

Repliker och kommentarer

I den här avdelningen välkomnas kommentarer till tidigare bidrag och korta inlägg med ekonomisk-politisk anknytning

SÖREN WIBE

Leder konkurrensen från friskolor till högre kvalitet i undervisningen?

En kritisk granskning av Bergström och Sandströms artikel

1. Inledning

I *Ekonomisk Debatt* nr 5 2001 publicerade två ekonomer från Handelns Utredningsinstitut, Mikael Sandström och Fredrik Bergström, en artikel om friskolor. Den återgav i koncentrerad form innehållet i en ESO-rapport (med samma författare); *Konkurrens bildar skolorna - en ESO-rapport om friskolornas betydelse för de kommunala skolorna*" (ESO [2001]).

I rapporten görs en statistisk studie över sambandet mellan andelen friskolor i landets kommuner och kvaliteten på undervisning i den kommunala skolan (mätt med elevernas betyg). Bergström och Sandström finner ett statistiskt säkerställt positivt samband. De förklarar detta med att konkurrensen från friskolor bidrar till bättre undervisning i alla skolor. Denna konkurrenseffekt är även mycket betydelse-

full ekonomiskt sett. En ökning av andelen friskolor i en kommun från noll till tio procent har enligt Bergström & Sandström [2001b] samma effekt på kvaliteten i skolan som en ökning av resurstilldelningen med mellan 25 och 50 procent. Kommunerna behöver alltså inte satsa på fler lärare eller bättre lokaler, allt som krävs är en ordentlig utbyggnad av friskolorna.

Sammantaget visar studien på kraftigt positiva effekter av friskolor, effekter som når alla elever. Det är detta som motiverar den rubrik som författarna valt för sin artikel i *Ekonomisk Debatt: Konkurrens mellan skolor - för barnens skull!*" . Bergström och Sandström väljer också att med full kraft föra ut resultatet i den politiska debatten i landet. Samma dag som ESO-rapporten publicerades skriver författarna en artikel på DN Debatt där de utan reservationer (och med versaler) presenterar huvudresultatet. Några veckor senare används studien i en debatt i riksdagen som ett viktigt vetenskapligt stöd för friskolorna.

Bergström och Sandström analyserar en viktig och allvarlig politisk fråga. De uttalar sig entydigt: vetenskapen stöder

SÖREN WIBE är fil dr och docent i nationalekonomi, professor i skogsekonomi vid Sveriges Lantbruksuniversitet i Umeå. Han har mest forskat kring frågor som rör produktionsteori och teknisk utveckling.

Tabell 1 Oberoende och beroende variabler i ESO-rapporten

”Individmaterialet”		”Kommunmaterialet”	
Beroende variabler	Oberoende variabler	Beroende variabel	Oberoende variabler
A-rätt	Kön	Medelbetyget i	Kostnaden per elev
B-rätt	Invandrarbakgrund	kommunerna	Andel lågutbildade i
C-betyg	Föräldrarnas utbildning	1992–97	kommunen
M-betyg	Lärartim/elev		Andel med utländsk
PG-betyg	Skolans storlek		bakgrund i kommunen
Betyg Ma	Skattekronor/inv		Andel friskolor
Betyg Eng	Förort		i kommunen
Betyg SV	Invånardistans		
Meritvärde	Andel friskolor i		
	kommunen		

Anmärkning: Bergström och Sandström skriver att de använt en variabel för ”Storstäder och förorter”. Det är emellertid fel: deras variabel avser endast förorter. ”Invånardistans” avser medelavståndet mellan kommunens invånare (i meter) om dessa tänkes spridda jämnt över kommunens yta.

hypotesen om den positiva friskoleffekten. När forskare så kraftfullt och reservationslöst ger sitt stöd åt en sida i en politiskt viktig fråga har man rätt att kräva att bevisen är otvetydiga. Det gäller inte i detta fall. Den statistiska analys som Bergström och Sandström gör tillåter inte de slutsatser som de drar. Det gör deras inlägg i friskoledebatten mer till politik än vetenskap och skadar på så sätt tilltron till ekonomer som objektiva vetenskapsmän. I denna kommentar redogör jag för den granskning jag gjort.

2. Datamaterial och databehandling

Den statistiska analysen i ESO-rapporten genomförs med två datamaterial. Dels ett ”individmaterial” som baseras på betyg och provsvar från ca 30 000 elever i åk 9 år 97/98, dels ett ”kommunmaterial” som avser medelbetygen i åk 9 i landets alla kommuner (dvs kommungenomsnitt) åren 1993–97.¹

Den statistiska analysen går ut på att korrelera betyg och resultat från nationella prov med ett antal variabler. Dessa presenteras i *Tabell 1*.

I princip analyserar alltså Bergström och Sandström (vad avser individmateria-

let) sambanden mellan nio olika betyg och nio förklarande variabler. Av de förklarande (oberoende) variablerna är tre individvariabler (de tre första), två skolvariabler (de två nästföljande) och fyra kommunvariabler (de fyra sista). De fem första beroende variablerna avser resultat från nationella prov i matematik (nationella prov genomförs i matematik, svenska och engelska). De tre nästa är renodlade betyg på en skala 1–4, och det sista (Meritvärde) är ett samlingsbetyg poängsatt från 0 till 320.

