Artikel

Högre utbildning
Vol. 2, Nr. 1, 2012, 33-46

Är förkunskaper verkligen viktiga för framgångsrika högre studier? En analys av studieresultatet på en grundläggande högskolekurs i statistik

Magnus Carlsson*
Ekonomihögskolan, Linnéuniversitetet

Studien undersöker vilken betydelse goda gymnasiala förkunskaper i matematik har för studieresultatet på en kurs i grundläggande statistik på högskolenivå. Särskilt fokus är på betydelsen av studenternas gymnasiala förkunskaper 1) i sådan matematik som ingår i behörighetskraven och 2) i form av extra kurser matematik som lästs utöver behörighetskraven. Data erhålls genom att på individnivå länka samman tentamensresultat på en statistikkurs med uppgifter kring studenternas gymnasiala meriter i ämnet matematik. Något förvånande – och i kontrast till tidigare studier – är det inget i resultaten som indikerar att extra kurser i matematik utöver behörighetskraven har betydelse för studieresultatet på statistikkursen. Istället tycks det vara vilket betyg studenterna har på de mer elementära matematikkurserna som ingår i behörighetskraven som är viktigt. En tolkning av resultaten är att en grundläggande högskolekurs i statistik inte kräver särskilt mycket matematiskt kunnande, utan att det är allmänna kognitiva förmågor som intuition, logiskt tänkande och problemlösningsförmåga som är viktiga och att betygen på de mest elementära matematikkurserna på gymnasiet speglar sådana förmågor.

Nyckelord: Statistik, matematik, förkunskaper, betyg, gymnasiala studier, högre studier, studieresultat, tentamensresultat

INLEDNING

Statistik på högskolenivå är ett matematiskt ämne och det ligger därför nära till hands att anta att gymnasiala förkunskaper i matematik är viktigt för studieresultatet i statistik. Universiteten och högskolorna i Sverige har alla liknande antagningskrav vad gäller antalet kurser gymnasiematematik som krävs för att antas till en grundläggande kurs i statistik.[1] Trots att alla studenter naturligtvis uppfyller de formella behörighetskraven finns det anekdotisk evidens för att statistikstudenternas förkunskaper i matematik ändå varierar mellan lärosäten.[2] Det tycks också finnas en motsvarande variation vad gäller studieresultaten. Syftet med denna studie är att mer noggrant undersöka betydelsen av förkunskaper i matematik för prestationen på en kurs i grundläggande statistik på högskolenivå. Idén till undersökningen uppstod när kursledningen på en grundläggande statistik-

1 Gymnasiekursen Matematik C tycks vara det vanligaste behörighetskravet på den första grundläggande kursen i statistik (en sådan kurs benämns typiskt Statistik A och ges av en institution för statistik). Det finns även en mängd specialkurser i grundläggande statistik som riktar sig till ekonomer, psykologer etc. Här är behörighetskravet i vissa fall istället Matematik B.

2 Detta är baserat på samtal med några universitetslektorer som undervisat i statistik vid Uppsala universitet, Lunds universitet och Linnéuniversitetet.

*Författarkontakt: magnus.carlsson@lnu.se

ISSN 2000-7558

http://www.hogreutbildning.se

Tidskriften tillämpar kollegial granskning för bidrag av typen ”Artikel”. Övriga bidrag granskas redaktionellt. För mer information hänvisas till http://hogreutbildning.se/page/om-tidskriften

kurs vid Linnéuniversitet observerade att en betydande andel av studenterna på kursen inte tycktes nå upp till de kunskapsmål som angavs i kursplanen. Man misstänkte att detta till viss del berodde på bristfälliga gymnasiala förkunskaper i matematik bland studenterna.[3] Av den anledningen genomfördes en intervention innan kursen i statistik påbörjades i form av en tvåveckors frivillig matematikkurs som syftade till att repetera grundläggande gymnasiematematik. Dock var det inte möjligt att utvärdera interventionen eftersom en rad andra faktorer ändrades samtidigt på kursen. Därför kommer en annan datakälla att användas i denna uppsats i syfte att analysera betydelsen av förkunskaper i gymnasiematematik för studieresultatet på en grundläggande kurs i statistik.

