

2026-09-23

Sammanfattning

Vi har undersökt effekter av Rektorsprogrammet genom att analysera data från 3 783 deltagare under åren 2009–2021. Studien tyder på att programmet inte ger märkbara förbättringar i elevresultat eller lärares bedömningar av rektorsers pedagogiska ledarskap. Resultaten ligger i linje med slutsatserna från den förhållandevis begränsade internationella forskningen om effekter av rektorsutbildningar. Våra analyser pekar vidare på att rektorer har begränsat inflytande över betygsvariationen jämfört med betydelsen av andra faktorer.

^a Rapporten har seminariebehandlats vid IFAU och granskats av Karin Edmark och Björn Öckert. Vi tackar dem, redaktören Martin Lundin och seminariedeltagarna vid IFAU för värdefulla kommentarer. Vi tackar också Anna Forssell och Anders Duvkär vid Skolverket. Projektet har finansierats med ett forskningsbidrag från IFAU (dnr 155/2021). Ansökan om etikprövning togs inte upp till prövning hos Etikprövningsmyndigheten (dnr 2022-02469-01).

^b Center for Educational Leadership and Excellence (CELE), Handelshögskolan i Stockholm, e-post: rebecka.persson@hhs.se.

^c Center for Educational Leadership and Excellence (CELE), Handelshögskolan i Stockholm, e-post: ema.demir@hhs.se.

^d Center for Educational Leadership and Excellence (CELE), Handelshögskolan i Stockholm, e-post: karl.wennberg@hhs.se.

Innehållsförteckning

1	Inledning	3
1.1	Rektorsers pedagogiska och administrativa ledarskap.....	4
1.2	Svenska och internationella studier om rektorsprogram.....	6
1.3	Det svenska rektorsprogrammet	9
2	Metod och data	11
2.1	Data.....	11
2.2	Urval och population	13
2.3	Statistisk analys	14
3	Resultat	15
3.1	Beskrivande statistik.....	15
3.2	Intraklasskorrelationer – en beskrivning av hur skolor och rektorer bidrar till elevresultat	16
3.3	Huvudanalys	18
3.4	Experimentdesign – en kompletterande analys	21
3.5	En beskrivning av pedagogiskt ledarskap, betyg och nationella prov vid olika lärosäten.....	22
4	Diskussion	24
	Referenser	27
	Appendix.....	32

1 Inledning

I denna rapport presenterar vi resultaten från en effektutvärdering av Skolverkets rektorsprogram. Skolverket har bedrivit Rektorsprogrammet sedan 2009, och i Skollagen blev utbildningen ett behörighetskrav för både verksamma och blivande rektorer från och med 2010. Samma år förtydligades att rektorer ska leda och samordna det pedagogiska arbetet på skolorna, och särskilt verka för att utveckla utbildningen (Skollagen 2010:800). Rektorsutbildningen ska ge deltagarna de kunskaper de behöver för att leda, organisera och förbättra verksamheten (Skolverket 2020). Flera undersökningar har lyft fram vikten av rektors ledarskap för skolors kvalitet och utveckling, bland annat en forskningsöversikt från Vetenskapsrådet (2011) och Skolinspektionens rapport ”Rektors ledarskap med ansvar för den pedagogiska verksamheten” (Skolinspektionen 2012).

Internationellt finns omfattande forskning om rektorers betydelse för elevers skolresultat. Till exempel visar Leithwood m.fl. (2004) i sin forsknings-sammanställning att rektorers skolledarskap kan förklara 3–5 procent av skillnaderna i elevers resultat mellan skolor. Detta utgör ungefär en fjärdedel av den totala variation som kan hänföras till skolskillnader. Forskare har undersökt pedagogiskt ledarskap i skolor för att identifiera och beskriva olika ledarskapspraktiker och hur dessa påverkar den pedagogiska personalens arbetsmiljö och prestationer, särskilt när det gäller elevers resultat. Det finns också forskning om olika typer av rektorsutbildningar, både för blivande och verksamma rektorer, framför allt i USA. Däremot saknas i Sverige och många andra länder forskning som undersöker om rektorers vidareutbildning faktiskt påverkar elevers skolresultat – som är skolans viktigaste mål.

I rapporten undersöker vi effekter av att delta i Rektorsprogrammet. Vi analyserar data från 3 783 deltagare under åren 2009–2021. Resultaten tyder på att deltagande i Rektorsprogrammet inte leder till mätbara förbättringar av elevernas studieresultat eller lärarnas bedömning av rektorers pedagogiska ledarskap. Även om den tidigare forskningslitteraturen är relativt begränsad och resultaten inte är helt entydiga, pekar forskningen på att utbildningsprogram för rektorer sällan har några tydliga effekter på elevers prestationer i skolan. Våra resultat ligger i linje med detta. Studien tyder därmed på att rektorsutbildning i sig inte räcker för att förbättra skolresultaten. Däremot kan Rektorsprogrammet ha ett egenvärde genom att introducera rektorer till yrket och de krav, regler och lagar som följer med yrkesrollen. Programmet kan också påverka andra utfall än dem vi har studerat i rapporten.

1.1 Rektorers pedagogiska och administrativa ledarskap

Rektorsprogrammet syftar till att förbättra rektorers *pedagogiska ledarskap*, vilket i sig är ett forskningsfält. Hallinger (2003) definierar pedagogiskt ledarskap som skolledarens förmåga att formulera skolans mål, leda den pedagogiska verksamheten och utveckla skolans lärandeklimat. Forskning om pedagogiskt ledarskap fokuserar ofta på mätbara förbättringar av elevernas lärande (Dutta & Sahney 2022). Skolinspektionen beskriver pedagogiskt ledarskap som rektors förmåga att ”tolka uppdraget, omsätta det i undervisning, leda och styra lärprocesser, samt skapa förståelse hos medarbetare för samband mellan insats och resultat” (Skolinspektionen, 2012, s. 6). Skolinspektionen följer årligen upp rektorers pedagogiska ledarskap genom skolenkäten, som lärare, elever och föräldrar besvarar (Skolinspektionen 2020).

Det finns omfattande internationell forskning om rektorers ledarskap och hur detta relaterar till elevers skolresultat. Som noterats i inledningen visar Leithwood m.fl. (2004) att rektorers skolledarskap kan förklara mellan tre och fem procent av skillnaderna i elevers resultat mellan skolor, vilket utgör ungefär en fjärdedel av allt som skolfaktorer tillsammans förklarar.

Flera forskare har sammanfattat meta-analyser av relationen mellan rektorers ledarskap och elevresultat i så kallade ”meta-metaanalyser”¹. Wu och Shen (2022) analyserade 18 tidigare metastudier och fann positiva samband mellan elevresultat och ledarskap, både när de undersökte etablerade ledarskapsbegrepp som ”pedagogiskt ledarskap” och när de studerade konkreta rektorspraktiker. En annan meta-metastudie baserad på 12 tidigare meta-analyser påvisade liknande resultat, men menade att sambanden generellt är starkare när forskarna använder ledarskapsmodeller för att kontextualisera och mäta rektorers ledarskap, jämfört med studier som söker identifiera specifika ledarskapspraktiker (Tan m.fl. 2022). Tidigare forskning visar också att sambanden är indirekta genom rektorers pedagogiska ledarskap av lärare, eftersom rektorer sällan har direkt kontakt med eleverna (Meyer m.fl. 2022).

I en svensk studie som undersöker hur rektorer påverkar elevresultat – snarare än att bara studera samvariation – fann Böhlmark m.fl. (2016) att en standardavvikelses skillnad i rektorers skicklighet motsvarade 0,05–0,1 standardavvikelsers skillnad i elevers skolresultat. Författarna beskriver den positiva effekten som substantiell och jämförbar i storleksordning med både lärareffekter och en klasstorleksminskning på två till fyra elever per klass. Det är dock svårt att identifiera exakt vad som kännetecknar framgångsrikt ledarskap bland

¹ ”Meta-metaanalyser” sammanställer resultat från tidigare metaanalyser. Eftersom dessa i sin tur till övervägande del bygger på tvärsnittsstudier rör det sig i huvudsak om korrelationssamband, inte om orsakssamband (dvs. inte om ”kausala effekter”).

rektorer, men författarna föreslår att rektorers personliga kompetens kan vara en möjlig förklaring.

Att studera effekter av rektorers ledarskap är utmanande, delvis på grund av generella svårigheter med att mäta ledarskapseffekter (Grissom & Loeb 2009). Till skillnad från lärare som har direktkontakt med elever och ansvarar för deras lärande, påverkar rektorer elevers lärande och skolresultat indirekt. Rektorer ansvarar för arbetsmiljön för lärare och elever genom beslut om undervisningens innehåll och upplägg utifrån läroplanens mål, organisation, bemanning, budget, lärares professionella utveckling, disciplin, närvaro samt skolpersonalens mål och handledning (Grissom & Loeb 2009; Leithwood m.fl. 2020). Det är därmed svårt att identifiera och mäta specifika mekanismer genom vilka rektorers pedagogiska ledarskap påverkar elevers lärande och skolresultat. Dessutom kan rektorers påverkan vara svår att särskilja från andra väletablerade faktorer som lärarnas insatser, föräldrarnas utbildningsnivå, närområdets karaktär samt grupp-effekter bland eleverna. En metodologisk utmaning är också att rektorer inte fördelas slumpmässigt mellan skolor, utan ofta aktivt väljer att söka sig till – eller undvika – skolor som är mer eller mindre välfungerande och elevgrupper som är mer eller mindre högpresterande.

Rektorers pedagogiska ledarskap påverkas även av deras position i ”styrkedjan” mellan huvudmannen och lärarna på en skola. Styrkedjan omfattar alla aktörer som kopplar samman de nationella målen för skolan med eleverna, det vill säga staten, huvudmannen, huvudmannens förvaltning, rektorer och lärare. I fristående skolor är styrkedjan något kortare och består av staten, huvudmannen, skolområdeschef/verksamhetschef, rektorer och lärare. En statlig utredning från 2015 drar slutsatsen att många av de utmaningar som rektorer möter kan bero på otydlig kommunikation i styrkedjan, vilket försvårar arbetet med att upprätta och genomföra strategier för att skolorna ska nå målen (SOU 2015:22). Utredningen betonar att den lokala styrkedjan – länken mellan huvudmannen och lärare via rektor – kan se olika ut, och att en tät och tydlig lokal styrkedja är en viktig förutsättning för att rektorer ska kunna utöva pedagogiskt ledarskap. I denna rapport undersöker vi om rektorers pedagogiska ledarskap och elevers skolresultat påverkas av rektorers deltagande i Rektorsprogrammet, som finns specificerat i Skollagen som en del av styrkedjan.

En annan distinktion som är viktig i rapporten rör skillnaden mellan rektorers generella betydelse och betydelsen av att *utbilda* rektorer. Rektorer kan vara betydelsefulla för elevresultat utan att en specifik utbildningsinsats nödvändigtvis förändrar dem på sätt som syns i elevresultat. Rapportens huvudfråga gäller det senare, och den litteratur som mer direkt motsvarar vår studie presenteras i följande avsnitt.

1.2 Svenska och internationella studier om rektorsprogram

Sedan 2009 har mer än 12 000 rektorer examinerats från Rektorsprogrammet. Trots detta har ingen systematisk utvärdering av programmets effekter på skolutfall genomförts. De svenska studier som finns fokuserar huvudsakligen på rektorers upplevelser av utbildningen och deras självskattade lärande genom kvalitativa metoder. Skott och Törnsén (2018) undersökte i en kvalitativ IFAU-rapport hur tolv programdeltagare uppfattade sin ledarskapsidentitet och hur de skapade mening i relation till sitt lärande under utbildningen. Rapporten framhöll erfarenhetsutbytet betydelse för rektorernas lärande. Även Jerdborg (2023) bygger på intervjuer och observationer av programdeltagare, med särskilt fokus på orsakerna till avhopp. Resultaten visar att de rektorer som avbröt utbildningen ofta upplevde att kopplingen mellan deras egen rektorspraktik och innehållet i programmet var svag. Aas och Paulsen (2019) undersökte hur programdeltagare balanserar mellan autonomi och struktur. De fann att denna balansgång underlättades när rektorerna fick möjlighet att tillsammans reflektera över utbildningen.

Befattningsutbildningar eller förberedelseutbildningar för rektorer är inte unika för Sverige. Rektorsprogram har blivit allt vanligare globalt. Bara i USA har antalet rektorsutbildningar fördubblats mellan år 2000 och 2014 (Perrone & Tucker 2019). Trots denna expansion kvarstår frågetecken kring huruvida dessa program leder till förbättrade skolresultat för elever (Grissom m.fl. 2019). Även om rektorsprogrammen är utformade för att stärka skolledarskap och därigenom förbättra olika typer av lärandeutfall, bygger en stor del av den befintliga forskningen på indirekta mått. Detta lämnar centrala frågor om programmens direkta och långsiktiga effekter obesvarade.

Tidiga studier av rektorsprogram fokuserade främst på rektorers ledarskapspraktiker, som antas ha en positiv inverkan på elevresultat (Donmoyer m.fl. 2012). Weinstein m.fl. (2018) analyserade data från ett representativt urval av skolor i sju latinamerikanska länder för att undersöka hur erfarenhet från rektorsutbildning samvarierar med rektorers ledarskapspraktiker. Resultaten påvisade endast marginella skillnader mellan rektorer som hade genomgått utbildningen och de som inte hade gjort det. En annan forskningsinriktning har undersökt lärarrelaterade utfall såsom förbättrat lärarsamarbete och ökad arbetstillfredsställelse. Orphanos och Orr (2014) jämförde 175 lärare vars rektorer hade genomgått ett ”innovativt ledarskapsprogram” med 589 lärare vars rektorer hade en traditionell förberedande utbildning. Genom strukturella ekvationsmodeller fann de att deltagande i utbildningen var förknippat med ett direkt positivt samband med rektorernas ledarskapspraktiker och ett indirekt samband med lärarsamarbete och arbetstillfredsställelse. Resultaten visade också att lärare som uppfattade sin rektors ledarskap mer positivt rapporterade högre arbetstillfreds-

ställelse och upplevde mer kollegialt samarbete. Dessa typer av korrelationsstudier har dock inte undersökt rektorsprogramms direkta effekter på skolresultat.