Då det gäller kommunmaterialet analyseras ett samband med fyra förklarande variabler med den beroende variabeln

¹ Författarna skriver genomgående att materialet avser åren 1993–97. Det korrekta torde dock vara åren 1992 samt 1994–97 eftersom året 1992, men inte år 1993, innehåller alla variabler i den databas som författarna refererar till. För år 1993 saknas data för en variabel, nämligen ”andel lågutbildade”. Jag vet inte hur författarna har behandlat detta; jag har frågat utan att få svar. Jag har valt att utelämnat år 1993 från analysen och studerar alltså åren 1992 samt 1994–97. Allt detta påverkar dock endast i ringa grad resonemang och resultat, och vanligen skriver jag därför 1992–97 trots att detta inte är helt korrekt.

som är genomsnittsbetyget i kommunen för alla elever i åk 9. Betygen åren 1992–97 sattes efter en skala 1–5.

Bergström & Sandström [2001a, 2001b] använder flera olika statistiska modeller. På individmaterialet görs t ex 15 skattningar utan valmodeller (OLS, random-effects-modeller, ordered probit) samt även 2 st sk valmodeller. I min granskning har jag begränsat mig till de skattningar som baseras på vanlig OLS. Skälet till min begränsning är att OLS ger författarnas starkaste resultat, dvs vid OLS-skattningarna erhåller ”friskoleeffekten” den starkaste statistiska signifikansen. Det skall villigt medges att det är teoretiskt möjligt att granskning av de andra modellerna skulle gett ett annat resultat (och jag vill gärna uppmantra andra forskare att genomföra denna analys). Att jag själv inte gjort denna granskning beror på tidsbrist och på att jag är övertygad om att min OLS-granskning är tillfyllest för att dra slutsatsen att Bergström och Sandströms analys har sådana brister att slutsatserna inte är trovärdiga.

3. En analys av individmaterialet

Alla som sysslat med statistiska analyser vet hur lätt det är att erhålla ett signifikant samband mellan två talserier utan att ett orsakssamband föreligger. Man kan faktiskt visa att det (nästan) alltid går att finna ett statistiskt signifikant samband mellan en beroende och 10 oberoende variabler, trots att alla talserier är helt slumpmässiga. I en statistisk analys är det därför nödvändigt att kontrollera att resultaten inte är tillfälliga. Man måste kontrollera att de står sig även om man ändrar variabeluppsättning och metod något, att datamaterialet är representativt, att bortfallet inte påverkar resultatet osv. Allt detta är naturligtvis extra viktigt om det är en betydelsefull samhällsfråga man undersöker och om man vill gå ut med ett otvetydigt resultat.

Betydelsen av denna robusthet kan illustreras med flera exempel i Bergström och Sandströms analys (se Bergström & Sandström [2001a, 2001b]). De använder t ex OLS på två nationella prov (A-rätt och B-rätt) med motiveringen att dessa är kontinuerliga² och att OLS därför är en godtagbar metod. Men nu finns det ytterligare fem nationella prov med kontinuerliga svar i databasen (i engelska), och om friskoleeffekten finns borde vi rimligen märka den också på dessa prov. Bergström och Sandström testar inte sin modell på dessa, men om de hade gjort detta skulle de funnit att friskoleeffekten genomgående blev statistiskt insignifikant, dvs att deras hypotes borde förkastas.³ För totalt sju nationella prov (med kontinuerliga provsvar) finns alltså en statistiskt signifikant effekt enbart för de två prov som Bergström och Sandström presenterar, även om vi använder exakt samma modell som författarna. Redan detta kastar vissa tvivel över deras resultat.

Ett annat och mer allvarligt exempel på godtycklighet rör Bergström och Sandströms val av förklarande variabler. Det finns många skilda variabler att tillgå i databasen, och det är mycket svårt att förstå författarnas kriterier vid urvalet. Det blir t ex oförklarat varför en del variabler används i ”individanalysen” men inte i ”kommunanalysen” och vice versa.⁴ En speciell omständighet är det faktum att medan det finns drygt 20 000 observatio-

² Svaren ges i poäng 1–45 respektive 1–30.

³ Av utrymmesskal visar jag inte resultaten för dessa skattningar. Intresserade läsare kan dock på begäran erhålla datautskriften från alla mina skattningar. Och givetvis kan jag ställa allt mitt datamaterial till den intresserades förfogande.

⁴ Det förklaras inte heller varför man aldrig beaktat vissa variabler som (när jag analyserat materialet) visar sig vara statistiskt stabila och relevanta i samtliga skattningar, t ex ”andelen pedagogiskt utbildad personal”.