Tidigare studier som utvärderat studieresultat på statistikkurser har främst fokuserat på andra faktorer än förkunskaper i matematik. T.ex. har man analyserat betydelsen av användandet av skriftliga tolkningar av statistiska resultat (Garfield, Hogg, Shau & Whittinghill, 2002; Magel, 1996; Stromberg & Ramanathan, 1996; Utts, Sommer, Acredolo, Jaher & Matthews, 2003), närvaro på kursen (Magel, 1996), aktiv problemlösning (Hillmer, 1996), webbaserad inlärning (Ward, 2004) och examinering (Krieg & Uyar, 2001).[4]

Det finns även några studier som fokuserar på betydelsen av just matematiska förkunskaper. Woodward & Galagedera (2006) analyserar studieresultaten på en grundläggande statistikkurs som gavs på ett universitet i Sri Lanka i slutet av 90-talet. Här finner man att inneboende förmåga, arbete och motivation är de enda av de undersökta variablerna som signifikant påverkar studenternas resultat. Man drar slutsatsen att statistiskt tänkande på grundläggande nivå främst är intuitivt och icke-matematiskt och att man därför kan tona ned behovet av förkunskaper i matematik. Green, Stone, Zegeye & Charles (2007) utnyttjar en förändring i behörighetskraven i mängden matematik som ska ha lästs för att se hur det påverkar studieresultaten på en kurs i statistik. Deras strategi baseras på att det skulle vara slumpmässigt vilka studenter som påverkades av förändringen i behörighetskraven och att det därför är möjligt att mäta effekten som matematikkunskaper har på studieresultatet. Resultaten visar att mängden matematik studenterna läst före kursen i statistik har en positiv påverkan på studieresultatet. Även Cohn (1972) utvärderar hur antalet kurspoäng matematik en student läst innan statistikkursen är relaterat till studieresultatet. Författaren finner att det finns en positiv association mellan antalet kurspoäng i matematik som studerats och resultatet på statistikkursen.[5]

Liksom en del av de tidigare studierna är ett syfte med nuvarande studie att undersöka vilken betydelse förkunskaper i matematik har för prestationen på en kurs i statistik. Bidraget till den befintliga litteraturen är att denna studie även försöker särskilja betydelsen av hur väl man behärskar matematiken som ingår i behörighetskraven från betydelsen av den mängd matematik som studerats utöver behörighetskraven. Att kunna särskilja mellan dessa två faktorer har viktiga implikationer för hur en förändring av behörighetskraven i matematik mest effektivt

3 Niss (1994) listar i ett ämnesdidaktiskt ramverk för matematik tre problem som en lärare måste beakta i undervisningen. Ett problem är huruvida ett område är möjligt att lära ut, t.ex. som i detta fall beroende på studenternas förkunskaper. De andra två problemen är 1) att läraren bör kunna motivera inför studenterna vilka områden som ska tas upp på kursen och 2) hur undervisningen ska implementeras; hur ska material förberedes och vilka resurser ska finnas till studenternas förfogande givet begränsningar i högskolesystemet, lärarens förmåga etc.

4 Några andra studier som behandlar olika aspekter kring studieresultat på grundläggande statistikkurser på högskolenivå är Becker (1998), Carlson (1999), Hakeem (2001) och Krieg & Uyar (1997).

5 En annan studie är Johnson & Kuennen (2006) som undersöker vilka typer av matematiska färdigheter som är associerade med framgångsrika studier på en kurs i statistik. Dock tittar de inte specifikt på mängden matematik som lästs eller specifikt på matematikbetyg.

kan öka andelen studenter som lever upp till kunskapsmålen: antingen genom att öka antalet matematikkurser i behörighetskraven eller genom att kräva ett högre betyg på de matematikkurser som ingår i nuvarande behörighetskrav. Det har också implikationer för huruvida man kan förvänta sig att en intervention i statistikundervisning liknande den som nämndes ovan är ett effektivt sätt att få fler studenter att nå de uppsatta kunskapsmålen.

Unika och detaljerade data kring dels studenternas betyg på gymnasiets matematikkurser dels antalet matematikkurser som studerats på gymnasiet utnyttjas för att separera mellan de två effekterna av intresse. Det faktum att samtliga studenter på statistikkursen läst de mest elementära behörighetsgivande matematikkurserna gör att det inte finns någon selektionsproblematik kring vilka studenter som valt att läsa dessa kurser. Därför bör man utan risk för selektionsbias kunna mäta effekten på studieresultatet på statistikkursen av gymnasiebetyget på de behörighetsgivande matematikkurserna.