Nyare forskning som utvärderar rektorsprogramms effekter genom att använda elevresultat som utfallsmått har i kvasiexperimentella analyser gett blandade resultat. Motsvarande studier finns ännu inte i Sverige – de flesta effektutvärderingarna är genomförda i USA. Corcoran m.fl. (2012) studerade New Yorks ”Skolledarskapsakademi” – ett 14-månaders skolledarprogram som förbereder blivande rektorer för att leda lågpresterande innerstadsskolor. Forskarna utvärderade treårseffekter bland 109 rektorer från programmet och jämförde dem med 331 rektorer i en kontrollgrupp genom en så kallad ”skillnad-i-skillnader”-modell. De fann att rektorerna från Skolledarskapsakademin, som oftare ledde mer utmanande skolor, bytte jobb något oftare än kontrollgruppsrektorerna under sina första tre år. Forskarna upptäckte även att bland de rektorer som stannade kvar på skolorna närmade sig skolresultaten i engelska de mer högpresterande jämförelseskolorna, men inte skolresultaten i matematik. Steele m.fl. (2021) genomförde en annan studie med liknande design. De undersökte effekten av ett kortare (5-veckors) rektorsprogram för blivande rektorer i en mindre grupp amerikanska innerstadsskolor med mestadels icke-vita elever från låginkomsthushåll. Resultaten visade att de skolor där rektorer genomgått rektorsutbildningen hade förbättrade elevresultat, men samtidigt fler elevavstängningar. En större studie av Doss m.fl. (2022) undersökte sambandet mellan deltagande i ett ettårigt program för blivande rektorer i mestadels segregerade innerstadsskolor. De fann statistiskt signifikanta samband med förbättringar i elevers skolresultat, närvaro och sannolikheten för rektorerna att stanna längre på skolan.

Rektorsutbildning i USA finns inte bara för skolor med stora utmaningar utan även som mer generella program i olika former. Corcoran (2017) undersökte det nationella *School Leadership's Executive Development Program* som har vissa likheter med det svenska nationella Rektorsprogrammet. Båda är kohortbaserade yrkesprofessionella program för lärare som redan tillträtt som rektorer. Det amerikanska programmet omfattar ett år av högskolekurser i internatform (cirka två träffar i månaden) parallellt med individuell onlineutbildning och gemensamt arbete inriktat på praktiska utmaningar i rektorsrollen. I studien jämfördes 124 deltagande rektorer som ledde grundskolor i Wisconsin med 62 rektorer som verkade vid jämförbara skolor men som inte genomgått utbildningen. Forskarna kunde dock inte påvisa några mätbara förbättringar i elevers skolresultat för skolor där rektorer deltagit i programmet.

Det finns även randomiserade fältstudier av ledarskapsprogram för rektorer i USA. Master m.fl. (2022) jämförde i en sådan studie effekten av ett ettårigt

deltidsprogram för blivande rektorer med ett program som även inkluderade efterföljande handledning². Deras analys visade att rektorsprogrammet i sig förbättrade rektorerers ledarskapspraktiker, men fann ingen direkt inverkan på elevers skolresultat. När programmet kombinerades med handledning kunde forskarna observera små förbättringar i elevers läs- och skrivkunskaper, medan ingen effekt kunde påvisas i matematik.

Studier om rektorerers utbildning och kvalifikationer – dock inte specifika rektorsprogram – har även genomförts i andra länder. En korrelationsstudie med över 3 000 elever från 182 skolor i Burkina Faso visade att rektorerers utbildning, erfarenhet och pedagogiska fortbildning samvarierade positivt med elevers skriv- och matematikresultat (Sanfo 2020). Gümüş m.fl. (2024) rapporterar positiva samband i en tvärsnittsstudie av sambandet mellan rektorerers kvalifikationer och elevers prestationer i matematik, läsning och naturkunskap med data från TALIS 2018 och PISA 2018. Studien omfattade elever från 497 skolor i Australien, Colombia, Danmark, Georgien, Malta, Tjeckien och Turkiet. Rektorsutbildning som genomförs efter att rektorer tillträtt sin tjänst – vilket ofta är fallet i Sverige – hade enligt studien positiva samband med elevresultat i samtliga ämnen.

En betydande begränsning i den befintliga forskningen är beroendet av observationsdata. Befintliga kausala studier (främst från USA) har ofta baserats på ett begränsat antal skolor i specifika skolkontexter, såsom starkt segregerade skolor, vilket minskar studiernas generaliserbarhet. Den studie som i urval och ansats ligger närmast vår är Corcorans (2017) studie av 186 rektorer i Wisconsin, som inte fann några märkbara förbättringar i elevers skolresultat för skolor där rektorer deltagit i USA:s nationella utbildningsprogram för rektorer.

Forskare som Gümüş m.fl. (2024) och Doss m.fl. (2022) betonar behovet av kausala ansatser för att fullt ut förstå hur rektorsprogram påverkar elevresultat i rektorernas skolor. För att hantera dessa begränsningar använder vi en longitudinell, kvasiexperimentell forskningsdesign med ett stort antal svenska skolor. Även om vår analys också har vissa osäkerheter, är möjligheterna att fånga effekten av deltagande i det svenska Rektorsprogrammet på pedagogiskt ledarskap (bedömt av lärare på den aktuella skolan) och elevresultat (avgångsbetyg och nationella provresultat) bättre än vad som vanligtvis är fallet i den existerande litteraturen.

² Master m.fl. citerar även flera randomiserade fältstudier som genomförts på samma ledarskapsprogram samt andra amerikanska program. De flesta påvisar nolleffekter på elevresultat.

1.3 Det svenska rektorsprogrammet

Rektorsprogrammet i Sverige är en statlig befattningsutbildning på avancerad akademisk nivå som Skolverket ansvarar för. Utbildningen regleras i Förordning (2011:183). Programmet startade 2009 och blev obligatoriskt för nya rektorer från 2010. Genom en lagändring 2018 (Proposition 2017/18:182) likställdes förskolans ledarskap med skolans och titeln ”förskolechef” ersattes med ”rektor” i skollagen. Sedan 2019 är Rektorsprogrammet obligatoriskt även för rektorer i förskolan. Även andra skolledare med motsvarande ledningsansvar kan delta, trots att de har andra titlar än rektor. Detta inkluderar exempelvis biträdande rektorer, förutsatt att de uppfyller behörighetskraven. Behörighetskraven innebär att minst 50 procent av tjänsten ska innefatta ledningsansvar, att den sökande har personalansvar samt mandat att leda och samordna det pedagogiska arbetet på skolan. Biträdande rektorer och andra som uppfyller dessa kriterier kan erbjudas en plats på Rektorsprogrammet i mån av plats.

En nyutträd rektor måste påbörja utbildningen senast ett år efter tillträdet, medan rektorer i förskolan har upp till två år på sig att inleda utbildningen. Rektorsprogrammet ska vara avslutat senast fyra år efter påbörjad utbildning för rektorer i skolan, och inom fem år för rektorer i förskolan. Huvudmannen ansvarar för att anmäla rektorer till programmet och säkerställa att de har förutsättningar att genomföra utbildningen på deltid inom ramen för sitt arbete. Huvudmannen ansvarar också för praktiska frågor som omkostnader och val av lärosäte.

I Sverige finns Rektorsprogrammet idag vid sju olika universitet: Göteborgs universitet, Karlstads universitet, Linnéuniversitetet, Mittuniversitetet, Stockholms universitet, Umeå universitet och Uppsala universitet. Umeå universitet och Karlstads universitet erbjuder programmet på två olika orter. Utbildningen sträcker sig över tre år, och deltagarna förväntas använda 20 procent av sin arbetstid till programmet samtidigt som de arbetar på sina skolor under utbildningstiden.

Enligt Förordning (2011:183, §2) syftar Rektorsprogrammet till att ge rektorer och annan personal med motsvarande ledningsfunktion kunskap om de krav som anges i gällande föreskrifter samt att utveckla deras ledarskapsförmåga så att verksamhetens kvalitet kan säkerställas. Målet med utbildningen är att deltagarna ska få sådana kunskaper att de kan:

- Ansvara för att elever och barn får en likvärdig och rättssäker utbildning
- Skapa förutsättningar för måluppfyllelse på både individ- och verksamhetsnivå
- Ansvara för att verksamheten som helhet utvecklas och förbättras

Utbildningen omfattar totalt 30 högskolepoäng, fördelade på tre kurser om vardera 10 högskolepoäng: *Skoljuridik och myndighetsutövning*, *Styrning, organisering och kvalitet* samt *Skolledarskap – pedagogisk ledning* (Förordning 2019:562).

Skolverket beskriver det mer preciserade innehållet i de tre kurserna inom Rektorsprogrammet i sitt måldokument (Skolverket 2020). Kursen *Skoljuridik och myndighetsutövning* syftar till att ge rektorer en djupare förståelse för de regleringar som styr skolväsendet. Kursen *Styrning, organisering och kvalitet* syftar till att stärka rektorers roll som chef med personalansvar och deras förmåga att styra och organisera utbildning samt bedöma kvalitet i pedagogisk verksamhet. Den tredje kursen *Skolledarskap – pedagogisk ledning* fokuserar på att utveckla och systematisera rektorers pedagogiska ledarskap, med målet att skapa goda förutsättningar för elevers lärande. Här betonas rektorers förmåga att främja undervisningsskicklighet och skapa en trygg arbetsmiljö för både elever och personal. Kursen lägger även vikt vid kritisk reflektion över det egna ledarskapet, de strukturella förutsättningarna för pedagogisk ledning samt komplexiteten i rektorsuppdraget inom en målstyrd organisation. Kursbeskrivningen betonar också vikten av att rektorer förstår, tar ansvar för och aktivt samverkar i styrkedjan.

Skolhuvudmannen anmäler deltagare till Rektorsprogrammet via Skolverket, som därefter förmedlar förteckningar över antagna deltagare till de lärosäten som anordnar utbildningen. Antagning och kursstart sker varje termin. Lärosätena genomför undervisningen främst genom internat, ofta på kursgårdar med övernattning. Enligt Skolverkets nuvarande måldokument är utbildningen erfarenhetsbaserad och processinriktad. Deltagarnas professionella erfarenheter utgör en central utgångspunkt för studierna och fördjupas genom litteraturstudier och föreläsningar. Utbildningen omfattar även uppsatsskrivande, gruppövningar och uppgifter som knyts till deltagarnas egna skolverksamheter och relateras till vetenskaplig teori.

Programmets upplägg kan variera något beroende på vilket lärosäte som genomför utbildningen. Vid Stockholms universitet ges Rektorsprogrammet som en kurs med tre delkurser, där ansvaret är fördelat mellan fem olika institutioner. Juridiska institutionen, tillsammans med Barn- och ungdomsvetenskapliga institutionen, administrerar *Skoljuridik och myndighetsutövning*. Företagsekonomiska institutionen, tillsammans med Statsvetenskapliga institutionen, ansvarar för *Styrning, organisering och kvalitet*. Institutionen för pedagogik och didaktik har ansvar för *Skolledarskap – pedagogisk ledning*. Undervisningen består av tolv internat förlagda på en kursgård, kombinerade med seminariedagar på plats eller online. Vid Göteborgs universitet genomförs

däremot utbildningen helt genom internat, med två tillfällen per termin om tre dagar vardera. Umeå universitet har valt att skifta ordningen på kurserna så att programmet avslutas med Styrning, organisering och kvalitet som normalt ges under år två. Både Karlstads universitet och Mittuniversitetet integrerar kurserna så att det finns inslag av exempelvis ledarskap i alla tre kurserna.

Ofta delar lärosätena in utbildningsgrupperna i mindre dialog-/basgrupper för övningar och erfarenhetsutbyte. Gruppindelningen ändras ofta inför varje ny kurs för att bredda rektorernas kontaktnät och möjliggöra fler perspektivutbyten. Även om litteraturlistorna varierar mellan lärosätena finns vissa återkommande titlar. Några av de mest förekommande är *Skoljuridik* av Boström (2023), *Rektor i fokus: kunskap, värden och verktyg* av Blossing m.fl. (2020) samt *Att leda mot skolans mål: vägval och möjligheter* av Johansson och Svedberg (2016).

2 Metod och data

2.1 Data

Vår undersökning utgår från en förteckning över alla som deltagit i Rektorsprogrammet från starten år 2009 till och med år 2021. Förutom denna deltagarförteckning har vi också haft tillgång till anmälningslistorna till Rektorsprogrammet. Därmed har vi kunnat jämföra sökande med antagna och på så vis identifiera personer som inte direkt blivit antagna till Rektorsprogrammet utan i stället stått på en väntelista innan de påbörjat programmet. Listorna över sökande och antagna kommer från Skolverket och har via Statistiska centralbyrån matchats med Lärarregistret, så att vi har kunnat följa deltagare i Rektorsprogrammet över tid.

Data från Lärarregistret låter oss se vilka av deltagarna i Rektorsprogrammet som varit aktiva som rektorer eller biträdande rektorer, och på vilka skolor, under perioden 2009–2021. Dock ändrades Lärarregistret 2013 så att man först då började registrera vilka som är rektorer respektive biträdande rektorer. Dessförinnan registrerades båda yrkesrollerna i kategorin ”Rektor och övriga skolledare (arbetsledning)”. Från och med 2013 finns benämningarna ”Rektor” och ”Övrig arbetsledning” i Lärarregistret. Vi kallar övrig arbetsledning för biträdande rektorer, eftersom det är den grupp som utöver rektorer blir antagen till Rektorsprogrammet. Lärarregistret uppdateras en gång om året, i oktober, vilket gör att eventuella skol- eller rollbyten kan ske utan att det direkt syns i registret. Vi har använt data så att rektorer och biträdande rektorer som

registrerats på en skola kopplas till skolans utfall samma läsår³. Det vill säga, om en rektor finns i Lärarregistret på en viss skola höstterminen 2015, så kommer rektorn i vår data att länkas till denna skolas utfall för läsåret 2015/2016, till exempel slutbetygen i årskurs nio för vårterminen 2016.

Vi undersöker tre olika utfall i relation till rektorers medverkan i Rektorsprogrammet. Dessa är elevernas slutbetyg och nationella provresultat i årskurs nio, samt lärares bedömningar av rektorers pedagogiska ledarskap. Vi har hämtat uppgifter om betyg och nationella prov från Skolverkets offentligt tillgängliga statistikdatabas. De tillgängliga uppgifterna är aggregerade till skolnivå, det vill säga betygen eller provpoängen som redovisas är medelvärden per skola och år. Betygen per skola är medelvärden av alla tillgängliga betyg för en skolas elever i alla ämnen. Nationella prov presenteras ämnesvis i den offentliga databasen. Våra data för nationella provresultat per skola och år är medelvärden av provpoängen i matematik, svenska och engelska. Uppgifter som bygger på mindre än tio elevers betyg eller provpoäng maskeras i den offentliga databasen och är därför inte tillgängliga. I statistiken från Skolverket har betygsskalan inte översatts, så uppgifterna fram till och med 2012 är inte jämförbara med betygsskalan för nationella prov som finns från och med 2013. Data före 2013 maskeras dessutom systematiskt, vilket gör att det inte går att räkna ut betygssnitt per skola. Därför använder vi data om nationella prov från och med år 2013. På grund av covid-19 genomfördes inte de nationella proven åren 2020 och 2021. År 2018 finns inga resultat för matematikproven registrerade i den offentliga statistiken eftersom provuppgifterna spridits innan provtillfället, vilket antogs ha kunnat påverka resultatens tillförlitlighet.