Tabell 2 OLS-skattningar av A-rätt mot ett antal olika variabeluppsättningar. (21 103 observationer)

Förklarande variabler	Kön	Invandrarbakgrund	Andel friskolor
Bergström & Sandström	-1,06***	-0,738***	0,093***
+Kostnad/elev	-1,06***	-0,755***	0,066**
+Andel lågutbildade	-1,06***	-0,766***	0,038
+Andel utlänningar	-1,05***	-0,790***	0,025
+Andel högutbildade	-1,06***	-0,788***	0,010

Anmärkning: En, två respektive tre asterisker anger signifikansnivåerna på 10, 5 respektive 1 procentsnivå. Modellen har skattats med följande variabler: Kön, moderns och faderns utbildningsnivå, invandrarbakgrund, skattekonor per invånare, invånardistans, förort, samt (givetvis) andelen friskolor i kommunen.

ner avseende det som skall förklaras (dvs de individuella betygen/proven), finns blott 34 olika datapunkter rörande den viktigaste förklarande variabeln, nämligen "andel friskolor". Här finns troligen en stor del av förklaringen till författarnas friskoleeffekt. Det visar sig nämligen att effekten är beroende av vilka övriga kommunvariabler som tas med i analysen. Läger man till någon eller några (t ex "andel högutbildade") eller om man byter ut eller tar bort någon eller några så försvinner friskoleeffekten och den blir inte sällan negativ.

En mycket speciell betydelse har t ex måttet "invånardistans" som mäter det fysiska medelavståndet mellan invånarna i en kommun. Författarna inkluderar denna variabel med motiveringen att den skall uttrycka graden av konkurrens: i en tätort (alltså med liten "invånardistans") så finns mycket konkurrens mellan skolorna och omvänt i glesbygd.

Nu gäller emellertid att denna distansvariabel alltid uppträder med fel tecken i Bergström och Sandströms analyser. Den är positiv och oftast signifikant, något som enligt författarnas hypotes betyder att hårdare konkurrens ger lägre (!) betyg. Författarna har alltså funnit både att konkurrensen ger högre betyg (genom en ökad andel friskolor) och att den ger lägre betyg (genom minskad "invånardistans"). Det är ett resultat för mycket.

Användandet av invånardistansen kan

ifrågasättas. Tanken att betygen i landets skolor skulle variera linjärt med det fysiska avståndet mellan människorna i en kommun känns lite skruvad. Pajala har t ex 1 000 i invånardistans mot cirka 20 för en vanlig storstad, vilket skulle betyda att "distanskonkurrensen" är 50 gånger större i storstaden. Men även om medelavståndet mellan människor i Pajala kommun är mycket stort så betyder det ju inte att invånarna bor så glest. Måttet baseras på antagandet att kommunens invånare är jämnt fördelade över kommunens yta, men huvuddelen av Pajalaborna bor faktiskt i Pajala samhälle. Och om man startar en friskola i Pajala tätort så är det fysiska avståndet mellan den och den kommunala skolan kanske mindre än motsvarande avstånd i Stockholm.

För att se betydelsen av vilka variabler som inkluderas/exkluderas har jag gjort ett antal regressioner vars resultat visas i *Tabell 2* och *Tabell 3*. Det avser OLS-skattningar av matematikprov A-rätt, det prov som ger starkast resultat hos Bergström och Sandström. Jag har först skattat samma modell som Bergström och Sandström därefter lagt till en skolvariabel och slutligen ett antal kommunvariabler.

Notera först den stora stabiliteten i (estimaterna för) variablerna Kön och Invandrarbakgrund. Deras värde påverkas inte märkbart av vilka förklarande variabler som ingår. Men detta är inte fallet med skattningarna för andelen friskolor. Då yt-

Tabell 3 Friskoleeffekten testat med två olika variabeluppsättningar. Parameter-estimat (t-värde). (21 103 obs)

	A-rätt	B-rätt	Meritvärde
Bergström & Sandström	0,093 (3,26)	0,083 (3,67)	0,149 (1,19)
Invånardistans och Förort bytt mot Kostnad/elev och Andel högutbildade	0,0044 (0,14)	0,0097 (-0,40)	0,063 (0,47)

Anmärkning: Det jag kallar ”Bergström & Sandström” är min replikering av deras modell och data som de beskriver dem. Eftersom jag inte tillåtit använda deras databas (se nedan) har jag heller inte exakt lyckats replikera deras resultat.

terligare kommunvariabler läggs till så faller värdet och blir raskt insignifikant (och det kan också bli negativt, se nedan). Hade Bergström och Sandström testat sin modell med några fler kommunvariabler (som redan finns i databasen) så hade man alltså fått ett helt annorlunda resultat jämfört med vad man presenterar.

Det är inte bara antalet kommunvariabler som har betydelse utan även vilka variabler man väljer. Bergström och Sandström väljer tre: Antal skattekronor, invånardistans och en dummy för förorterna. De båda sistnämnda är dock problematiska i den meningen att skattningarna varierar kraftigt vad avser värde, tecken och signifikans då modell, data-mängd eller variabeluppsättning varierar något. Författarna hade därför goda skäl att välja andra variabler eller åtminstone testa en annan uppsättning. Detta görs i *Tabell 3*, där jag bytt ut ”Invånardistans” och ”Förort” mot kommunens kostnad per elev samt andelen högutbildade i kommunen.

Här framgår tydligt skattningarnas känslighet. Om man bara byter två dåligt motiverade variabler mot två andra (som lättare försvarar sin plats) så försvinner signifikansen helt och estimaten kan till och med bli negativa.