Gruppen studenter som valt att läsa fler kurser gymnasiematematik är å andra sidan troligen ett selektivt urval. Anledningen är att det förmodligen inte är slumpmässigt vilka individer som väljer att studera fler kurser gymnasiematematik. Det kan t.ex. vara individer med hög matematisk förmåga[6] som väljer att läsa ytterligare kurser. Om en sådan selektion förekommer finns det risk för att felaktigt dra slutsatsen att antalet kurser man studerat utöver de behörighetsgivande har en positiv inverkan på prestationen, fast förklaringen i själva verket är att studenter med hög matematisk förmåga läser fler kurser i matematik. Detta potentiella problem skulle dock kunna kringgås ifall det i analysen vore möjligt att jämföra individer med lika allmän matematisk förmåga. Tack vare att samtliga statistikstudenter läst de mest grundläggande kurserna i matematik och att betygen finns tillgängliga kan en proxyvariabel för matematisk förmåga skapas med hjälp av betygen på dessa kurser. Tanken är att man då kan skatta effekten av antalet studerade kurser baserat på studenter med *lika* allmän matematisk förmåga genom att konstanthålla betyget på de mest elementära gymnasiekurserna i matematik.

Resultaten visar något förvånande – och i motsats till Green, Stone, Zegeye & Charles (2007) – att antalet kurser i matematik som studerats utöver behörighetskraven inte tycks ha betydelse för studieresultatet på en grundläggande kurs i statistik. Istället tycks det vara hur väl studenterna behärskar innehållet på de matematikkurser som ingår i behörighetskraven som är viktigt. En tolkning är att en grundläggande högskolekurs i statistik inte kräver särskilt mycket matematiskt kunnande, utan att det är allmänna kognitiva förmågor som intuition, logiskt tänkande och problemlösningsförmåga som är viktiga. En implikation av resultaten tycks vara att om ett universitet eller en högskola önskar öka andelen godkända studenter på en grundläggande kurs i statistik, är det effektivast att anta studenter med högre betyg på de elementära matematikkurserna på gymnasiet, och inte att öka antalet kurser gymnasiematematik som ingår i behörighetskraven. Inte heller tycks en intervention i statistikundervisning i form av repetition av grundläggande gymnasiematematik vara ett effektivt sätt att förbättra studieresultaten.

Nästa avsnitt presenterar datamaterialet som används och därefter följer den empiriska analysen. Sist i uppsatsen återfinns en diskussion av resultaten samt några slutsatser.

---

6 Krutetskii (1976) beskriver matematisk förmåga med en rad olika förmågor i en omfattande longitudinell studie av barn och vuxna. Wistedt & Lagergren m.fl. (2006) sammanfattar dessa förmågor i tre kategorier: 1) ”Insamla matematisk information – förmåga att tänka matematiskt och fånga den formella strukturen i ett problem” (sid. 16), 2) ”Bearbeta matematisk information – förmåga att tänka logiskt och förstå matematiska symboler; förmåga att generalisera matematiska objekt, relationer och operationer; förmåga till ett flexibelt tankesätt där man lätt växlar mellan strategier och representationer; förmåga att förkorta och förenkla matematiska resonemang och operationer” (sid 16) och 3) ”Bevara matematisk information – förmåga att minnas matematiska relationer och metoder för problemlösning” (sid 17.)

DATA

Statistikkursen vars tentamensresultat analyseras behandlar grundläggande statistik inom området statistisk teori. Kursen gavs med samma struktur vid två tillfällen år 2010 under nuvarande kursledning.[7] Av 140 studenter som påbörjade kursen detta år var 53 studenter från Marknadsföringsprogrammet och 87 studenter från andra program inom ämnet företagsekonomi.[8] Examinationen på kursen utgjordes av tentamen – vars resultat kommer att analyseras – samt en obligatorisk datalaboration. På tentamen var tonvikten ungefär till lika stor del på statistiska beräkningar som på intuition och tolkning av resultat.