Det tredje utfallet som vi undersöker är rektorers pedagogiska ledarskap. Data om detta har vi fått från Skolinspektionen, som sedan 2010 administrerar en årlig enkät till personal vid alla Sveriges skolor som en del av Skolenkäten. Att fylla i enkäten är frivilligt, men skolor är skyldiga att ge lärare möjlighet att delta i undersökningen. Rektorers pedagogiska ledarskap mäts i lärarenkäten genom en skala som omfattar tre påståenden: Lärarna ombeds att skatta i vilken utsträckning rektorn (1) har god kunskap om det dagliga arbetet i skolan, (2) arbetar för att främja elevernas kunskapsutveckling, och (3) tar ansvar för det pedagogiska arbetet på skolan. Vi har fått tillgång till rådata om pedagogiskt ledarskap från Skolinspektionen för åren 2010–2021 och har genomfört en

³ Denna matchningsregel innebär att rektorer kan tillskrivas en skolas utfall även om de tillträtt eller lämnat skolan vid annan tidpunkt under läsåret. Sådant mätfel tenderar att dra effektskattningar mot noll snarare än att skapa systematisk snedvridning, förutsatt att felet inte är korrelerat med programdeltagande. Eftersom huvudresultaten är nära noll innebär detta att eventuella sanna effekter möjligen är något större än vad våra skattningar antyder.

valideringsstudie eftersom skalan inte tidigare använts i forskningssyfte och därför inte fått sitt innehåll prövat. Valideringsstudien beskrivs i metodbilagan.

Förutom våra utfallsvariabler inkluderar studien kontrollvariabler som delvis finns tillgängliga i Skolverkets offentliga statistik. För data om föräldrarnas utbildningsnivå har vi använt föräldrars genomsnittliga utbildningsbakgrund per skola och år från Skolverkets databas SALSA, där vi också hämtat data om andelen nyinvandrade elever per skola och år från år 2013. Ur Skolverkets Personalstatistik har vi hämtat data om antalet rektorstjänster per skola och om antalet lärare per skola för att kontrollera för skolornas storlek. Andra kontrollvariabler som vi hämtat från Rektorsprogrammets deltagarlistor eller från Lärarregistret är: kön, om personen är rektor eller biträdande rektor, om skolans huvudman är kommunal eller privat samt hur länge man tjänstgjort som skolledare innan Rektorsprogrammet. Vi har också konstruerat kontrollvariabler för när rektorer byter skola och hur lång tjänstetid man har på en skola under observationsperioden.

2.2 Urval och population

Under åren 2009–2021 har 13 718 personer påbörjat Rektorsprogrammet. Av dem har 80 procent en yrkesbakgrund som lärare. Mer än två tredjedelar (68 procent) är kvinnor. Enligt ansökningsformulären till Rektorsprogrammet har många som påbörjar programmet tidigare erfarenhet som skolledare: 54 procent har haft mellan tre och fem års erfarenhet som skolledare, medan cirka 20 procent haft mindre än ett års erfarenhet eller ingen alls. Över tid har deltagargruppen förändrats något. Under de första åren deltog främst rektorer i utbildningen, eftersom den var obligatorisk för dem. Antalet biträdande rektorer har ökat över tid. En annan grupp som successivt ökat i programmet är rektorer i förskolan (som tidigare kallades förskolechefer). I vår undersökning studerar vi dock enbart deltagare i Rektorsprogrammet som arbetat som rektorer eller biträdande rektorer under perioden 2009–2021 på skolor med årskurs nio, vilket är 3 783 personer. Av dessa tjänstgjorde nästan hälften (1 746 personer) som rektorer redan innan de påbörjade programmet. De 12 kohorterna av programdeltagare som vi har analyserat innefattar ungefär 140–400 deltagare årligen (Tabell 1). Den största kohorten (år 2009/2010) utgör 11 procent av de 3 783 deltagarna; den minsta (2011/2012) 4 procent av deltagarna.

Tabell 1 Kohortdistribution (2010–2021)

Kohort	Antal deltagare (samtliga)	Antal rektorer i urvalet (verksamma på skolor med Åk 9)
2009/2010	1 387	411
2010/2011	1 093	331
2011/2012	1 109	139
2012/2013	988	315
2013/2014	964	263
2014/2015	1 030	254
2015/2016	1 033	305
2016/2017	1 100	376
2017/2018	1 263	403
2018/2019	1 226	347
2019/2020	1 312	330
2020/2021	1 313	309
Totalt	13 718	3 783

2.3 Statistisk analys

Vår metod för att undersöka vilka effekter Rektorsprogrammet har på skolors betyg, nationella prov och lärares skattningar av rektors pedagogiska ledarskap går ut på att studera hur utfallen utvecklas per år, innan och efter det att rektorer påbörjar Rektorsprogrammet, vilket ibland kallas för en *eventstudie*. Rektorer jämförs med sig själva innan och efter programstart, och med kontrollgrupper. För att skapa kontrollgrupper och samtidigt dra nytta av att deltagare påbörjar programmet vid olika tillfällen har analysen sin förebild i Callaway och Sant’Annas (2021) dynamiska skillnad-i-skillnader-analys.⁴

Analyserna av Rektorsprogrammets effekter görs med flernivåanalys, som är vanlig i psykologi och ledarskapsforskning. Flernivåanalysen säkerställer att enheten som undersöks är rektorer. En utförligare beskrivning av analysen och vår statistiska modell finns i appendix. Flera alternativa metodval är möjliga. En renodlad tillämpning av Callaway och Sant’Annas (2021) modell hade varit ett naturligt alternativ, men skulle ha gett skattningar på kohortnivå snarare än rektornivå. Det är inte helt ändamålsenligt när rektorerers olikheter är centrala i studien. Vi har provat variationer av specifikationer i modellerna som redovisas,

⁴ En kausal tolkning av programeffekten (dvs. att det finns ett orsakssamband) förutsätter antaganden om parallella trender, ingen påverkan av deltagare i behandlingsgruppen på kontrollgruppen (s.k. ”spillover”), samt att oobserverade rektors- och skolspecifika faktorer inte är korrelerade med tidsvarierande förklaringsvariabler. Identifikationen utnyttjar två egenskaper i data: att programmet rullas ut stegvis över tid med olika behandlingskohorter och att cirka en tredjedel av rektorena byter skola under observationsperioden. Det senare gör det möjligt att separera rektors- och skoleffekter. Antagandena diskuteras närmare i appendix.

utan att slutsatserna ändras. En mer ingående redogörelse för dessa val och tillhörande tester återfinns i metodbilagan.

Den ”behandling” som vi studerar är att rektorer påbörjar Rektorsprogrammet, inte att de fullföljer eller examineras från det. År 0 i analysen är alltså året då deltagaren påbörjar programmet. Det är värt att komma ihåg att behandlingen har förhållandevis begränsad omfattning och är distribuerad över en längre tid. Rektorsprogrammet omfattar 30 högskolepoäng fördelade över tre år, vilket motsvarar 20 procent av deltagarnas arbetstid. Deltagarna arbetar parallellt som rektorer eller biträdande rektorer på sina skolor, och studietakten kan variera mellan deltagare och lärosäten. Vidare är kedjan från behandlingen till de utfall vi studerar relativt lång: programmet ska påverka rektors ledarskapspraktik, som i sin tur ska påverka lärares arbete, som i sin tur ska påverka elevers resultat. Givet behandlingens omfattning och denna indirekta verkningsmekanism är det rimligt att förvänta sig att eventuella effekter på betyg och nationella prov är små, medan effekter på rektors pedagogiska ledarskap (som ligger närmare behandlingen i orsakskedjan) kan tänkas stå i mer direkt relation till behandlingen.

Vi följer rektorerna i tre år innan de börjar på Rektorsprogrammet och i fem år efter programstart. Förutom att vi studerar effekten av Rektorsprogrammet över tid undersöker vi om effekter skiljer sig åt beroende på ledarposition. Rektorer och biträdande rektorer har olika roller och ansvarsnivåer i skolorganisationen, där rektorer har det övergripande ledningsansvaret medan biträdande rektorer ofta har mer operativa uppgifter. Därför skulle effekterna av Rektorsprogrammet kunna skilja sig åt mellan rektorer och biträdande rektorer.

3 Resultat

3.1 Beskrivande statistik

Data som används för att undersöka effekterna av Rektorsprogrammet består av 18 563 årsvisa observationer från 3 783 deltagare i programmet. Vi har sammanställt uppgifter om respektive deltagare över tid, både före och efter programstart. Beskrivande statistik för urvalet som vi analyserar visas i Tabell A1 i appendix. Av deltagarna i urvalet är 68 procent kvinnor och 32 procent män, och 83 procent har gått en lärarutbildning. Vid ansökningstillfället angav 49 procent att de omfattas av obligatoriet att gå programmet, och 37 procent att de ansvarar för det systematiska kvalitetsarbetet på sin skola. En knapp femtedel (20 procent) hade vid ansökningstillfället mellan noll och ett års rektorserfarenhet, drygt hälften (58 procent) hade 3–5 års erfarenhet, medan endast 10 procent hade mer

än fem års erfarenhet. Rektorstjänster varierar något i omfattning, men nästan hälften av deltagarna har en heltidstjänst som rektor. Deltagarna är verksamma vid 2 147 skolor, varav 22 procent är fristående. Rektorerna stannar på skolan i genomsnitt cirka 2,3 år, och drygt en tredjedel bytte skola under observationsperioden.

I Tabell 2 redovisar vi beskrivande statistik för våra utfallsvariabler per skola över tid. Vi redovisar också föräldrars genomsnittliga utbildningsnivå per skola över tid.

Tabell 2 Medelvärde för genomsnittsbetyg och nationella prov (årskurs 9), pedagogiskt ledarskap, samt föräldrarnas utbildningsnivå

	Antal	Medelvärde	Standard- avvikelse	Min	Max
Betyg	2 025	220,8	27,0	20,0	299,2
Nationella prov	1 832	13,3	1,5	2,5	17,9
Pedagogiskt ledarskap	1 573	1,8	0,4	1,0	4,0
Föräldrars utbildning	1 966	2,3	0,2	1,3	2,9

Not: *Betyg* avser genomsnittligt meritvärde i årskurs 9, det vill säga summan av elevens 16 högsta slutbetyg där varje betygssteg ger poäng enligt A = 20, B = 17,5, C = 15, D = 12,5, E = 10 och F = 0. Maximalt meritvärde är 320 poäng (samma maximivärde gäller även under det tidigare betygssystemet G/VG/MVG, där G = 10, VG = 15 och MVG = 20). Genomsnittet beräknas på skolnivå och hämtas från Skolverkets statistik. *Nationella prov* avser skolans genomsnittliga provbetygs-poäng i matematik, svenska och engelska i årskurs 9, omvandlad till samma A–F-poängskala (A = 20, ..., F = 0), där 20 poäng motsvarar genomsnittsbetyget A. *Pedagogiskt ledarskap* är ett skolnivåindex baserat på enkätfrågor besvarade av lärare. Ett högre värde innebär att ledarskapet uppfattas som bättre. *Föräldrars utbildning* är hämtad från Skolverkets SALSA-databas och anger genomsnittlig högsta utbildningsnivå för elevernas föräldrar vid respektive skola. Skalan har tre nivåer: 1 = genomgången folk- eller grundskola, 2 = genomgången gymnasial utbildning, 3 = eftergymnasial utbildning. Beräkningen baseras på genomsnittet av föräldrarnas högsta utbildningsnivå (vid uppgift om båda föräldrarna), och uppgifterna hämtas från SCB:s utbildningsregister. Skolvärdet utgör genomsnittet av elevernas individuella värden.

Skolorna i det studerade urvalet har i genomsnitt 27,5 lärartjänster med en standardavvikelse på 13,7 lärartjänster. Av de 3 783 personerna i urvalet är 2 247 personer någon gång registrerade i Lärarregistret som rektorer och 2 084 personer som biträdande rektorer. Innan 2013, då ingen åtskillnad mellan rollerna gjordes i Lärarregistret, fanns 1 286 personer i urvalet registrerade som ”rektor och övriga skolledare”. Det finns viss rörlighet mellan de båda rollerna, där 673 personer går från att vara biträdande rektorer till att vara rektorer, och 356 personer går från att vara rektorer till att vara del av ”övrig arbetsledning”.

3.2 Intraklasskorrelationer – en beskrivning av hur skolor och rektorer bidrar till elevresultat

I våra data representerar skolor och rektorer samma nivå, det vill säga det finns betygsgenomsnitt, genomsnittspoäng på nationella prov samt lärares skattningar

av rektorers pedagogiska ledarskap per skola och år som utgör utfall för både skolan och skolans rektor. Men eftersom rektorer byter jobb mellan skolor relativt ofta (Skolverket 2024) går det att urskilja i vilken mån skolors resultat över tid kan tillskrivas faktorer som handlar om skolor respektive rektorer, och om rektorers inflytande förändras med deltagande i Rektorsprogrammet.

För att få en bild av rektorers inflytande över skillnader mellan skolor i betyg, nationella prov och lärares skattningar av pedagogiskt ledarskap, i förhållande till andra skolfaktorer, har vi räknat fram så kallade intraklasskorrelationer för rektorer och biträdande rektorer tillsammans med intraklasskorrelationer för skolor. Analyserna har alltså två hierarkiska grupperingar av data på samma nivå – skolor och rektorer. Intraklasskorrelationerna baseras på skolors respektive rektorers medelvärden över tid. Resultaten kan tolkas i procent och visar hur mycket variation i utfallen som kan tillskrivas rektorer respektive skolor. Notera att analysen är beskrivande och att den inte ska tolkas som en analys av orsaks-samband.

Tabell 3 redovisar procenttal för rektorer (utanför parentes) och skolor (inom parentes). Vi har gjort analysen först för hela urvalet av programdeltagare och sedan för rektorer och biträdande rektorer var för sig, både före och efter påbörjat Rektorsprogram. Vi har bara använt data från och med 2013, då båda rollerna (rektor och biträdande rektor) finns med i data. Underlaget omfattar 3 436 personer.⁵

Tabell 3 visar att andra faktorer på skolorna förklarar mer av variationen i de tre utfallsvariablerna mellan olika skolor än vad rektorer gör. Till exempel förklarar rektorer och biträdande rektorer bara 1 procent av variationen i skolors betygsmedelvärden, medan 80 procent av variationen mellan skolor förklaras av andra skolfaktorer.

Ett annat resultat är att rektorer har större inflytande än biträdande rektorer vad gäller alla tre utfallen. Analysen visar vidare att rektorers inflytande över pedagogiskt ledarskap är det enda resultatet som konsekvent är större än skolors inflytande. Skillnaden mellan rektorer och biträdande rektorer i just detta utfall är stor, med 19 procent för rektorer och 6 procent för biträdande rektorer.