Den känslighet vad gäller skattningen av friskoleeffekten som visas i *Tabell 2* och *Tabell 3* gäller inte bara de prov/betyg som återges i tabellerna. *Tabell 2* kan

upprepas för andra prov (t ex B-rätt) och vad det gäller variabelbytet som illustreras i *Tabell 3* så har jag upprepat detta också för de rena betyg som ges, dvs C-betyg, PG-betyg, betyg i svenska samt betyg i engelska.⁵ Då blev friskoleeffekten faktiskt negativ för samtliga betyg utom matematik (där koefficienten blev positiv men insignifikant). Då vi kompletterar variabeluppsättningen ger en statistisk analys således visst stöd för åsikten att andelen friskolor inverkar negativt (!) på kommunala elevers betyg. Men en sådan tolkning är självfallet oberättigad. När skattningarna är så känsliga för variabelvalet är den enda rimliga slutsatsen att det inte går att dra någon säker slutsats.

4. En analys av kommunmaterialet

Då det gäller kommunmaterialet måste man först ställa sig frågan om val av tidsperiod. Författarna har tillgång till fem årgångar tvärsnittsdata, nämligen för åren 1992, 1994, 1995, 1996 och 1997. Det resultat man funnit kan kontrolleras genom en mängd olika kombinationer, t ex genom att skatta åren var för sig, genom att skatta två på varandra följande år, tre olika år etc. Dessutom kan man genomföra en mycket viktig och relevant skattning, nämligen

⁵ Dessa skattningar genomfördes med OLS.

Tabell 4 Antal elever i åk 9 år 1997 i olika kommungrupper samt antal elever i ESO:s undersökning

	Antal elever i åk 9	Antal elever i ESO	Representativitet, %
Storstäder	11 698	12 234 ⁷	100
Förortskommuner	14 255	2 065	14
Större städer	25 284	9 434	37
Medelstora städer	15 367	2 442	16
Industrikommuner	9 102	1 014	11
Landsbygdskommun	4 894	0	0
Glesbygdskommun	2 709	256	9
Öriga större kommuner	7 672	200	3
Öriga mindre kommuner	5 333	353	7
Totalt	96 314	27 998	29

Anmärkning: Kommungrupperna följer Kommunförbundets indelning.

med **förändringen** av respektive variabel. Författarnas slutsats är ju den, att en ökande andel friskolor – allt annat lika – leder till högre betyg. Då borde rimligen en statistisk analys av **förändringen** av medelbetyg i en kommun vara statistiskt korrelerad med **förändringen** i andelen friskolor.

Också valet av förklarande variabler väcker frågor. I individanalysen användes t ex variabeln ”Skattekronor” men inte ”Kostnad per elev”. Här har nu den förra utgått och ersatts med den senare. Allt finns i samma databas, och det är svårt att förstå varför en variabel som t ex ”Skattekronor” som är starkt positivt signifikant (och robust) i de flesta skattningar på individmaterialet nu inte finns med som förklarande variabel. Just variabeln ”Skattekronor” är betydelsefull även ur den synpunkten att den kan fånga andelen högutbildade i en kommun.

Jag genomförde ett stort antal regressio- ner på det material som Bergström och Sandström använde både för enskilda år och för hela eller delar av perioden. Jag skattade även modellen med ”förändring- ar” av respektive variabel som argument . Då framkom tre huvudresultat. Det första var att den statistiska signifikansen försvann (i nio fall av tio) för alla andra tids- perioder än den författarna analyserar, dvs då vi delade upp materialet år från år, två år i rad etc. Det andra resultatet var att signi-

fikansen alltid försvann (dvs även för den period som ESO-rapporten studerar) om vi lade till en variabel som på något sätt tog hänsyn till andelen högutbildade i kom- munen, t ex variabeln ”Skattekronor”. Och det tredje var att då vi använde förändring- ar som argument så försvann också alltid signifikansen. Hade författarna gjort såda- na kompletterande undersökningar (och presenterat dem) skulle det vetenskapliga underlaget till deras budskap blivit mer ny- anserat och resultatet mer tvetydigt.⁶

5. Datamaterialets representativitet

Att resultaten är robusta är ett krav på en trovärdig statistisk analys. Ett annat är att det material man arbetar med är represen- tativt. Men redan en enkel analys av data- materialet visade att det inte uppfyllde detta krav. Totalt fanns cirka 30 procent av landets elever (i åk 9) i databasen, men representationen varierade kraftigt med kommuntypen (se *Tabell 4*).

⁶ Dessa resultat redovisas av utrymmesskäl ej här. De kan erhållas från författaren på begäran.

⁷ Antalet i ESO:s undersökning är här större än vad som uppgivits vara totala antalet elever som avslutade åk 9 i grundskolan enligt Skol- verkets hemsida.