Beroendevariabeln som kommer att användas i analysen är en kontinuerlig variabel och utgörs av tentamensresultatet som finns tillgängligt för 110 studenter – de studenter som skrev tentamen.[9] Maximal poäng på tentamen var 32 poäng och betygsgränserna för godkänt och väl godkänt resultat var 19 respektive 25 poäng. Medlevärdet för tentamensresultatet var cirka 22 poäng.

De oberoende variablerna av intresse är antalet kurser i matematik på gymnasienivå som studenterna läst och betyg på matematikkurserna som utgör behörighetskraven. Båda variabler konstrueras utifrån de uppgifter som finns tillängliga i det nationella antagningsystemet NyA. Där finns uppgifter kring studenternas samtliga meriter som ligger till grund för antagningen till universitetet.

Den första oberoende variabeln är antalet gymnasiekurser i matematik. I det nuvarande gymnasiala betygssystemet finns det i princip fem matematikkurser, vilka betecknas Matematik A-E. Kurserna måste läsas i ordning och varje avklarad kurs motsvarar ett visst antal gymnasiepoäng. Behörighetskravet för programstudenter på Marknadsföringsprogrammet är att ha läst Matematik B. Det innebär att samtliga marknadsföringsstudenter på statistikkursen hade läst minst t.o.m. Matematik B. Behörighetskravet för övriga programstudenter är att ha läst Matematik C och således hade samtliga dessa studenter läst minst t.o.m. Matematik C.[10] För enkelhetens skull konstrueras den oberoende variabeln som mäter antalet matematikkurser som en diskret variabel som antar heltalen i intervallet 0-4, där 0 motsvarar Matematik A osv.[11]

Matematik A är den mest elementära matematikkursen på gymnasiet. Här behandlas på en grundläggande nivå bl.a. geometriska begrepp, statistiska data, algebraiska uttryck, formler, funktioner, linjära ekvationer och enkla potensekvationer.[12] På Matematik B läser eleverna t.ex. klassisk geometri, sannolikheter vid slumpförsök, lägesmått för statistiska material, spridningsmått, statistiska undersökningar, andragradsekvationer, räta linjens ekvation i olika former, linjära olikheter, ekvationssystem och algebraiska metoder.

---

7 Kursen var i själva verket uppdelad i två delar: deskriptiv statistik och statistik teori. Här analyseras tentamensresultatet på statistisk teori. Skälet är att examinationen kring deskriptiv statistik var i form av inlämningsuppgifter och en hemtenta. Nästan samtliga studenter godkändes på deskriptiv statistik, varför det knappt finns någon variation i beroendevariabeln i detta fall.

8 Även ett fåtal studenter läste kursen som fristående kurs vid detta tillfälle.

9 Fokus i uppsatsen är på det första tentamenstillfället som gavs.

10 I analysen kommer hänsyn att tas till det faktum att det kan finnas icke-observerbara variabler som är systematiskt korrelerade med egenskapen att vara marknadsförare – studenter som i genomsnitt läst mindre gymnasiematematik – och som samtidigt bestämmer tentamensresultatet. Hänsyn till detta tas genom att i analysen inkludera en kontrollvariabel i form av en dummyvariabel som indikerar ifall studenten är marknadsförare eller ej (en s.k. fix effekt). Utan en sådan kontrollvariabel riskerar man att få ett systematiskt fel i skattningen.

11 En sådan konstruktion med nollan som minsta värde gör konstanten i regressionsanalysen tolkningsbar.

12 Samtliga uppgifter kring kursplanerna för Matematik A-E är hämtade från skolverkets hemsida: se http://www.skolverket.se

Av samtliga 140 studenter hade 73 studenter även läst Matematik C (förutom Matematik A och B). På Matematik C studerar eleverna områden som logaritmer, potenser med reella exponenter, polynom, polynomekvationer av högre grad, faktorisering, matematiska modeller, ändringskvot, derivata, egenskaper hos funktioner, deriveringsregler, talet e, slutsatser om en funktions derivata och sambandet mellan en funktions graf och dess derivata.