De tre utfallen skiljer sig åt på så vis att rektorer och biträdande rektorer ser ut att ha litet men ändå visst inflytande över betygsskillnader mellan skolor och

⁵ Av de 15 488 observationerna (rektorer/skolor per år) var 3 155 observationer från 1 726 personer före det att de börjat på Rektorsprogrammet och 12 293 var observationer från 3 094 deltagare efter det att de påbörjat Rektorsprogrammet. I urvalet fanns också 8 405 observationer från 2 217 personer då de varit rektorer, och 7 043 observationer från 2 058 personer när de varit biträdande rektorer. Av personerna i urvalet var 2 357 personer förknippade med bara en skola, medan 1 079 personer återfanns på mer än en skola. I genomsnitt var de som arbetat på mer än en skola under tidsperioden på 2,3 skolor med en standardavvikelse på 0,5 skolor.

mycket litet eller inget inflytande över resultaten i nationella prov. Rektorer och biträdande rektorer har olika stort inflytande över pedagogiskt ledarskap, där rektorer har ett stort inflytande och biträdande rektorer ett mindre men ändå märkbart inflytande.

Resultaten fördelade före och efter påbörjat Rektorsprogram visar en negativ trend för både betyg och pedagogiskt ledarskap. För nationella prov syns en liten positiv förändring för rektorer, men inte för biträdande rektorer. Förändringen är bara 2 procentenheter, men noterbar eftersom intraklasskorrelationerna i övrigt för nationella prov, både före och efter programstart, är noll.

Tabell 3 Rektorers och biträdande rektorers relationer till utfallet relativt andra skolfaktorer i procent (intraklasskorrelationer)

Ledarposition	Betyg		Nationella prov		Pedagogiskt ledarskap	
Rektor + bitr. rektor	1 (80)		0 (70)		6 (18)	
Rektor	4 (76)		1 (63)		19 (13)	
Biträdande	2 (81)		0 (75)		6 (15)	
	Före	Efter	Före	Efter	Före	Efter
Rektor + bitr. rektor	3 (80)	1 (81)	0 (74)	0 (70)	26 (18)	10 (18)
Rektor	7 (73)	4 (77)	0 (67)	2 (63)	47 (26)	21 (14)
Biträdande	3 (81)	2 (83)	0 (76)	0 (75)	32 (12)	8 (19)

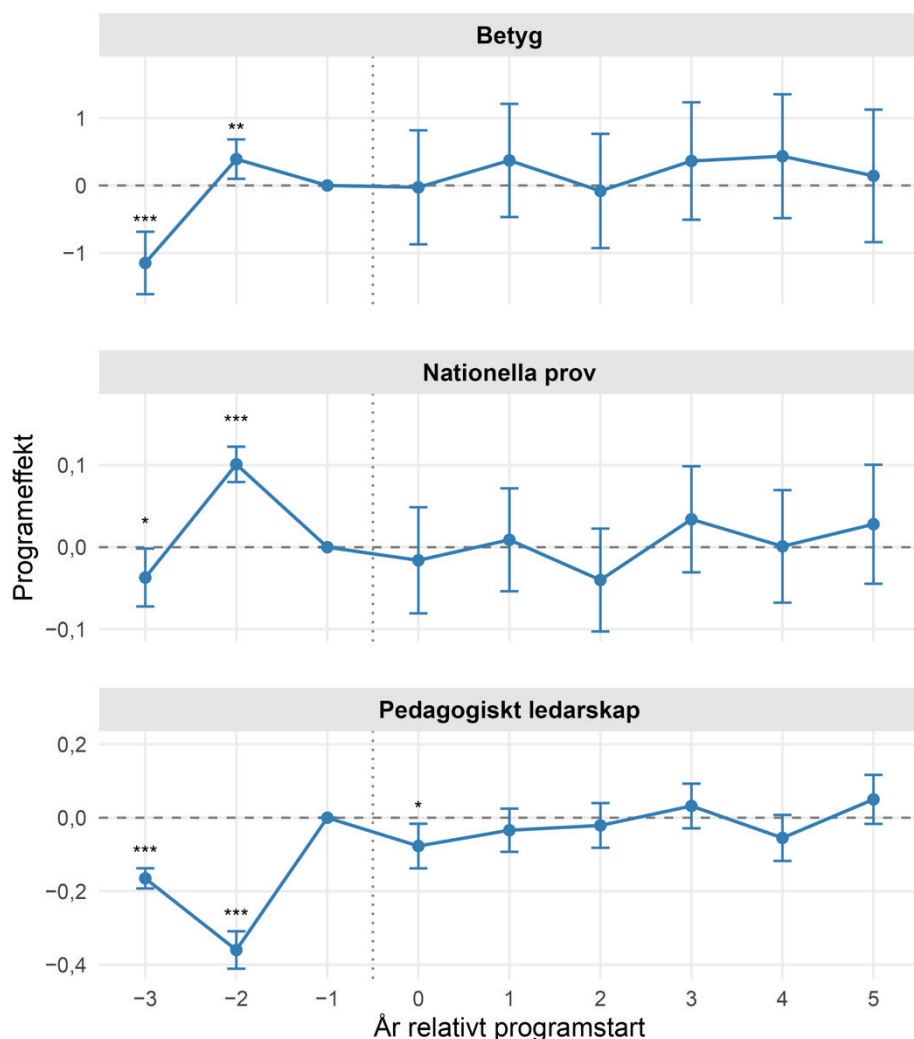
Not: Värdena anger intraklasskorrelationer (ICC) uttryckta i procent från flernivåmodeller där varje utfall (betyg, nationella prov, pedagogiskt ledarskap) dekomponeras i variationskomponenter på olika nivåer. Värdet utanför parentes anger hur stor andel av variationen i utfallet som tillskrivs den aktuella ledarpositionen (rektor och/eller biträdande rektor), medan värdet inom parentes anger andelen variation som tillskrivs andra skolfaktorer. Procentandelarna kan inte summeras till 100, eftersom återstoden utgörs av inomskolvvariation och en felterm. Den övre panelen redovisar ICC över hela observationsperioden; den nedre redovisar separata skattningar för perioden före respektive efter rektorns deltagande i Rektorsprogrammet. Analyserna baseras på 3 436 personer för vilka data finns från och med 2013, då både rollerna rektor och biträdande rektor finns dokumenterade i data.

3.3 Huvudanalys

3.3.1 Rektorsprogrammets effekter

Huvudanalyserna har gjorts i tre olika modeller, en per utfall: betyg, nationella prov och pedagogiskt ledarskap. Figur 1 illustrerar resultaten och bygger på skattningar som redovisas mer detaljerat i Tabell A2 i appendix. Figuren visar utfallens utveckling över tid, år för år, i relation till tidpunkten då rektorerna började på Rektorsprogrammet och i relation till kontrollgruppen. År 0 är året då programmet startade.

Figur 1 Huvudresultat för Rektorsprogrammets effekt på genomsnittsbetyg och nationella prov (årskurs 9) samt pedagogiskt ledarskap



Not: Punktskattningar med 95 procents konfidensintervall. Programeffekterna är skattningarna från och med programstart år 0. År -1 är referenskategori. Skattningarna anges i respektive utfallsmåtts ursprungliga enheter: betygspoäng (genomsnitt ≈ 221), provspoäng för nationella prov (genomsnitt ≈ 13), och skalpoäng för pedagogiskt ledarskap (skala 1–4). *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.

Till att börja med ska åren före programstart observeras. För analysens trovärdighet bör skillnaderna mellan programdeltagande och kontrollgruppen under föreperioden vara små och inte uppvisa några tydliga mönster över tid. För alla tre utfallen handlar det om förhållandevis små skillnader, vilket stärker tillförlitligheten i jämförelserna mellan programdeltagande och kontrollgrupp.

Exempelvis visar skattningen ungefär 1,15 betygspoäng lägre i behandlingsgruppen tre år före programstart, vilket kan sättas i relation till utfallets genomsnittliga betygspoäng på cirka 221. Det bör dock påpekas att även om skillnaderna framstår som små, är de i de flesta fall statistiskt signifikanta och i vissa fall inte helt obetydliga (t.ex. vad gäller pedagogiskt ledarskap). Dessutom varierar skattningarna en hel del under föreperioden. Dessa mönster föranleder viss försiktighet när slutsatser om programeffekter dras.

Åren efter programstart finns det inga statistiskt signifikanta effekter för betyg under något av de observerade åren. Skattningarna ligger nära noll och varierar mellan ungefär $-0,1$ och $+0,4$ betygspoäng, vilket är mycket lite i relation till utfallets genomsnitt på cirka 221 poäng. För nationella provresultat är effekterna också mycket små och inte statistiskt signifikanta för något av de observerade åren. Skattningarna ligger samtliga år nära noll, mellan ungefär $-0,04$ och $+0,03$ poäng. För pedagogiskt ledarskap, som skattats av lärare på en skala från 1 till 4, noteras i de allra flesta fallen inga statistiskt signifikanta effekter. Det finns en signifikant negativ effekt vid programstart (år 0) på $-0,08$ poäng, motsvarande ungefär 2 procent av skalans omfång, vilket är en mycket liten effekt.

Det sammantagna intrycket från Figur 1 är att deltagande i Rektorsprogrammet inte nämnvärt verkar ha påverkat de utfall vi studerar. Det finns framför allt inga tecken på positiva effekter.⁶

3.3.2 Separata effekter för rektorer och biträdande rektorer

Analysen av hur effekten av Rektorsprogrammet varierar mellan ledarpositioner redovisas i Tabell A2 (se interaktionstermen mellan ledarposition och programeffekt). Analysen visar att biträdande rektorer och rektorer inte uppvisar signifikant olika effekter på elevresultat (betyg eller nationella prov). För pedagogiskt ledarskap däremot är programeffekten mer negativ för biträdande rektorer än för rektorer, med en liten effekt ($-0,16$ skalpoäng). Detta tyder på att programmet i något högre grad minskar det lärarskattade pedagogiska ledarskapet för biträdande rektorer än för rektorer, medan elevresultaten inte påverkas olika beroende på ledarposition.

⁶ Eftersom variabeln för ledarposition endast finns tillgänglig från 2013 och framåt, presenteras en robusthetstestning i appendix (Figur A1) där alla kohorter av rektorer (från och med läsåret 2009/2010) inkluderas för betyg som utfall. Analysen är baserad på 14 965 observationer från 3 422 rektorer och 1 900 skolor. Resultaten indikerar samma relationer som syns i huvudanalysen.

3.3.3 Olika effekter för olika rektorer

Tabell A2 presenterar även ”slumpeeffekterna” för rektorer. Dessa visar att det finns substantiell variation med avseende på hur rektorerna förändras av Rektorsprogrammet. Variansen i programeffekt är särskilt betydande för betyg, där standardavvikelsen är 10,77 poäng. För nationella prov är standardavvikelsen 0,80 poäng, och för pedagogiskt ledarskap 0,69 skattningspoäng. Detta betyder att det finns betydelsefulla skillnader mellan deltagarnas utfall av programmet i termer av betyg, nationella prov och pedagogiskt ledarskap, även om de genomsnittliga effekterna är små. Vissa deltagare uppvisar alltså positiva effekter medan andra uppvisar negativa effekter, vilket resulterar i ett genomsnitt nära noll.

Variationen mellan skolor är också substantiell, särskilt för betyg (standardavvikelse 14,10) och för nationella prov (standardavvikelse 0,84), vilket visar att skolkontexten spelar stor roll för elevresultaten oberoende av programdeltagande.

För betyg och nationella prov har vi genomfört analyser för att undersöka om programeffekterna varierar beroende på skolornas socioekonomiska sammansättning, mätt genom föräldrars genomsnittliga utbildningsnivå. Resultaten indikerar att programeffekterna systematiskt varierar med skolornas socioekonomiska status. För betyg ökar programeffekten med cirka 5,8 betygspoäng för varje enhetsökning i föräldrars genomsnittliga utbildningsnivå. För nationella prov är motsvarande effekt cirka 0,8 provpoäng per enhet. Detta innebär att skolor med högre socioekonomisk status har mer positiva (eller mindre negativa) programeffekter vad gäller både betyg och nationella prov. Mönstret tyder på att programmet i större utsträckning gynnar skolor med socioekonomiskt starkare elevsammansättning.

3.4 Experimentdesign – en kompletterande analys

Vi har också genomfört en analys av en mindre grupp rektorer som följer en experimentell design (se Tabell A3–A6 i appendix). Kontrollgruppen utgörs av de deltagare i Rektorsprogrammet som stått på väntelista via lottning innan programstart (n=123) och behandlingsgruppen av de som under samma tidsperiod ansökt och kommit in direkt (n=737). Vi jämför utfallsvariablerna på de skolor vars rektorer börjar på programmet direkt när de ansökt, före och efter programstart, med utfallsvariablerna på de skolor vars rektorer står på väntelista, före och efter ansökan. Även i dessa analyser kontrollerar vi för tid, föräldrarnas genomsnittliga utbildningsnivå per skola och år, skolstorlek, antalet rektorstjänster per skola, typ av huvudman och hur länge rektorerna varit på väntelista respektive inskrivna på Rektorsprogrammet.

I analysen finner vi ingen skillnad mellan kontroll- och behandlingsgruppen med avseende på utfallet pedagogiskt ledarskap. För betyg och nationella prov finner vi mindre skillnader som dock inte är statistiskt signifikanta.