Tabell 5 Betydelsen av vissa kommuner. Resultat då data från angivna kommuner uteslutits för att göra materialet mer representativt (OLS-skattningar). Beroende variabel: Meritvärde. (Ca 21 200 obs i utgångsläget)

	Föräldrars utbildningsnivå	Kön	Andel friskolor
Bergström & Sandström	28,6***	19,8***	+0,418***
Ej Pajala	28,0***	19,8***	+0,011
Ej Pajala +Gbg	28,3***	20,1***	-0,31*
Ej Gbg + Upps.	29,3***	20,3***	-0,50**
Ej Sthlm,Link,Umä	28,5***	20,8***	+0,58***
Ej Gbg	29,8***	20,4***	-0,030
Ej Sthlm, Gbg	29,7***	21,2***	-0,12

Anmärkning: En, två respektive tre asterisker anger signifikansnivåer på 10, 5 respektive 1 procentsnivån.

Från *Tabell 4* framgår tydligt att materialet är kraftigt överrepresenterat, framförallt vad avser storstäder och större städer. Författarna påstår sig i rapporten delvis ta hänsyn till detta genom att i sina regressioner införa en s k dummyvariabel för "Storstäder och förorter" men som påpekats avser den dummin enbart förorter.

Storstädernas överrepresentation har avgörande betydelse för resultatet. Totalt ingår endast 34 kommuner i databasen och det är naturligtvis särskilt viktigt att dessa kommuner är representativa, och att resultaten inte ändras om vi t ex plockar bort en eller två kommuner ur datasetet. Men här är resultaten tvärtom mycket känsliga för vilka kommuner som ingår. I *Tabell 5* visas parameterskattningar då samlingsbetyget "Meritvärde" används som beroende variabel. För varje skattning har vi uteslutit en eller flera kommuner. Som en jämförelse visar vi parameterestimatet för två andra variabler, nämligen föräldrarnas utbildningsnivå och individens kön.⁸

Observera först att parameterskattningarna vad avser variablerna för föräldrarnas utbildningsnivå och kön är mycket stabila oberoende om vi tar bort en kommun eller två, vilket är rimligt om det är en trovärdig effekt. Om "Föräldrars utbildningsnivå" har en viss effekt i Linköping så borde denna effekt vara densamma i Härnösand.

Men det motsatta är fallet för parametern för "Andel friskolor". Där behöver vi endast ta bort ett hundratal observationer i Pajala (av totalt drygt 20 000) för att skattningen skall falla till en bråkdel av det tidigare värdet och förlora all statistisk signifikans. Genom att ta bort en eller ett par kommuner så kan parameterskattningen för friskolevariabeln variera mellan signifikant positiv till signifikant negativ. Att resultatet är så tydligt beroende av exakt vilka kommuner som studeras avslöjar att det inte är robust och stärker åter misstanken att Bergström och Sandströms friskoleeffekt är en statistisk villa. Om det verkligen fanns ett samband mellan friskolor och goda betyg så borde rimligtvis detta inte påverkas av om Pajala eller Göteborg ingick i urvalet.⁹

⁸ För fullständiga resultat, kontakta författaren. Samma variabler som i Bergströms och Sandströms analys har använts, med den skillnaden att den korrekta variabeln för "Storstad" använts.

⁹ I brev till mig har Sandström påpekat att han efter min kritik genomfört 24 068 regressioner med "B-rätt" och "Meritvärdet" som beroende och där han tagit bort en, två och tre kommuner. I endast ett fall blev friskoleparametern negativ och signifikant. Jag kan inte närmare kommentera detta, men de resultat som jag återger ovan är hur som helst helt korrekta.

Tabell 6 Enkel korrelation mellan betyget "Meritvärde" och "Andel friskolor" för elever med föräldrar med olika utbildningsnivå

	Föräldrarnas utbildningsnivå		
	Förgymnasial	Gymnasial	Eftergymnasial
Korrelation	-0,0144	0,0247	0,0428

6. Storstäder och andel högutbildade

Att storstäderna är överrepresenterade har troligen stor betydelse för det statistiska resultatet. Det hänger samman med att korrelationen mellan betyg och andel friskolor varierar kraftigt med föräldrarnas utbildning. Det visar sig nämligen att om det överhuvud finns något samband mellan friskolor och betyg så gäller det bara för de mer välutbildade grupperna. Förhållandet kan illustreras med de enkla korrelationerna i *Tabell 6*.

Här framgår tydligt den stora skillnaden som finns mellan olika utbildningsgrupper. Dessa skillnader går också igen om mer sofistikerad regressionsanalys görs. Jag genomförde t ex totalt 36 olika regressioner (OLS) med den "förgymnasiala" gruppen (ca 2 500 obs), med olika variabeluppsättningar, men i alla fallen blev friskoleeffekten negativ och i flera statistiskt signifikant. Om vi alltså tror på den modell och typ av analys som Bergström och Sandström stöder sig på så är friskolornas effekt positiv för högutbildade hushåll men negativ för (barn till) de lågutbildade. En ökande andel friskolor skulle med andra ord endast gynna elever som kommer från välutbildade hem och missgynna barn till lågutbildade; alltså förstärka segregationen i samhället!