Elva och åtta statistikstudenter hade även läst Matematik D respektive E. Matematik D behandlar enhetscirkeln, trigonometriska begrepp, trigonometriska ekvationer och dess lösningar, grafer till trigonometriska funktioner, periodiska förlopp, beräkningar av sidor och vinklar i en godtycklig triangel, härledning av deriveringsregler för trigonometriska funktioner, logaritmfunktioner, sammansatta funktioner samt produkt och kvot av funktioner, andraderivata, numerisk ekvationslösning, differentialekvation, integraler, sambandet mellan integral och derivata och numerisk integration m.m. Kursplanen för den mest avancerade gymnasiekursen i matematik – Matematik E – innehåller bl.a. komplexa tal, polynomekvationer med komplexa rötter, lösning av problem som kräver bestämning av derivator och integraler, beräkning av volymer med hjälp av integraler, differentialekvationer som modeller för verkliga situationer, exakta lösningar till några enkla differentialekvationer och numeriska lösningar till differentialekvationer.

Den andra oberoende variabeln är betyget på de matematikkurser som utgör behörighetskraven på statistikkursen (Matematik A och B). I nuvarande gymnasiala betygssystem erhåller eleverna ett betyg på varje matematikkurs; betyget är fristående från de betyg man fått på eventuella tidigare kurser i matematik. Betygsskalan är icke-godkänd (IG), godkänd (G), väl godkänd (VG) och mycket väl godkänd (MVG). Betygskriterierna för G, VG och MVG speglar i hög grad vilken kunskapsnivå[13] eleven uppnått inom de områden som anges i kursplanen.

Betyget IG – som en elev får ifall man inte lever upp till betygskriterierna för betyget G – är något problematiskt att tolka när man ska räkna antalet kurser i matematik som en student läst på gymnasiet. Skälet är att det är oklart ifall eleven fått betyget IG p.g.a. att denne ej fullföljt kursen eller p.g.a. att man ej levt upp till betygskriterierna för betyget G trots att man fullföljt kursen. Det faktum att samtliga studenter som antogs till statistikkursen har en gymnasieexamen talar för att man valt en annan kurs i stället. Betyget IG utgör mindre än 3 procent av samtliga matematikbetyg i datamaterialet; för enkelhetens skull kodas betyget IG i analysen som om att man inte läst kursen.[14] För varje matematikkurs skapas en diskret betygsvariabel, vilken antar värden i intervallet 0-2 där varje heltal motsvarar G, VG respektive MVG.[15]

Sex stycken statistikstudenter hade ett gymnasiebetyg från ett tidigare betygssystem. I det systemet var matematikämnet inte uppdelat i separata kurser med enskilda betyg på varje kurs. Istället lästes matematik kontinuerligt under hela gymnasietiden och eleven fick ett slutbetyg som speglade elevens totala studieresultat under gymnasiet. De gymnasiala meriterna i matematik för dessa individer är svåra att jämföra med meriter erhållna i det nuvarande betygssystemet. Av den anledningen exkluderas dessa sex observationer från analysen. Återstår gör 134 observationer som kommer att användas i den empiriska analysen.

13 Kunskapsnivå syftar på begrepp som faktakunskap, begreppskunskap, procedurkunskap och metakognitiv kunskap (se Blooms reviderade taxonomi: Krathwohl, 2002).

14 Om man istället exkluderar dessa observationer från analysen ändras resultaten bara marginellt.

15 Faktum är att betyg mäts på en ordinal skala och därför vore det potentiellt mer korrekt att istället koda betyg med hjälp av dummy-variabler. I en känslighetsanalys som diskuteras senare i artikeln visar det sig dock att resultaten inte påverkas av om betyg kodas som en diskret kvantitativ variabel eller med hjälp av dummy-variabler.

EMPIRI

Avsnittet analyserar hur studenternas poäng på tentamen i statistik påverkas dels av antalet kurser i matematik som lästs utöver behörighetskraven och dels av studenternas betyg på de matematikkurser som ingår i behörighetskraven. Först diskuteras den strategi som används för att skatta sambanden och därefter följer analys och resultat.

Identifieringsstrategi

När det gäller att skatta effekten av antalet kurser i gymnasiematematik som studenten läst utöver behörighetskraven vore ett naivt sätt att helt enkelt estimera en regression där tentamensresultatet är den beroende variabeln och antalet kurser den oberoende variabeln. Det uppenbara problemet med det tillvägagångssätt är att det troligen inte är slumpmässigt vilka individer som väljer att läsa ytterligare kurser i matematik på gymnasiet. Man kan mycket väl föreställa sig att individer med t.ex. hög allmän matematisk förmåga väljer att läsa fler kurser matematik och att dessa individer också presterar bättre på statistikkursen. Då är det möjligt att finna ett positivt samband mellan antalet kurser i matematik utöver behörighetskraven och studieresultat på statistikkursen utan att det för den skull är antalet kurser som i sig har en påverkan på studieresultatet. Ett sådant resultat kan likväl bero på att individer med hög allmän matematisk förmåga väljer att läsa fler kurser i matematik på gymnasiet.