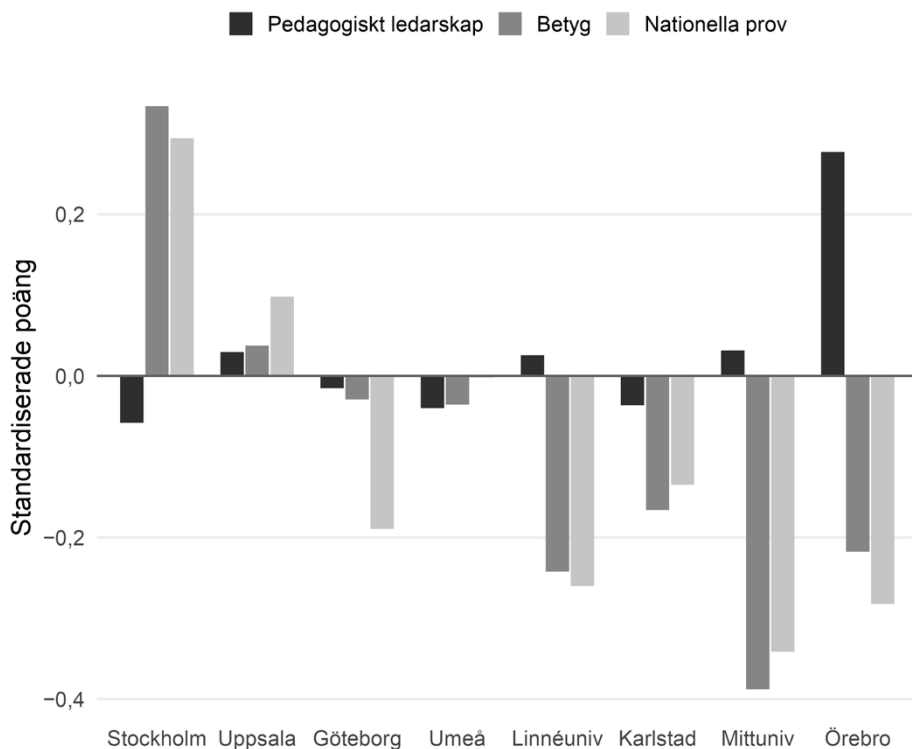
Rektorer som stått på väntelista kunde söka till programmet igen och blev ofta antagna en termin eller ett år senare, vilket gör att skillnader mellan kontroll- och behandlingsgrupp bara testat kortsiktiga effekter av programdeltagande. Två andra begränsningar är att det är få rektorer i kontrollgruppen och att det finns skillnader mellan behandlingsgruppen och kontrollgruppen med avseende på vissa bakgrundsfaktorer (se Tabell A3). Även om analysens resultat ligger väl i linje med huvudresultaten bör slutsatserna därför omgärdas av vissa reservationer.

3.5 En beskrivning av pedagogiskt ledarskap, betyg och nationella prov vid olika lärosäten

Vi har avslutningsvis undersökt om det finns skillnader i de tre utfallen – pedagogiskt ledarskap, betyg och nationella provresultat – beroende på vid vilket lärosäte deltagarna har gått Rectorsprogrammet. Detta har vi gjort i en ren beskrivande analys. Skillnader mellan lärosätena kan bero på en mängd olika faktorer och analysen svarar därför inte på frågan om vilken kvalitet de olika rektorsutbildningarna håller.

Sammanlagt har åtta lärosäten gett Rectorsprogrammet, men bara sex av dem har fortfarande programmet: Göteborgs universitet, Karlstads universitet, Linnéuniversitetet, Stockholms universitet, Umeå universitet och Uppsala universitet. Två lärosäten i våra data har tidigare haft Rectorsprogrammet: Mittuniversitetet och Örebro universitet, av vilka det senare drevs i samarbete med Mälardalens och Dalarnas högskolor. Vi har jämfört lärosätenas medelvärden, baserat på aggregerade data över tid per deltagare från och med att de börjat på Rectorsprogrammet, i en ANOVA-analys. Figur 2 illustrerar medelvärdena med standardiserade poäng.

Figur 2 Utfall per lärosäte, medelvärden över tid



Not: Medelvärden per lärosäte för de tre utfallen, beräknade på aggregerade data över tid per deltagare från och med programstart. Utfallen är standardiserade (medelvärde 0, standardavvikelse 1), vilket innebär att staplarna visar avvikelser från genomsnittet för samtliga deltagare i standardavvikelser. Linnéuniv = Linnéuniversitetet, Mittuniv = Mittuniversitetet. Örebro avser Örebro universitet i samarbete med Mälardalens och Dalarnas högskolor. Skillnaderna mellan lärosätena har prövats med ANOVA och parvisa jämförelser mellan lärosätena (se texten). Figuren är beskrivande och skillnaderna kan inte tolkas som skillnader i utbildningarnas kvalitet.

Den största skillnaden i genomsnittsbetyg på skolor vars rektorer gått Rectorsprogrammet finns mellan Mittuniversitetet och Stockholms universitet. I analysen som jämför lärosätenas medelvärden är denna skillnad (drygt 20 poäng) statistiskt signifikant. Betygssnittet vid skolor vars rektorer gått programmet vid Stockholms universitet är signifikant högre än vid skolor där rektorerna gått programmet vid något av de andra lärosätena. Också medelvärdet för nationella prov är signifikant högre vid skolor vars rektorer gått Rectorsprogrammet vid Stockholms universitet jämfört med övriga lärosäten, förutom för Uppsala universitet. En sådan skillnad beror sannolikt på relativt höga betygsresultat i Storstockholm och Uppsala jämfört med de nationella genomsnitten och framför allt jämfört med skolor i glesbygd och småstad (Holmlund m.fl. 2020).

Lärares bedömning av pedagogiskt ledarskap för rektorer skiljer sig oftast inte mellan lärosätena, förutom för rektorer som gått Rektorsprogrammet vid Örebro universitet. Skillnaden mellan Örebro universitet och medelvärdet för alla lärosäten är statistiskt signifikant, men motsvarar bara 0,14 skalpoängs skillnad. Vår analys kan inte fastställa om Rektorsprogrammet vid Örebro universitet är bättre på att träna rektorer i pedagogiskt ledarskap, eller om resultaten påverkas av andra faktorer som skiljer sig åt mellan skolorna där rektorerna arbetar. För att göra det krävs andra analysmetoder och mer data.

4 Diskussion

Vi har utvärderat sambandet mellan rektors och biträdande rektors deltagande i Rektorsprogrammet och utfall på skolenivå. Våra resultat tyder på att deltagande i programmet i genomsnitt inte leder till förbättringar av elevers studieresultat (betyg och resultat på nationella prov) eller i det pedagogiska ledarskapet på kort till medellång sikt (upp till fem år efter påbörjad utbildning).

Den tidigare forskningen om effekter av ledarskapsutbildningar för rektorer är begränsad, och det saknas särskilt studier genomförda i en svensk kontext. Den internationella litteraturen domineras dessutom av korrelationsstudier. Mot denna bakgrund utgör vår studie ett viktigt bidrag.

Resultaten bör samtidigt tolkas med viss försiktighet. Analysen bygger på en kvasiexperimentell design där tolkningen vilar på antaganden som inte fullt ut kan prövas, bland annat att deltagar- och kontrollgrupp skulle ha utvecklats likartat i frånvaro av programmet. Trots att skattningarna för åren före programstart i huvudsak är små, uppvisar de viss variation och i flera fall finns det statistiskt signifikanta skillnader. Det skulle kunna innebära att grupperna inte är helt jämförbara. Det är också svårt att helt utesluta att andra förändringar som sammanfaller med programstarten påverkar utfallen eller att det finns mätfel eller ”spillover” mellan deltagare och jämförelsegruppen som påverkar resultaten. Slutsatsen att programmet inte ger några tydliga positiva effekter framstår som robust, men bör alltså omgärdas med vissa reservationer. De skattade effekterna ska därför betraktas som indikationer snarare än definitiva effektstorlekar.

Våra resultat belyser en långvarig utmaning inom ledarskapsforskningen: klyftan mellan ledarskapsutbildning och dess faktiska genomslag i ledares beteende och olika utfall (Collins & Holton III 2004; Skylar Powell & Yalcin 2010). Trots att våra resultat kan uppfattas som nedslående ligger de i linje med

tidigare forskning och pågående debatter om huruvida ledarskapsutbildningar faktiskt bidrar till konkreta resultatförbättringar (Martin m.fl. 2021).⁷

Pedagogiskt ledarskap står i direkt relation till rektorers arbete på skolan. Våra inledande beskrivande analyser visade viss negativ förändring i rektorers pedagogiska ledarskap efter påbörjat deltagande i Rektorsprogrammet. Dessa beskrivande resultat kan dock bero på andra faktorer, till exempel att många rektorer som påbörjar programmet är nya i rektorsrollen och därför inte hunnit etablera ett pedagogiskt ledarskap på sina skolor. En annan tänkbar förklaring kan vara att rektorerna under utbildningen använder 20 procent av sin arbetstid till studier. Det kan innebära att de blir mindre närvarande i det dagliga arbetet på sina skolor, vilket lärare kan uppfatta som försämrat ledarskap.

En aspekt av rektorers betydelse i skolan som också framkommer i de beskrivande analyserna är att cirka 4 procent av variationen i betyg kan tillskrivas rektorer. Detta kan ses som ett mått på hur stor del av alla faktorer som påverkar skolors prestationer som kan kopplas till just rektorer. Resultatet – att rektorer bara påverkar några få procent av den totala variationen – stämmer väl överens med tidigare forskning. I den beskrivande analysen framgår vidare att rektorer tycks ha större betydelse än biträdande rektorer. Detta kan ses som en indikation på att rektorsfunktionen, med det yttersta ansvaret för att leda och samordna det pedagogiska arbetet på skolan, är viktigare för elevresultaten än andra ledningsfunktioner.

Resultatskillnader mellan skolor förklaras givetvis också av andra faktorer än rektorer. Vi vet till exempel från tidigare forskning att lärarkvalitet är en betydande faktor. I grundskolan kan en högeffektiv lärare förbättra elevresultat med 5–10 procentenheter per år, och effekten ackumuleras över tid (Chetty m.fl. 2014; Hanushek & Rivkin 2012). I ljuset av dessa faktorer är rektorers betydelse liten.

Ändå bör man komma ihåg att rektorer kan vara viktiga för att skapa ett positivt skolklimat och en fungerande verksamhet som förhåller sig till de lagar och regler som gäller för skolan. Till exempel kan Rektorsprogrammet spela en viktig roll i att förbereda skolledare för administrativa uppgifter och skolans övergripande ledning. Vår analys säger inget om programmets effekter inom dessa dimensioner. En annan viktig aspekt som lyfts fram i kvalitativ forskning om Rektorsprogrammet är betydelsen av utbildningens sociala dimension. Skott och Törnsén (2018) betonar att rektorers lärande både är en individuell och en kollektiv process, där kollegiala nätverk spelar en viktig roll för professionell

⁷ Det kan dock vara intressant att notera att en del tidigare forskningen har visat att program som kombineras med efterföljande handledning kan ge vissa förbättringar i elevresultat (Masters m.fl. 2022).

utveckling. Även Aas och Paulsen (2019) fann att deltagarna lättare kunde koppla utbildningen till sina egna erfarenheter och utmaningar när de fick möjlighet att reflektera tillsammans med andra rektorer. I Skolverkets redovisning av Rektorslyftet framhålls att många rektorer saknar tillgång till sådana nätverk (Skolverket 2012), vilket visar på behovet av strukturer för kollegialt lärande. Rektorsprogrammet skulle därför kunna ses som en stödstruktur för sådant lärande.

Trots att vi inte hittar några positiva effekter av Rektorsprogrammet på de utfall som vi studerar kan programmet alltså ändå vara viktigt för skolväsendet. Programmet kan spela en central roll som introduktion till uppdraget och som stöd i den komplexa övergången från lärare till skolledare. Många rektorer bär med sig erfarenhet från klassrummet, men rektorsrollen innebär ett ledarskap och ett juridiskt ansvar som sträcker sig bortom lärarprofessionens ramar. Huruvida Rektorsprogrammet lyckas möta dessa behov – och i vilken utsträckning – är frågor som kräver vidare forskning. Det är först då vi fullt ut kan förstå programmets verkliga värde.

Referenser

- Aas, M. & Paulsen, J. M. (2019), National strategy for supporting school principal's instructional leadership: A Scandinavian approach, *Journal of Educational Administration*, 57(5), 540–553.
- Antonakis, J., Bastardo, N. & Rönkkö, M. (2021), On ignoring the random effects assumption in multilevel models: Review, critique, and recommendations, *Organizational Research Methods*, 24(2), 443–483.
- Blossing, U., Berg, G., Bergviken Rensfeldt, A., Duek, S., Ekholm, M., Gustafsson, A., Gustavsson, S., Hylander, I., Håkansson, J., Johansson-Hidén, B., Liljenberg, M., Pihlgren, A. S., Ryffé, D., Scherp, H.-Å., Svedberg, L., Wrethander, M. & Ärlestig, H. (2020), *Rektor i fokus: kunskap, värden och verktyg*. Studentlitteratur.
- Böhlmark, A. Grönqvist, E. & Vlachos, J. (2016), The headmaster ritual: The importance of management for school outcomes, *The Scandinavian Journal of Economics*, 118(4), 912–940.
- Boström, V. (2023), *Skoljuridik*. Liber.
- Callaway, B. & Sant'Anna, P. H. (2021), Difference-in-differences with multiple time periods, *Journal of Econometrics*, 225(2), 200–230.
- Chetty, R. Friedman, J. N. & Rockoff, J. E. (2014), Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood, *American Economic Review*, 104(9), 2633–2679.
- Collins, D. B. & Holton III, E. F. (2004), The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982 to 2001, *Human Resource Development Quarterly*, 15(2), 217–248.
- Corcoran, R. P. (2017), Preparing principals to improve student achievement, *Child & Youth Care Forum*, 46, 769–781.
- Corcoran, S. P. Schwartz, A. E. & Weinstein, M. (2012), Training your own: The impact of New York City's aspiring principals program on student achievement, *Educational Evaluation and Policy Analysis*, 34(2), 232–253.
- Donmoyer, R. Yennie-Donmoyer, J. & Galloway, F. (2012), The search for connections across principal preparation, principal performance, and student achievement in an exemplary principal preparation program, *Journal of Research on Leadership Education*, 7(1), 5–43.

- Doss, C. Zaber, M. A. Master, B. K. Gates, S. M. & Hamilton, L. S. (2022), The Relationship Between Measures of Preservice Principal Practice and Future Principal Job Performance, *Educational Evaluation and Policy Analysis*, 44(1), 3–28.
- Dutta, V. & Sahney, S. (2022), Relation of principal instructional leadership, school climate, teacher job performance and student achievement, *Journal of Educational Administration*, 60(2), 148–166.
- Grissom, J. A. & Loeb, S. (2009), Triangulating Principal Effectiveness: How Perspectives of Parents, Teachers, and Assistant Principals Identify the Central Importance of Managerial Skills. Working Paper 35, National Center for Analysis of Longitudinal Data in Education Research.
- Grissom, J. A. Mitani, H. & Woo, D. S. (2019), Principal preparation programs and principal outcomes, *Educational Administration Quarterly*, 55(1), 73–115.
- Gümüş, S. Şükrü Bellibaş, M. Şen, S. & Hallinger, P. (2024), Finding the missing link: Do principal qualifications make a difference in student achievement?, *Educational Management Administration & Leadership*, 52(1), 28–51.
- Hallinger, P. (2003), Leading educational change: Reflections on the practice of instructional and transformational leadership, *Cambridge Journal of Education*, 33(3), 329–352.
- Hanushek, E. A. & Rivkin, S. G. (2012), The distribution of teacher quality and implications for policy, *Annual Review of Economics*, 4(1), 131–157.
- Holmlund, H. Sjögren, A. & Öckert, B. (2020), *Jämlikhet i möjligheter och utfall i den svenska skolan*. IFAU-rapport 2020:7.
- Jerdborg, S. (2023), Participation in the Swedish national principal training programme: How does it intertwine with principals' practice?, *Educational Management Administration & Leadership*, 51(4), 932–948.
- Johansson, O. & Svedberg, L. (2016), *Att leda mot skolans mål: vägval och möjligheter*. Gleerups.
- Leithwood, K. Harris, A. & Hopkins, D. (2020), Seven strong claims about successful school leadership revisited, *School Leadership & Management*, 40(1), 5–22.