Bergström och Sandström gör en stor poäng av att friskolor är positiva för alla barn. Man skriver t ex på flera ställen, såväl i ESO-rapporten som i *Ekonomisk Debatt*-artikeln, att den positiva friskoleeffekten "gäller även om man studerar endast de svagaste eleverna i de kom-

munala skolorna". Detta är ett politiskt viktigt påstående eftersom kritiken mot friskolorna främst riktat in sig på effekten på elever från hem med låg studievana. Nu förklarar Bergström och Sandström inte närmare hur de nått sin slutsats och i själva verket finns i deras arbeten inget som helst belägg för påståendet. De delar inte upp materialet på olika utbildningsklasser; "friskoleeffekten" är i deras modeller lika för alla barn oberoende av betyg eller härkomst. Bergström och Sandströms påstående om just de svaga eleverna saknar inte bara grund; som ovanstående resultat visar finns det tvärtom goda skäl att tro att friskoleeffekten – om den nu över huvud taget finns – är radikalt olika för barn med olika klass- och utbildningsbakgrund.¹⁰

Om man har de olika korrelationer som visas i *Tabell 6* så förstår man att korrelationen för hela gruppen beror på hur den är sammansatt. Det är här som storstädernas tyngd blir avgörande eftersom andelen högutbildade är långt större i storstäder och större städer än i övriga riket. I hela riket är andelen med eftergymnasial utbildning ungefär 25 procent, men i

¹⁰ Bergström och Sandström kan med talet om "de svagaste eleverna" mena att de genomför ett test avseende vilka elever som blir godkända (i alla ämnen) eller ej. Deras resultat (ESO [2001, s 92]) visar där en insignifikant men positiv effekt för andelen friskolor. Men detta resultat är också beroende av ett mycket speciellt urval av förklarande variabler. Om man byter variabler som i *Tabell 3* så erhåller vi en **negativ** effekt på andelen godkända elever (probit-analys).

Stockholm är den 41, i Göteborg 34 och i t ex Lund 55 procent.

Överrepresentationen av storstäder medför alltså samtidigt att de högutbildade blir överrepresenterade.¹¹ I åldrarna 35–55 år (rimliga värden för föräldrar till barn i åk 9) och för år 1998 är fördelningen mellan de tre utbildningsgrupperna i Sverige ungefär 25-50-25 (procent).¹² I datamaterialet är emellertid fördelningen för den utbildningsvariabel som används ungefär 10-45-45.¹³

För att kontrollera för detta konstruerade jag ett tiotal undergrupper (var och en med ca 4 500 observationer) där jag viktade om observationerna.¹⁴ Därefter gjorde jag regressioner på samma sätt som görs av Bergström och Sandström. Som jämförelse tittar jag på parameterskattningarna för två andra variabler: Föräldrarnas utbildningsnivå och elevens kön. Skattningarna för variablerna "Kön" och "Föräldrarnas utbildningsnivå" var nivåmässigt stabila och signifikanta och ungefär av samma storlek som då vi skattade för hela materialet. Det är också vad vi borde förvänta oss om estimaten återspeglar ett verkligt orsakssamband. Vad beträffar parametern för "Andel friskolor" så stärker resultaten inte Bergströms och Sandströms sak. Skattningarna gjordes med de tre kontinuerliga variablerna A-rätt, B-rätt samt Meritvärde. Ungefär en tredjedel av skattningarna var negativa (alla för variabeln "B-rätt" var dock positiva), knappt någon var signifikant och skattningens värden fluktuerade kraftigt. Dessutom fanns flera signifikant negativa skattningar, och detta trots att jag arbetat med samma förklarande variabler som i ESO-studien.¹⁵

Resultaten ovan ger näring åt misstanken att det samband som Bergström och Sandström tror sig ha funnit är något annat, nämligen det att det i storstäder finns både gott om högutbildade (som har barn med höga betyg) och en stor andel friskolor. Detta samband – som med stor sannolikhet är korrekt – kan lätt spilla över i en

statistisk korrelation som inte är ett orsakssamband utan en statistisk villa. Ett "riktigt" orsakssamband skulle inte gett dessa instabila resultat.

7. Antalet observationer

På ett flertal ställen anger Bergström och Sandström att de arbetar med data som omfattar ca 30 000 elever (se t ex ESO [2001, s 3, 66, 119] och Bergström & Sandström [2001a, s 324]). Nu är det i och för sig korrekt att det antal poster som ingår i databasen är ca 30 000. Men författarna undviker sorgfälligt att berätta att långt ifrån alla dessa poster ingår i de statistiska analyserna. Data saknas nämligen i stor omfattning så den databas som Bergström och Sandström använder i de statistiska analyserna omfattar maximalt 23 000 observationer, oftast ännu lägre, kanske drygt 20 000. Bortfallet är alltså

¹¹ De högutbildade finns i långt högre grad i friskolorna. Elever från hem med högsta utbildningsgruppen utgör 68 procent av eleverna i friskolorna mot 48 procent i den kommunala skolan (i dataurvalet). Elever från hem med den lägsta utbildningen utgör 11 procent i den kommunala skolan mot 5 procent i friskolorna.

¹² Se Statistisk Åbok "Utbildning och Forskning". De exakta siffrorna för män var 26, 52, 22 och för kvinnor 23, 64 och 14 procent. (Här räknas gymnasial utbildning som 3 år).