Det experiment man i teorin skulle vilja genomföra för att kunna mäta effekten av antalet kurser vore att gå ut på gymnasieskolorna och slumpmässigt bestämma hur många kurser gymnasiematematik som en elev ska läsa utöver behörighetskraven. Därefter skulle man låta individerna skriva tentamen i statistik. Med ett sådant experiment – vilket förstås inte är möjligt att genomföra i praktiken – skulle det vara möjligt att skatta effekten på ett mer korrekt (väntevärdesriktigt) sätt. Tankeexperimentet kan dock hjälpa oss att inse att ifall det finns ett annat sätt att bryta korrelationen mellan matematisk förmåga och antalet kurser som lästs utöver behörighetskraven borde det fortfarande vara möjligt att skatta den verkliga effekten. Det sätt – den s.k. identifieringsstrategin – som kommer att användas är att jämföra individer som läst olika mycket matematik på gymnasiet, men som har *lika* matematisk förmåga. Detta åstadkoms genom att inkludera en proxyvariabel för matematisk förmåga i analysen.

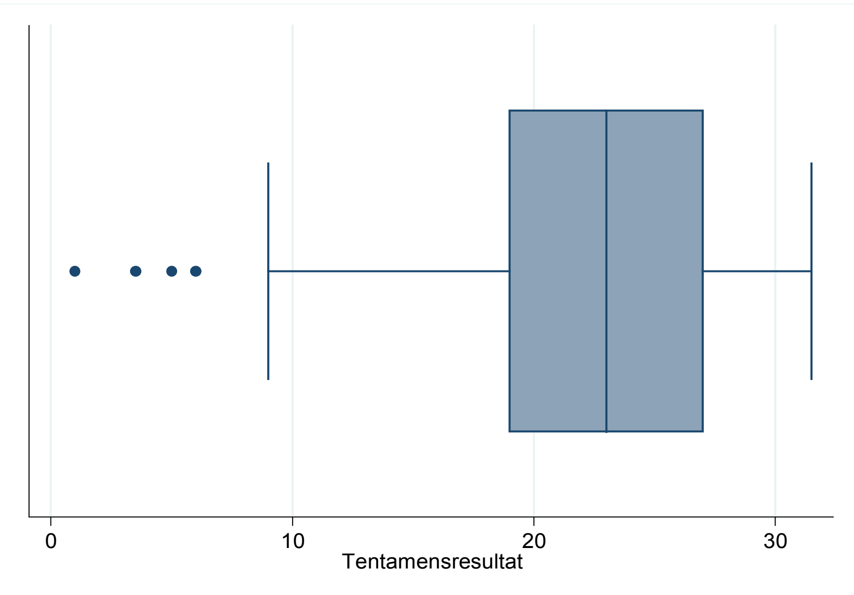
Hur konstrueras en proxyvariabel för allmän matematisk förmåga? Här utnyttjas det faktum att samtliga studenter som antogs till kursen i statistik har läst Matematik A och B på gymnasiet. Med andra ord finns det ingen selektionsproblematik vad gäller vilka individer som valt att läsa dessa kurser. Detta faktum kan utnyttjas genom att använda gymnasiebetyget på dessa kurser som proxyvariabel för allmän matematisk förmåga. Med denna proxy inkluderad som en kontrollvariabel i analysen bör studenter med lika allmän matematisk förmåga kunna jämföras.

Notera att man inte kan kringgå selektionsproblematiken genom att t.ex. istället använda betyget på matematikkursen på den högsta nivån som studenten läst som en proxyvariabel. Det finns flera anledningar till att det inte fungerar. En anledning är t.ex. att det eventuellt är svårare att få ett högt betyg på en mer avancerad gymnasiekurskurs i matematik. Då kommer ett och samma betyg på två olika kurser inte motsvara samma allmänna matematiska förmåga. Alltså är en sådan variabel ingen proxyvariabel för allmän matematisk förmåga. På liknande sätt kan man argumentera för varför det inte heller fungerar att använda studentens genomsnittliga betyg för matematikkurserna som studerats på gymnasiet.