- Leithwood, K. Seashore, K. Anderson, S. & Wahlstrom, K. (2004), Review of research: How leadership influences student learning, University of Minnesota, Center for Applied Research and Educational Improvement.
- Martin, R. Hughes, D. J. Epitropaki, O. & Thomas, G. (2021), In pursuit of causality in leadership training research: A review and pragmatic recommendations, *The Leadership Quarterly*, 32(5), 101375.
- Master, B. K. Schwartz, H. Unlu, F. Schweig, J. Mariano, L. T. Coe, J. Wang, E. L. Phillips, B. & Berglund, T. (2022), Developing school leaders: Findings from a randomized control trial study of the executive development program and paired coaching, *Educational Evaluation and Policy Analysis*, 44(2), 257–282.
- Meyer, A. Richter, D. & Hartung-Beck, V. (2022), The relationship between principal leadership and teacher collaboration: Investigating the mediating effect of teachers' collective efficacy, *Educational Management Administration & Leadership*, 50(4), 593–612.
- Mundlak, Y. (1978), On the pooling of time series and cross section data, *Econometrica: Journal of the Econometric Society*, 69–85.
- Myung, I. J. (2003), Tutorial on maximum likelihood estimation, *Journal of Mathematical Psychology*, 47(1), 90–100.
- Orphanos, S. & Orr, M. T. (2014), Learning leadership matters: The influence of innovative school leadership preparation on teachers' experiences and outcomes, *Educational Management Administration & Leadership*, 42(5), 680–700.
- Perrone, F. & Tucker, P. D. (2019), Shifting profile of leadership preparation programs in the 21st century, *Educational Administration Quarterly*, 55(2), 253–295.
- Persson, R. (2023), *Ledarskapsmätning i skolan – en explorativ studie om svenska myndigheters mätningar av rektors pedagogiska ledarskap*. Center for Educational Leadership and Excellence.
- Proposition 2017/18:182 (2018), *Samling för skolan*.
- Sanfo, J.-B. M. (2020), A three-level hierarchical linear model analysis of the effect of school principals' factors on primary school students' learning achievements in Burkina Faso, *International Journal of Educational Research*, 100, 101531.

- Skolinspektionen (2012), *Rektors ledarskap med ansvar för den pedagogiska verksamheten*. Skolinspektionen.
- Skolinspektionen (2020), *Statistik från Skolenkäten*. Skolinspektionen.
- Skollagen (2010:800), *Skollagen*. Sveriges riksdag.
- Skolverket (2012), Redovisning av uppdrag att svara för fortbildning för rektorer – Rektorslyftet (Redovisning av regeringsuppdrag Dnr 68-2011:466). Skolverket. <https://www.skolverket.se/download/18.6bfaca41169863e6a65a019/1553964578862/pdf2885.pdf>.
- Skolverket (2020), *Rektorsprogrammet måldokument 2021–2027*. Skolverket. ISBN: 978-91-7559-359-3
- Skolverket (2024), *Rektorsomsättning – en utmaning på många grundskolor*. <https://www.skolverket.se/publikationer?id=12319>.
- Skott, P. & Törnsén, M. (2018), *Rektorer, utbildning och lärande*. IFAU rapport 2018:11.
- Skylar Powell, K. & Yalcin, S. (2010), Managerial training effectiveness: A meta-analysis 1952-2002, *Personnel Review*, 39(2), 227–241.
- Snijders, T. A. B. & Bosker, R. J. (2011), *Multilevel analysis: An introduction to basic and advanced multilevel modeling*, 2nd ed., Sage Publishers.
- Steele, J. L. Steiner, E. D. & Hamilton, L. S. (2021), Priming the leadership pipeline: School performance and climate under an urban school leadership residency program, *Educational Administration Quarterly*, 57(2), 221–256.
- Tan, C. Y. Gao, L. & Shi, M. (2022), Second-order meta-analysis synthesizing the evidence on associations between school leadership and different school outcomes, *Educational Management Administration & Leadership*, 50(3), 469–490.
- Vetenskapsrådet (2011), *Rektor – en forskningsöversikt 2000–2010*. Vetenskapsrådet.
- Weinstein, J. Azar, A. & Flessa, J. (2018), An ineffective preparation? The scarce effect in primary school principals’ practices of school leadership preparation and training in seven countries in Latin America, *Educational Management Administration & Leadership*, 46(2), 226–257.
- Wooldridge, J. M. (2016), *Introductory econometrics a modern approach*, South-Western Cengage Learning.

Wu, H. & Shen, J. (2022), The association between principal leadership and student achievement: A multivariate meta-meta-analysis, *Educational Research Review*, 35, 100423.

Appendix

I detta appendix presenteras kompletterande metodologiska resonemang. Vi redovisar också studiens huvudanalys mer detaljerat, en robusthetsanalys och resultaten från den kompletterande analysen som bygger på en experimentell ansats.

A. Valideringsstudie om pedagogiskt ledarskap

Syftet med en valideringsstudie är att undersöka om en grupp enkätfrågor faktiskt fångar det som avses – i vår analys rektorers pedagogiska ledarskap. Validitetsundersökningar handlar också om huruvida frågorna samvarierar på ett sätt som visar att de mäter samma sak. Man undersöker om frågorna tillsammans betar sig som förväntat när man jämför dem med något annat redan känt, vilket ger stöd för att innehållet är det avsedda. Frågorna i den aktuella skalan, som ingick i Skolenkäten 2010–2021, är: (1) ”Rektor har god kunskap om det dagliga arbetet i skolan”, (2) ”Rektor arbetar för att främja elevernas kunskapsutveckling” och (3) ”Rektor tar ansvar för det pedagogiska arbetet på skolan”. Lärare har angett sina svar på en fyrgradig skala från ”Stämmer inte alls” till ”Stämmer helt och hållet”.

I vår valideringsstudie (Persson 2023) samvarierade frågorna om pedagogiskt ledarskap i förväntad grad, det vill säga de kunde sägas vara reliabla. Vi kunde påvisa reliabilitet både mellan frågorna i skalan ($\alpha=0,89$) och mellan lärare på skolor ($ICC(1)=0,21-0,27$). Siffrorna betyder att frågorna samvarierar sinsemellan i tillräcklig grad – alltså att de verkar handla om samma sak – och att det finns en samvariation mellan lärare som bedömer samma rektor. Samstämmigheten bland lärare på samma skola och år vad gäller rektorers pedagogiska ledarskap ligger i genomsnitt mellan 21–27 procent.

Vi jämförde sedan skalan om pedagogiskt ledarskap med andra skalor som mäter samma företeelse, samt med lärares intentioner att stanna på sina skolor och i yrket. De olika skalorna om samma sak var högt korrelerade. Relationen mellan rektorers pedagogiska ledarskap och lärares intention att stanna på skolan var betydligt starkare ($r=0,62$) än relationen mellan rektorers pedagogiska ledarskap och intentionen att stanna i yrket ($r=0,30$), vilket var förväntat och påvisar att skalans innehåll är det förväntade.

B. Statistisk analys

I den statistiska analysen undersöker vi hur skolors betyg, nationella prov och lärares skattningar av rektors pedagogiska ledarskap utvecklas i relation till det att deltagare påbörjar Rektorsprogrammet. Förutom den longitudinella upp-

ställningen av data årsvis före och efter det att rektorer påbörjar programmet, drar analysstrategin också nytta av att deltagare på programmet påbörjar utbildningen vid olika tidpunkter.

Identifikation och antaganden

Identifikationen av effekten bygger på två egenskaper i data. För det första rullas Rektorsprogrammet ut stegvis över tid, så att deltagare påbörjar programmet i olika kohorter mellan 2009 och 2021. Detta gör det möjligt att konstruera kontrollgrupper av ännu icke-behandlade rektorer för varje behandlingskohort, i linje med Callaway och Sant’Annas (2021) ansats. För det andra byter cirka en tredjedel av rektorerna i urvalet skola under observationsperioden. Denna rörlighet skapar en delvis korsad datastruktur som gör det möjligt att separera variation som tillskrivs rektorer från variation som tillskrivs skolor.

En kausal tolkning av programeffekten villkoras av flera antaganden: För det första antar vi parallella trender, vilket betyder att i frånvaro av Rektorsprogrammet hade utfallen för deltagare och kontrollgrupp utvecklats likadant över tid. Detta antagande prövas indirekt genom koefficienterna för åren före programstart. Små och icke-signifikanta skattningar utan tydliga tidstrender ger stöd för antagandet, men flera av skattningarna är statistiskt signifikanta (se Figur 1 och Tabell A2), vilket motiverar viss försiktighet i tolkningen. För det andra antar vi att kontrollgruppens utfall inte påverkas av att andra rektorer deltar i programmet (ingen ”spillover” mellan deltagare och kontrollgrupp). För det tredje, vilket diskuteras nedan, antar vi att oobserverade rektors- och skolspecifika faktorer inte är korrelerade med tidsvarierande förklaringsvariabler, vilket är ett antagande som vi prövar formellt och hanterar genom att specificera korrelerade slumpeffekter.

Av de identifierande antagandena bedömer vi att de största riskerna rör parallella trender och selektion in i programmet. Rektorer som varit verksamma innan de påbörjat programmet utgör kontrollgruppen i den dynamiska analysen och det är möjligt att denna grupp selekterats till programmet annorlunda än de som påbörjat utbildningen i samband med påbörjad rektorstjänst.

Antagandet om spillover mellan deltagare och kontrollgrupp kan i viss mån ifrågasättas, eftersom rektorer hos samma huvudman eller inom samma geografiska område kan utbyta erfarenheter. Det skulle kunna påverka även icke-behandlade rektorsers praktik och föra behandlings- och kontrollgrupp närmare varandra, vilket skulle dra våra skattningar mot noll. Detta innebär att vi inte kan utesluta att verkliga effekter underskattas. Den korsade datastrukturen, där både rektorer och skolor modelleras som separata variationskällor, mildrar dock spillover som sker inom en och samma skola, eftersom sådana delade effekter

delvis absorberas i skolans variationskomponent. Spillover mellan skolor hanteras däremot inte av strukturen och kan inte uteslutas helt.

Kontrollgrupp

Att Rektorsprogrammet har olika kohorter av deltagare gör att vi kan konstruera en kontrollgrupp. Kontrollgruppen tas fram så att varje kohort efter påbörjat program jämförs med de deltagare som ännu inte påbörjat Rektorsprogrammet. Denna analysaspekt har motsvarigheter i både ekonomisk och psykologisk forskningsmetod, där den i ekonomisk metod hanteras i ”dynamisk skillnad-i-skillnader-analys” och i psykologisk forskning med ”multipla baslinjer”.

Den ekonomiska forskningen har utvecklat metoden för att konstruera kontrollgrupper när flera tidpunkter för behandling finns. Callaway och Sant’Anna (2021) påvisar att resultat blir felaktiga om datapunkter för dem som redan påbörjat en behandling behålls i analysen när kontrollgrupper för efterföljande kohorter konstrueras, om det samtidigt sker en förändring av effekten över tid. I deras modell hanteras detta genom att en effekt per kohort och år räknas fram, för att sedan aggregeras.

Callaway och Sant’Annas modell är dock svår att applicera på ledarskapsforskning, där den individuella variationen är huvudintresset för analys. För att säkerställa att rektorer är tillgängliga som enhet i analysen, istället för årsvisa kohorter av rektorer, använder vi flernivåanalys (Snijders & Bosker 2011). Vi konstruerar kontrollgrupper enligt Callaway och Sant’Annas metod där deltagare som påbörjat programmet utesluts ur kontrollgruppen för efterföljande kohorter, men använder flernivåanalys där standardfelen skattas enligt hur rektorer varierar, inte enligt estimatens osäkerhet per kohort och år. Beroendet mellan observationer som upprepas i flera kohortjämförelser hanteras i slumpdelen av flernivåmodellerna, så att standardfelen inte underskattas.

Analysmetod

Flernivåanalysen har en struktur med två nivåer: mätningar över tid inom rektorer och skillnader mellan rektorer. Parametrarna skattas med maximal sannolikhetsestimering (Myung 2003), som kan hantera obalanserade data och låter oss behålla datapunkter som annars hade exkluderats (t.ex. när relativt färre datapunkter finns för en tidpunkt jämfört med en annan).

I våra data utgör rektorer och skolor två överlappande grupperingar då ett resultat, till exempel ett betygsvärde, kan tillskrivas både en skola och en rektor verksam på samma skola innevarande år. Ett betygsvärde ett givet år kan tillhöra en hierarkisk struktur där skolan har samma rektor vid alla mättillfällen, eller en korsad struktur där mer än en rektor har tjänstgjort på skolan under observations-

perioden. Då rektorer relativt ofta byter arbetsplats (cirka 29 procent i vårt urval byter arbetsplats under observationsperioden) är strukturen delvis korsad, vilket gör det möjligt att separera variation i utfallsvariablerna som beror på rektorer från variation som beror på skolor. I den beskrivande analysen utnyttjar vi denna struktur för att jämföra hur mycket av den totala variationen i skolors utfall som kan tillskrivas skolor respektive rektorer. De intraklasskorrelationer som resulterar kan tolkas som den andel av variationen som finns hos rektorer respektive skolor. Att explicit skatta varianserna för både skolor och rektorer ökar den statistiska styrkan och möjliggör att storleksordningen för varianskomponenterna kan bestämmas i relation till den totala variansen. Metodiken hanterar information mellan skolor och rektorer så att enheter med olika mängd information, till exempel rektorer med olika långa uppföljningsperioder, hanteras utan att introducera bias. Detta innebär att både rektorer som byter skola och de som stannar på samma skola analyseras tillsammans. Att exkludera rektorer som inte byter skola skulle riskera att introducera snedvridna resultat då rörlighet sannolikt korrelerar med rektorsegenskaper eller omständigheter på skolor.

Utfallsvariablerna (avgångselevens betyg och resultat på nationella prov, samt lärares skattningar av rektors pedagogiska ledarskap) är genomsnittliga värden per skola och år. Det betyder att analyserna gäller skillnader mellan skolor och rektorer, inte elev- eller lärarvariationen som också omger de genomsnittliga utfallsvärdena. Våra skattningar representerar den genomsnittliga betydelsen av rektorer respektive skolor för skolornas resultat.