¹³ Den exakta fördelningen beror på exakt vilka data som används (och det går inte att få reda på), men detta är värdena för det dataset som Bergström och Sandström troligen använder sig av.

¹⁴ Viktningen gick till så att jag tog varje observation från den lågutbildade gruppen, varannan från "medelklassen" och var tredje från de högutbildade. Därefter togs var femte observation från storstäderna och de större städerna och samtliga från övriga kommuntyper.

¹⁵ På grund av utrymmesbrist publiceras ej dessa resultat, men de kan erhållas från författaren.

Tabell 7 Bortfallsanalys

	Antal observationer	Medelbetyg matematik	Medeltal meritvärde	Medeltal andel friskolor
Hela databasen	29 336	2,45	199	5,15
Bortfallet	5 539	2,12	146	5,90
Årstående i databasen	23 797	2,52	210	4,98

minst cirka 20 procent. I strid med all praxis vad avser statistisk analys nämns inte detta bortfall på något ställe, och man anger heller inte i någon av de många skattningarna hur många observationer man egentligen använt.

Ett bortfall på 20–25 procent behöver inte betyda att en statistisk analys leder vilse. Om bortfallet har en fördelning som överensstämmer med hela databasen så ändras inte den statistiska analysen. Men om bortfallet är skevt så blir den statistiska analysen felaktig. Om man i en undersökning t ex plockar bort de bästa eleverna från en skola så kommer denna skola att framstå som speciellt dålig, eftersom bara de sämre eleverna är kvar i urvalet.

En noggrann analys av det aktuella bortfallet visar att det finns en klar systematik. Dels gäller att bortfallet omfattar elever med låga betyg, dels gäller att det i högre grad kommer från kommuner med mycket friskolor. Betydelsen av detta är uppenbar. Låt oss som exempel tänka oss att i kommuner med 0 procent friskolor så ingår alla elever i undersökningen. I kommuner med 1 procent friskolor så tar vi bort 1 procent av de sämsta betygen, i kommuner med 2 procent friskolor tar vi bort 2 procent av de sämsta betygen osv. Gör vi på detta sätt får vi naturligtvis en positiv statistisk korrelation mellan andelen friskolor och goda betyg.

Det aktuella bortfallets betydelse för två betyg (matematik, meritvärde) framgår av *Tabell 7*.

I *Tabell 7* framgår tydligt hur bortfallet har påverkat databasen. Betygen har höjts och andelen friskolor sänkts. Det betyder

att vi tagit bort observationer (elever) med svaga betyg och som bor i kommuner med en stor andel friskolor. Denna systematiska rensning har sänkt den genomsnittliga andelen friskolor till 4,98 procent och samtidigt höjt snittbetygen, i matematik t ex med 0,07 enheter och för meritvärdet med 11 enheter.

Nu kan databortfallet uppkomma av flera orsaker: data saknas t ex om föräldrarnas utbildning, eleven har inte deltagit i det aktuella provet eller också har han/hon deltagit, men av olika anledningar så har resultatet inte rapporterats till Skolverket. En intressant iakttagelse är den starka korrelationen mellan bortfall och andelen friskolor. Jag utförde en sk probit-analys med "Rapporterat deltagande i prov" (ca 25 000 obs) som den beroende variabeln (1 = resultat finns, 0 = resultat finns ej). Analysen utfördes för samtliga nationella matematikprov och var samtliga negativa och signifikanta på högsta nivå. Det betyder att det finns ett statistiskt säkerställt samband och tolkningen är att om det finns en hög andel friskolor i kommunen så har man en klart högre andel elever där resultat från de nationella proven inte finns. Fler friskolor innebär alltså fler bortfall av prov om vi skall tro statistiken. Initiellt avfärdade jag detta resultat som ytterligare ett fall av statistisk tillfällighet. Men en kollega med starka band till skolvärlden ansåg att det fanns en rationell förklaring, nämligen den att man inte sällan underlätit att sända in dåliga provresultat och att presen att handla så ökade då konkurrensen om eleverna hårdnade. Konkurrensen från friskolorna skulle alltså leda till att dåliga

provresultat gömdes undan. Att man då kan finna en statistiskt positiv "friskoleeffekt" är inte svårförståeligt, speciellt om man kan välja variabler och modeller med viss urskiljning. Men förklaringen till effekten är kanske en helt annan än den som Bergström och Sandström tillhandahåller.

8. Sammanfattning och lärdomar

Den granskning av Bergströms och Sandströms arbete som jag här redovisat är naturligtvis inte invändningsfri: Jag har t ex inte testat alla de olika modeller och estimeringsmetoder som Bergström och Sandström använder.¹⁶ Men min slutsats är ändå att friskolestudien innehåller allvarliga statistiska brister, nämligen:

- 1) Modellerna är selektiva. Om man lägger till, tar bort eller byter ut någon eller några av de förklarande variablerna så försvinner den statistiska signifikansen och sambandet kan lika gärna bli negativt (och statistiskt signifikant).
- 2) Datamaterialet är inte representativt för Sveriges skolor. Det finns en stor överrepresentation av storstäder och elever från hem med hög utbildning. Om hänsyn tas till detta försvinner det statistiska sambandet.
- 3) Det finns ett stort och systematiskt bortfall som inte nämns och som i extremfall uppgår till cirka 40 procent. Om hänsyn tas till detta försvinner det statistiska sambandet. Bortfallet är starkt korrelerat med andelen friskolor i en kommun.