Eventstudien som är basen i analyserna ger också indikationer om antagandet om parallella trender. I analysen skattas en parameter för utfallet varje år före och efter programstart, med året innan programstart som referens. Koefficienterna för perioder före behandlingen (ofta kallade *leads* i den ekonometriska litteraturen) fungerar som direkta test av parallella trender: icke-signifikanta koefficienter indikerar att det inte finns systematiska skillnader i utfallsvariabelns utveckling innan behandling påbörjas.

En viktig skillnad mellan flernivåanalys och ekonometriska skillnad-i-skillnader-modeller är hur variationen som handlar om rektorer hanteras. Skillnad-i-skillnader-modeller använder vanligtvis så kallade fixa effekter. Vår modell använder i stället slumpmässiga effekter, där variansen för rektorer och skolor estimeras som egna parametrar. Slumpmässiga effekter uppfyller samma syfte som fixa effekter, det vill säga att kontrollera för systematiska faktorer mellan rektorer, men med skillnaden att all varians görs explicit och tolkningsbar. I tabellerna avser ”fixa effekter” flernivåmodellens regressions-

koefficienter, till skillnad från slumpeffekterna som fångar variationen mellan rektorer och skolor.

Statistisk modell

Våra huvudmodeller i den statistiska analysen specificeras som:

$$Y_{ijt} = \beta_0 + \sum_{k=-3}^{+5} (\gamma_k \text{Tid_relativt_programstart_ijt}^k + \sum_t \delta_t \text{År_t} + \beta_1 \text{Ledarposition_ijt} + \beta_2 (\text{Programdeltagande} \times \text{Ledarposition})_{ijt} + X'_{ijt} \beta + u_j + v_i + w_i \cdot \text{Programdeltagande_ijt} + \varepsilon_{ijt}$$

Y_{ijt} är utfallsvariabeln (betyg, nationella prov eller pedagogiskt ledarskap) för rektor i vid skola j och tidpunkt t . $\text{Tid_relativt_programstart_ijt}^k$ är en uppsättning indikatorvariabler för tid relativt till programstart (år -3 till år $+5$, där år -1 är referens), vilka fångar programeffektens utveckling över tid. År_t är en uppsättning dummyvariabler för kalenderår, där 2013 är referensår, och kontrollerar för generella tidstrender. Ledarposition_ijt anger om deltagaren är rektor eller biträdande rektor. Interaktionen $\text{Programdeltagande} \times \text{Ledarposition}$ testas om programeffekten skiljer sig mellan rektorer och biträdande rektorer. $\text{Programdeltagande_ijt}$ är en indikator som antar värdet 1 från och med programstart. Den ingår inte som egen fix effekt, eftersom indikatorerna för tid relativt programstart redan fångar nivåskillnaden före och efter programstart. Däremot ingår den i interaktionen ovan och i den rektorsspecifika slumpmässiga lutningen w_i . X'_{ijt} är en vektor av kontrollvariabler. Vektorn innehåller även klustermedelvärden för de tidsvarierande kontrollvariablerna, vilket är det som gör modellen till en CRE-specifikation. u_j är slumpmässiga effekter för skolor (intercept). v_i är slumpmässiga effekter för rektorer (intercept). $w_i \cdot \text{Programdeltagande_ijt}$ är slumpmässiga effekter för programeffekten per rektor, vilket tillåter heterogenitet i hur olika rektorer påverkas av programmet. ε_{ijt} är en residual.

Huvudmodellerna i denna rapport är estimerade som flernivåmodeller med slumpeffekter på rektors- och skolnivå. En grundläggande förutsättning för att en sådan modell ska ge tillförlitliga skattningar är det så kallade *random effects*-antagandet: att rektors- och skolspecifika oobserverade faktorer (representerade av slumpeffekterna) inte är korrelerade med de tidsvarierande förklaringsvariablerna i modellen. Om detta antagande inte håller blir skattningarna av programeffekten snedvridna och kan inte tolkas trovärdigt (Antonakis m.fl. 2021; Wooldridge 2016). I praktiken är detta antagande ofta orealistiskt i data.

En rektor som arbetar på en skola med särskilda karaktärsdrag, vad gäller elevsammansättning, lärarkår eller organisatoriska förutsättningar, möter sannolikt också systematiskt olika villkor när det gäller de kontrollvariabler vi observerar. Antagandet om okorrelerade slumpeffekter behöver därför prövas snarare än förutsättas.

För att hantera detta använder vi en så kallad *correlated random effects*-specifikation (CRE), ursprungligen formulerad av Mundlak (1978). Tekniken innebär att klustermedelvärden för tidsvarierande kontrollvariabler inkluderas som ytterligare kontroller i modellen. När detta görs blir skattningen av program-effekten matematiskt ekvivalent med vad som i ekonometrin kallas en *fixed effects*-skattning, det vill säga en skattning som bygger på variation inom rektor respektive inom skola över tid, och som är robust mot godtycklig korrelation mellan oobserverade rektors- eller skolspecifika faktorer och de tidsvarierande oberoende variablerna (Antonakis m.fl. 2021; Wooldridge 2016).

För att en variabel ska kunna ingå som klustermedelvärde i CRE-specifikationen krävs att den uppvisar tillräcklig variation inom kluster över tid. För denna analys uppfyller två av de tidsvarierande kontrollvariablerna detta kriterium: andelen nyinvandrade elever (variation inom skola) och rektorns tjänstgöringstid vid skolan (variation inom rektor). Övriga kontrollvariabler uppvisar mycket begränsad variation inom kluster över tid och fungerar därför i praktiken som kontroller för stabila skolnivåskillnader.

Antagandet om okorrelerade slumpeffekter prövas formellt med ett Wald-test på den gemensamma signifikansen av klustermedelvärdena. För samtliga huvudmodeller förkastas antagandet statistiskt (genomsnittsbetyg $\chi^2(2) = 8,43$, $p = 0,015$; nationella prov $\chi^2(2) = 33,35$, $p < 0,001$; pedagogiskt ledarskap $\chi^2(2) = 16,57$, $p < 0,001$), vilket motiverar valet av CRE-specifikationen som huvudmodell. De praktiska skillnaderna mellan en specifikation med okorrelerade slumpeffekter och CRE-specifikationen är dock små för betyg och måttliga för övriga utfall, vilket stärker robustheten i de redovisade resultaten.

Beskrivande statistik för deltagare i Rektorsprogrammet

Som komplement till resultatredovisningen sammanfattar Tabell A1 centrala bakgrundsuppgifter för de 3 783 deltagare som ingår i analysurvalet, det vill säga de programdeltagare som under perioden 2009–2021 varit verksamma som rektorer eller biträdande rektorer på skolor med årskurs nio. Tabellen redovisar urvalets sammansättning med avseende på kön, utbildningsbakgrund, roll och erfarenhet vid ansökningstillfället samt skoltyp och anställningsförhållanden.

Tabell A1 Beskrivande statistik för deltagare i Rektorsprogrammet

	Antal	Procent
Urvalets storlek		
Antal deltagare	3 783	—
Antal årsvisa observationer	18 563	—
Antal skolor	2 147	—
Kön		
Kvinna	2 563	67,8 %
Man	1 220	32,2 %
Utbildningsbakgrund		
Har gått lärarutbildning	3 121	82,5 %
Roll vid ansökningstillfället		
Omfattas av obligatoriet	1 858	49,1 %
Ansvarar för systematiskt kvalitetsarbete	1 416	37,4 %
Rektorserfarenhet vid ansökningstillfället		
0–1 år	744	19,7 %
1–3 år	482	12,7 %
3–5 år	2 180	57,6 %
Mer än 5 år	377	10,0 %
Skoltyp (baserat på 2 147 skolor)		
Fristående	481	22,4 %
Kommunal	1 666	77,6 %
Övriga uppgifter om rektorsers anställning		
Genomsnittligt antal skolor för rektorer som bytt skola	2,3	—
Bytte skola under observationsperioden	1 079	29 %

Not: Strecket (—) anger att procentangivelse inte är meningsfull.

Urvalet domineras av kvinnor (67,8 procent) och av deltagare med lärarutbildning i botten (82,5 procent), vilket speglar rektorskårens sammansättning i stort. Vid ansökningstillfället omfattades knappt hälften (49,1 procent) av det formella obligatoriet att genomgå programmet. Erfarenhetsfördelningen visar att en majoritet hade tre till fem års erfarenhet som skolledare (57,6 procent), medan grupperna med kortast respektive längst erfarenhet var mindre. Deltagarna var verksamma vid 2 147 skolor, varav drygt en femtedel (22,4 procent) var fristående. Rörligheten mellan skolor var betydande: knappt en tredjedel (29 procent) bytte skola minst en gång under observationsperioden, vilket är den variation som möjliggör separationen av rektors- och skoleffekter i analysen.

C. Resultat från flernivåanalyserna

Tabell A2 presenterar resultaten från flernivåanalyserna som ligger till grund för Figur 1 i resultatredovisningen.

Tabell A2 Huvudresultat för Rektorsprogrammets effekt på genomsnittsbetyg och nationella prov (årskurs 9), samt pedagogiskt ledarskap

	Genomsnittsbetyg	Nationella prov	Pedagogiskt ledarskap
Fixa effekter			
(Intercept)	132,63*** (1,80)	8,24*** (0,13)	1,81*** (0,11)
År relativt programstart			
År -3	-1,15*** (0,24)	-0,04* (0,02)	-0,16*** (0,01)
År -2	0,39** (0,15)	0,10*** (0,01)	-0,36*** (0,03)
Programstart	-0,03 (0,43)	-0,02 (0,03)	-0,08* (0,03)
År 1	0,37 (0,43)	0,01 (0,03)	-0,03 (0,03)
År 2	-0,08 (0,43)	-0,04 (0,03)	-0,02 (0,03)
År 3	0,36 (0,44)	0,03 (0,03)	0,03 (0,03)
År 4	0,43 (0,47)	0,00 (0,03)	-0,05 (0,03)
År 5	0,14 (0,50)	0,03 (0,04)	0,05 (0,03)
Kontrollvariabler			
Kalenderår 2014	2,05*** (0,26)	-0,22*** (0,02)	0,31*** (0,03)
Kalenderår 2015	14,27*** (0,29)	-0,45*** (0,02)	0,12*** (0,03)
Kalenderår 2016	13,76*** (0,32)	0,12*** (0,02)	0,23*** (0,03)
Kalenderår 2017	16,31*** (0,35)	-0,40*** (0,03)	0,11** (0,03)
Kalenderår 2018	18,50*** (0,38)	1,00*** (0,03)	0,22*** (0,03)
Kalenderår 2019	16,25*** (0,42)	-0,43*** (0,03)	0,13*** (0,04)
Kalenderår 2020	17,50*** (0,46)	—	0,02 (0,04)
Kalenderår 2021	16,44*** (0,50)	—	0,11** (0,04)
Kalenderår 2022	12,18*** (0,54)	-0,54*** (0,04)	—
Kalenderår 2023	10,30*** (0,59)	-0,12** (0,04)	—
Kön	0,10 (0,32)	0,01 (0,02)	0,00 (0,02)
Huvudman	-12,52*** (0,89)	-0,66*** (0,06)	0,24*** (0,03)
Föräldrars utbildningsnivå	40,69*** (0,61)	2,47*** (0,04)	-0,15*** (0,04)
Skolstorlek	0,04*** (0,01)	0,00*** (0,00)	0,00 (0,00)
Antal rektorstjänster	0,18 (0,19)	0,09*** (0,01)	0,07*** (0,01)
Erfarenhet i ledarposition	-0,08 (0,19)	0,01 (0,01)	0,00 (0,01)
Andel nyinvandrade elever	-0,95*** (0,01)	-0,02*** (0,00)	0,01*** (0,00)
Andel nyinvandrade elever (klustermedel)	-0,09** (0,03)	-0,01*** (0,00)	-0,01*** (0,00)
Skolbyte	0,28 (0,24)	0,09*** (0,02)	0,02 (0,02)
Tjänstgöringsår på skola	0,00 (0,07)	0,01* (0,01)	-0,02*** (0,00)
Tjänstgöringsår på skola (klustermedel)	0,02 (0,09)	0,00 (0,01)	0,01 (0,00)
Ledarposition × programeffekt	-0,09 (0,39)	0,05 (0,03)	-0,16*** (0,02)
Slumpeffekter			
Rektor (Intercept)	84,76 (9,21)	0,43 (0,65)	0,28 (0,53)
Rektor (programeffekt)	116,00 (10,77)	0,63 (0,80)	0,48 (0,69)
Skola (Intercept)	198,75 (14,10)	0,70 (0,84)	0,09 (0,30)
Residual	52,02 (7,21)	0,24 (0,49)	0,03 (0,18)

Forts. nästa sida

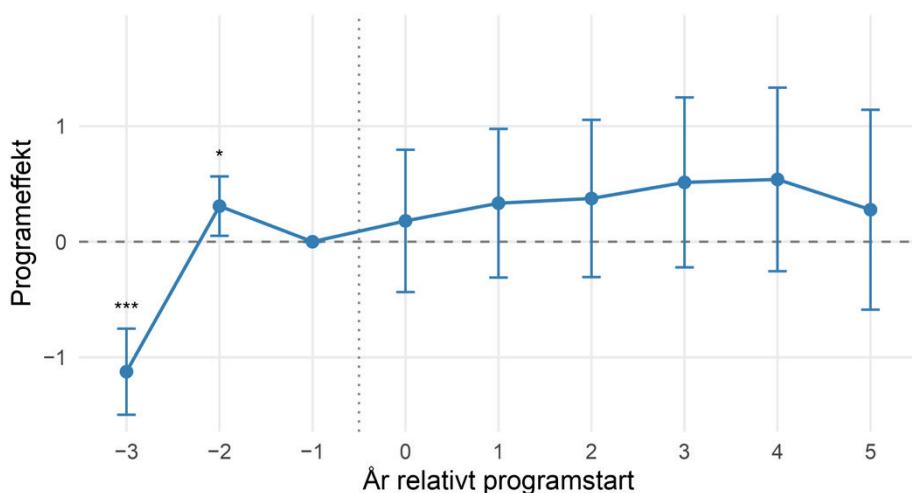
	Genomsnittsbetyg	Nationella prov	Pedagogiskt ledarskap
Antal observationer			
N observationer	12 595	9 973	4 458
N rektorer	3 110	3 001	2 382
N skolor	1 753	1 706	1 390

Not: Estimat med standardfel i parenteser. För slumpeffekterna redovisas standardavvikelse i parenteserna. Modellen är estimerad med korrelerade slump effekter (CRE) enligt Antonakis m.fl. (2021), där klustermedelvärden inom TimeGroup för andel nyinvandrade elever (skolnivå) och tjänstgöringsår på skola (rektorsnivå) inkluderats som kontroller. Wald-test för antagandet om okorrelerade slump effekter: Genomsnittsbetyg $\chi^2(2) = 8,43$, $p = 0,015$; Nationella prov $\chi^2(2) = 33,35$, $p < 0,001$; Pedagogiskt ledarskap $\chi^2(2) = 16,57$, $p < 0,001$ – samtliga test förkastar formellt antagandet, men praktiska skillnader mellan RE- och CRE-skattningar är små för betyg och måttliga för övriga utfall. I År relativt programstart är År -1 referensår och 2013 är referens för kalenderår. Streck (—) anger år som saknas i data för det specifika utfallet (nationella prov saknas för 2020–2021 på grund av covid-19; vi har inte data från lärarenkäten 2022–2023). * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

Nationella prov ställdes in våren 2020 och 2021 till följd av covid-19-pandemin, vilket innebär att nationella provresultat saknas för dessa två år. Eftersom analysen bygger på rektorernas tid i förhållande till programstart snarare än på enskilda kalenderår påverkas skattningarna av programeffekten endast marginellt av detta bortfall, och estimaten för nationella prov de kalenderår som omger pandemin uppvisar inga tecken på avvikande mönster.