Var och en av de tre sakerna tagna för sig är tillräckligt för att dra slutsatsen att Bergströms och Sandströms positiva friskoleeffekt saknar ordentligt vetenskapligt underlag. Tagna tillsammans räcker de mer än väl för att hävda att Bergström och Sandström inte kan visa någon positiv effekt på den kommunala skolan av ut-

byggnaden av friskolorna i landet. Det enda samband som jag funnit som gäller nästan oberoende av modellval är att bortfallet av provsvar är positivt korrelerat med andelen friskolor i kommunen. Till detta kan läggas att jag kontrollerat min analys med material från åren 1999 och 2000 och inte funnit något som får mig att ändra min slutsats: friskolorna kan med det föreliggande statistiska materialet inte visas ha haft någon positiv effekt på betyg/prov i den kommunala skolan.

Det finns två lärdomar att dra av denna historia. Den första är att man måste granska data i detalj innan man kan avgöra om en statistisk studie är trovärdig eller ej. Det kan finnas avgörande fel även om ett arbete är författat av kompetenta forskare, även om det är publicerat i en vetenskaplig (eller populärvetenskaplig) tidskrift, även om det är framlagt vid en stor vetenskaplig konferens. Alla som sysslar med forskning vet att granskningsprocessen sällan går så djupt som att ekonometriska skattningar görs om, att granskaren testar alternativa modeller eller att han/hon analyserar bortfallet. ~~Än~~ då det gäller de mest välrenommerade vetenskapliga tidskrifterna är det sällsynt att bedömarna själva detaljstuderar det underliggande datamaterialet.

Av denna anledning är det extra viktigt att iakta regeln att hålla data omedelbart tillgängliga så att andra forskare lätt skall kunna granska resultaten. Men här finns Bergströms och Sandströms svåraste för syndelse. Redan då rapporten publicerades frågade jag efter författarnas rådata, men fick beskedet att jag inte fick tillgång till dem, även om jag skaffade nödvändiga tillstånd (från Skolverket).¹⁷ Först efter

¹⁶ I en brevkommentar till min kritik menar författarna t ex att det är en stor brist att jag inte testat deras valmodeller.

¹⁷ Detta är en allvarlig beskyllning, men jag kan givetvis styrka mitt påstående genom bevarad e-post som jag gärna ställer till den intresserades förfogande.

ett mycket omfattande och tidsödande arbete från min sida kunde jag på annat sätt (via SCB och Skolverket) få tillgång till den databas som författarna använder sig av. Detta är oacceptabelt och anledningen till att denna granskning kommer mer än ett halvår efter det att Bergström och Sandström publicerat sin artikel i *Ekonomisk Debatt*.¹⁸

En andra slutsats är att man bör iaktta tillbörlig försiktighet då man i stor stil torgför politiskt viktiga resultat som objektiv vetenskap. Ingen förmenar Bergström och Sandström rätten att skriva på DN Debatt, men om man där hävdar att man talar enbart i Vetenskapens namn, utan politiska bitoner, ska man ha ordentligt på fötterna. Om man stödjer sig på ekonometri så bör man framhålla att ingen signifikansnivå innebär en "säkerhet", allt handlar om en glidande skala av sannolikheter där den visshet som teoretiska och rent matematiska analyser ger aldrig uppnås. Att få en statistisk signifikant korrelation är inte svårt, det svåra (och vetenskapliga) består i att kunna visa att det man funnit representerar ett verkligt underliggande orsakssamband.

Referenser

- Bergström, F & Sandström, F M, [2001a], "Konkurrens mellan skolor – för barnens skull!", *Ekonomisk Debatt*, årg 29, nr 5, s 319–331.
- Bergström, F & Sandström, F M, [2001b], "Friskolor gynnar alla barn", *Dagens Nyheter* (DN Debatt), 8 mars 2001.
- ESO, [2001], "Konkurrens bildar skola", Finansdepartementet, Ds 2001:12.

¹⁸ Författarnas vägran att lämna ut sina rådata har även den konsekvensen att jag inte exakt lyckats reproducera deras resultat eftersom jag inte lyckats utröna exakt vilka data de använder. Jag har fått försäkringar från Skolverket om att den databas jag fått är densamma men jag vet inte exakt hur Bergström och Sandström valt ut data ur hela databasen, hur de hanterat observationer med ofullständiga uppgifter osv. Jag har frågat Bergström och Sandström om exakt vilka data de egentligen använder, men inte fått något svar. Det är möjligt att de baserar sina statistiska analyser på något "filtrerat" material, eller att de på något sätt konstruerat nya data för att fylla observationer där uppgifterna är ofullständiga. Utan tillgång till deras exakta datauppsättning (och när de inte svarar på frågor) är det omöjligt att säkert veta hur de gått tillväga.