Figur A1 redovisar resultaten för betyg som utfall när alla kohorter i urvalet, från och med läsåret 2009/2010, är inkluderade. Analysen är baserad på 14 965 observationer från 3 422 rektorer och 1 900 skolor. Resultaten indikerar samma relationer som i huvudanalysen.

Figur A1 Resultat för Rektorsprogrammets effekt på genomsnittsbetyg (årskurs 9), 2009–2021



Not: Robusthetstestning med data 2009–2021. Modellen är skattad med okorrelerade slump effekter, inte med korrelerade slump effekter (CRE) som huvudanalysen, och kontrollerar endast för kalenderår, typ av huvudman och föräldrars utbildningsnivå, eftersom uppgifter om ledarposition och andel nyinvandrade elever saknas för de tidiga åren. Punktskattningar med 95 procents konfidensintervall. År -1 är referenskategori (noll). Kalenderår ingår som dummyvariabler med det första året i urvalet (2010) som referens. Signifikansnivåer: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.

D. Resultat från den experimentella analysen

I detta avsnitt redovisas resultaten från den kompletterande analys med experimentell design som beskrivs i resultatdelen (avsnitt 3.4). Vi presenterar först en beskrivande jämförelse av kontroll- och behandlingsgrupp för att pröva i vilken utsträckning grupperna är jämförbara vid utgångspunkten (Tabell A3). Därefter redovisar vi flernivåmodellerna, en per utfall, där den centrala parametern är korsnivåinteraktionen som testar om förändringen från före till efter programstart skiljer sig mellan direktantagna och väntelistekontroller (Tabell A4–A6).

Tabell A3 jämför väntelistekontrollerna ($n = 123$) med rektorer som påbörjade Rektorsprogrammet samma år som väntelistekohorten ($n = 737$), baserat på deras tidigaste observation i registret från och med 2013. Grupperna är väl balanserade på flera centrala dimensioner. Kön fördelningen är likartad (kvinnoandel 70,7 procent respektive 67,4 procent, $p = 0,469$), och andelen rektorer i offentlig verksamhet är jämförbar (73,2 procent mot 74,9 procent, $p = 0,683$). Föräldrarnas genomsnittliga utbildningsnivå vid de skolor rektorerna leder skiljer sig inte heller mellan grupperna ($M = 2,26$). Det innebär att de elevsammansättningsförhållanden som speglas i SALSA-måttet är jämförbara mellan kontroll- och behandlingsgrupp.

Tabell A3 Beskrivande jämförelse av väntelistekontroller och deltagare i Rektorsprogrammet vid första observationen

	Kontroll (n = 123)	Behandling (n = 737)	Test	p
Befattning			$\chi^2(1) = 13,78$	< 0,001
Biträdande rektor	98 (79,7 %)	460 (62,4 %)		
Rektor	25 (20,3 %)	277 (37,6 %)		
Kön			$\chi^2(1) = 0,53$	0,469
Kvinna	87 (70,7 %)	497 (67,4 %)		
Man	36 (29,3 %)	240 (32,6 %)		
Typ av huvudman			$\chi^2(1) = 0,17$	0,683
Enskild	33 (26,8 %)	185 (25,1 %)		
Offentlig	90 (73,2 %)	552 (74,9 %)		
Skolledartjänstens omfattning			$\chi^2(3) = 14,40$	0,002
10–49 %	0 (0,0 %)	5 (0,7 %)		
50–74 %	4 (3,3 %)	33 (4,5 %)		
75–99 %	57 (46,3 %)	217 (29,4 %)		
100 %	62 (50,4 %)	482 (65,4 %)		
Sammanlagd tid som skolledare (nivå 1–4)			$\chi^2(3) = 52,83$	< 0,001
1 (kortast)	12 (9,8 %)	99 (13,4 %)		
2	56 (45,5 %)	150 (20,4 %)		
3	43 (35,0 %)	463 (62,8 %)		
4 (längst)	12 (9,8 %)	25 (3,4 %)		
Lärosäte			$\chi^2(5) = 11,16$	0,048
Göteborgs universitet	14 (11,4 %)	95 (12,9 %)		
Karlstads universitet	8 (6,5 %)	93 (12,6 %)		
Mittuniversitetet	5 (4,1 %)	58 (7,9 %)		
Stockholms universitet	38 (30,9 %)	160 (21,7 %)		
Umeå universitet	29 (23,6 %)	194 (26,3 %)		
Uppsala universitet	29 (23,6 %)	137 (18,6 %)		
Föräldrars utbildningsnivå (SALSA), M (SD)	2,26 (0,24)	2,26 (0,24)	$t(794) = -0,16$	0,873

Not: N = 860 unika rektorer (123 kontroller, 737 behandlade). För kategoriska variabler anges värden som antal (kolumnprocent) och Pearsons chi-tvåtest. För den kontinuerliga SALSA-variabeln anges medelvärde (standardavvikelse) samt tvåsidigt t-test med antagande om lika varianser; Welch's t-test ($t(160,1) = -0,16$, $p = 0,874$) och Mann–Whitneys rangordningstest ($z = 0,43$, $p = 0,664$) gav samma slutsats. Saknade observationer på SALSA-variabeln ($n = 64$) exkluderades parvis.

Däremot framträder fyra systematiska skillnader. För det första är behandlingsgruppen i större utsträckning rektorer i formell mening (37,6 procent) jämfört med kontrollgruppen (20,3 procent), där biträdande rektorer dominerar (79,7 procent mot 62,4 procent). För det andra arbetar behandlade rektorer i större utsträckning på heltid (65,4 procent mot 50,4 procent), medan kontrollgruppen oftare har en tjänsteomfattning på 75–99 procent (46,3 procent mot 29,4 procent). För det tredje är den samlade skolledarerfarenheten ojämnt fördelad:

behandlingsgruppen är överrepresenterad i den näst högsta erfarenhetsnivån (62,8 procent i nivå 3 mot 35,0 procent), medan kontrollerna är överrepresenterade i nivå 2 (45,5 procent mot 20,4 procent). För det fjärde skiljer sig fördelningen över lärosäten gränssignifikant mellan grupperna: Stockholms universitet är överrepresenterat bland kontrollerna (30,9 procent mot 21,7 procent), medan Karlstads universitet är underrepresenterat (6,5 procent mot 12,6 procent).

Sammantaget visar resultaten att behandlings- och kontrollgrupperna är jämförbara vad gäller kön, huvudmannaskap och elevsammansättning (SALSA), men att de skiljer sig systematiskt avseende formell befattning, tjänstomfattning, samlad skolledarerfarenhet och, i mindre grad, lärosätesfördelning; dimensioner som väntelistans antagningsmekanism rimligen kan förklara. Det är till exempel sannolikt att man har prioriterat mer erfarna och mer formellt etablerade rektorer för omedelbara programplatser. Det kan också ha funnits regionala kapacitetsbegränsningar. Dessa obalanser bör beaktas vid tolkningen av experimentresultaten.

Tabell A4 till A6 redovisar de tre flernivåmodellerna för den experimentella jämförelsen, en för vart och ett av utfallen nationella provresultat, genomsnittligt meritvärde och pedagogiskt ledarskap. Modellerna kontrollerar för årgång, föräldrars utbildningsnivå (SALSA), lärartäthet, antal rektorstjänster och huvudmannatyp, samt inkluderar exponeringstid och samlad skolledarerfarenhet på mellannivå.

I samtliga modeller är korsnivåinteraktionen den parameter som motsvarar den experimentella programeffekten, det vill säga skillnaden mellan direktantagna och väntelistekontroller i förändring från före till efter programstart. För inget av de tre utfallen är denna interaktion statistiskt signifikant (nationella prov: estimat = 0,249, $p = 0,186$; meritvärde: estimat = -2,427, $p = 0,223$; pedagogiskt ledarskap: estimat = 0,035, $p = 0,774$). Resultaten överensstämmer därmed med huvudanalysen: vi finner inga statistiskt säkerställda effekter av programdeltagande på något av utfallen.

Sammantaget ger den experimentella analysen inget stöd för att Rektorsprogrammet påverkar elevresultat eller pedagogiskt ledarskap på kort sikt, vilket stärker tolkningen av huvudanalysens nollresultat. Samtidigt bör resultaten betraktas som indikativa snarare än definitiva – det finns obalanser mellan grupperna, antalet väntelistekontroller är begränsat och jämförelsen fångar i praktiken kortsiktiga skillnader eftersom väntelistan medger förnyad ansökan (se resultatdelen).

Tabell A4 Flernivåmodell med nationella provresultat i matematik, svenska och engelska (åk 9) som utfall (experimentell jämförelse, väntelista mot direktantagna till Rektorsprogrammet)

Parameter	Estimat	SE	Est./SE	p
Fixa effekter — inomnivå				
Årgång	-0,066	0,013	-4,883	< 0,001***
Föräldrars utbildning (SALSA)	4,619	0,157	29,367	< 0,001***
Lärartäthet	0,010	0,002	5,468	< 0,001***
Antal rektorstjänster	-0,182	0,064	-2,863	0,004**
Huvudmannatyp	-0,576	0,096	-5,974	< 0,001***
Fixa effekter — mellannivå				
Korsnivåinteraktion	0,249	0,188	1,323	0,186
Grupp	0,075	0,116	0,648	0,517
Exponering	-0,025	0,011	-2,332	0,020*
Sammanlagd tid	0,115	0,043	2,700	0,007**
Intercept				
Nationella prov	135,753	27,158	4,999	< 0,001***
Slumpterm (lutning Pre/Post)	-0,284	0,181	-1,564	0,118
Varianskomponenter				
Inomnivå (residual): Nationella prov	0,999	0,049	20,578	< 0,001***
Mellannivå (residual): Nationella prov	0,278	0,048	5,781	< 0,001***
Mellannivå (residual): slumpmässig lutning (Pre/Post)	0,049	0,071	0,689	0,491

Not: * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$. Korsnivåinteraktionen prövar om förändringen från före till efter i nationella provresultat skiljer sig mellan rektorer som antogs direkt (behandling) och rektorer som stått på väntelista (kontroll).

Tabell A5 Flernivåmodell med betyg (åk 9) som utfall (experimentell jämförelse, väntelista mot direktantagna till Rektorsprogrammet)

Parameter	Estimat	SE	Est./SE	p
Fixa effekter — inomnivå				
Årgång	-0,216	0,184	-1,175	0,240
Föräldrars utbildning (SALSA)	68,031	2,332	29,174	< 0,001***
Lärartäthet	0,077	0,032	2,399	0,016*
Antal rektorstjänster	-1,654	0,676	-2,445	0,014*
Huvudmannatyp	-14,913	1,675	-8,906	< 0,001***
Fixa effekter — mellannivå				
Korsnivåinteraktion	-2,427	1,992	-1,218	0,223
Grupp	1,350	1,925	0,701	0,483
Exponering	-0,410	0,196	-2,094	0,036*
Sammanlagd tid	0,272	0,681	0,399	0,690
Intercept				
Meritvärde	519,670	370,562	1,402	0,161
Slumpterm (lutning Pre/Post)	3,364	1,935	1,739	0,082
Varianskomponenter				
Inomnivå (residual): Meritvärde	140,968	6,303	22,365	< 0,001***
Mellannivå (residual): Meritvärde	163,124	20,758	7,859	< 0,001***
Mellannivå (residual): slumpmässig lutning (Pre/Post)	21,895	9,106	2,405	0,016*

Not: * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$. Korsnivåinteraktionen prövar om förändringen från före till efter i genomsnittligt meritvärde skiljer sig mellan rektorer som antogs direkt (behandling) och rektorer som stått på väntelista (kontroll).

Tabell A6 Flernivåmodell med pedagogiskt ledarskap som utfall (experimentell jämförelse, väntelista mot direktantagna till Rektorsprogrammet)

Parameter	Estimat	SE	Est./SE	p
Fixa effekter — inomnivå				
Årgång	-0,007	0,012	-0,558	0,577
Föräldrars utbildning (SALSA)	-0,110	0,071	-1,549	0,121
Lärartäthet	0,001	0,001	0,970	0,332
Antal rektorstjänster	0,027	0,037	0,733	0,464
Huvudmannatyp	0,275	0,038	7,257	< 0,001***
Fixa effekter — mellannivå				
Korsnivåinteraktion	0,035	0,123	0,287	0,774
Grupp	0,056	0,067	0,842	0,400
Exponering	-0,015	0,006	-2,805	0,005**
Sammanlagd tid	-0,007	0,026	-0,270	0,787
Intercept				
Pedagogiskt ledarskap	15,505	24,453	0,634	0,526
Slumpterm (lutning Pre/Post)	-0,026	0,122	-0,211	0,833
Varianskomponenter				
Inomnivå (residual): Pedagogiskt ledarskap	0,224	0,022	10,166	< 0,001***
Mellannivå (residual): Pedagogiskt ledarskap	0,046	0,013	3,488	< 0,001***
Mellannivå (residual): slumpmässig lutning (Pre/Post)	0,020	0,026	0,752	0,452

Not: * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$. Korsnivåinteraktionen prövar om förändringen från före till efter i pedagogiskt ledarskap skiljer sig mellan rektorer som antogs direkt (behandling) och rektorer som stått på väntelista (kontroll).

Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU) är ett forskningsinstitut under Arbetsmarknadsdepartementet med placering i Uppsala.

IFAU ska främja, stödja och genom forskning genomföra uppföljningar och utvärderingar.

Uppdraget omfattar effekter av arbetsmarknads- och utbildningspolitik, arbetsmarknadens funktionssätt och arbetsmarknadseffekter av socialförsäkringen.

I rapportserien presenteras såväl IFAU:s forskning som resultat av samarbeten med andra nationella och internationella forskningsorganisationer.

IFAU delar årligen ut bidrag till olika forskningsprojekt, vars resultat publiceras i rapportserien.

Rapporterna kan vara fristående eller publiceras tillsammans med ett Working paper.

Alla IFAU:s publikationer finns på www.ifau.se