De andra oberoende variablerna av intresse är gymnasiebetyget på Matematik A och B. Ovan argumenterades för att dessa variabler ska inkluderas i analysen p.g.a. att de är viktiga kontrollvariabler. Variablerna kommer därför att fungera både som kontrollvariabler och variabler vars parameterestimat är av intresse. Notera att betygsvariablerna kommer att fånga upp alla bakomliggande faktorer som bestämmer betyget såsom problemlösningsförmåga, allmän matematisk förmåga, motivation, förmåga till logiskt tänkande och ackumulerad matematisk kunskap före gymnasieskolan.

Deskriptiv statistik

Den beroende variabeln antar värden i intervallet 0-32, där 32 är maximal poäng på tentamen på statistikkursen. Lådagrammet i figur 1 visar fördelningen av tentamensresultaten. Där framgår att 50 procent av alla resultat finns i intervallet 19-27 poäng, 25 procent av resultaten är lägre än 19 poäng och 25 procent är högre än 27 poäng. Till vänster i lådagrammet kan också ses att det förekommer fyra stycken avvikande tentamensresultat (s.k. extremvärden). Det är observationer vars värden skiljer sig markant från övriga värden genom att de ligger långt från lådans vänstra kant.[16] Troligen förklaras de låga poängen (mellan ett och sex poäng) av att individerna inte tentamensläst utan enbart skrivit tentamen som en förberedelse inför omtentamen. Det är orimligt att få så låga poäng ifall man verkligen har tentamensläst. Av den anledningen exkluderas dessa individer från analysen av tentamenspoängen.[17] Återstår gör 102 studenter för vilka tentamenspoängen observeras.



Figur 1. Fördelningen av tentamensresultat.

16 Ett extremvärde definieras som en observation som ligger mer än 1.5 kvartilavstånd från lådans vänstra kant där ett kvartilavstånd är avståndet mellan lådans vänstra och högra kant.

17 Resultatet blir dock nästan identiskt om de inkluderas.

Tabell 1 redovisar medelvärdet för tentamensresultat betingat på antalet matematikkurser som lästs på gymnasiet. Noterbart är att det finns en tendens till att tentamensresultatet är högre ju fler kurser man läst. Tabellen visar också att majoriteten av studenterna läst antingen till och med Matematik B eller C. Det innebär att resultaten i denna studie särskilt kommer att spegla betydelsen av att ha läst Matematik C utöver Matematik B.

Tabell 1. Medelvärde tentamensresultat och antal matematikkurser.

Läst upp till och med	Tentamensresultet	Antal studenter
Matematik A	21.8	3
Matematik B	22.4	30
Matematik C	23.1	55
Matematik D	20.7	8
Matematik E	27.5	6
		N=102

Tabell 2 redovisar istället medelvärdet för tentamensresultat betingat på gymnasiebetyg på Matematik A respektive B. Medelvärdet tycks vara högre för studenter med högre betyg – det verkar gälla både för Matematik A och B.

Tabell 2. Medelvärde tentamensresultat och gymnasiebetyg.

	Efter betyg på Matematik A		Efter betyg på Matematik B	
Gymnasiebetyg	Tentamensresultat	Antal studenter	Tentamensresultat	Antal studenter
G	20.9	29	21.8	56
VG	23.0	48	23.4	28
MVG	25.1	25	26.4	15
		N=102		N=99

Not: Att antalet observationer skiljer sig åt för Matematik A och B beror på att det förekommer tre individer i data där det saknas uppgift kring betyget på Matematik B.

Regressionsanalys

Deskriptionen antyder att det finns ett positivt samband mellan tentamensresultatet på kursen i statistik och antalet kurser matematik utöver behörighetskraven som lästs på gymnasiet. Deskriptionen antyder också att det finns ett positivt samband mellan tentamensresultat och betyget på Matematik A och B. Nästa steg blir att med hjälp av regressionsanalys undersöka ifall sambanden existerar även i statistisk mening och om de i så fall har en kausal tolkning (vilket undersöks genom att inkludera proxyvariablerna i analysen).

Resultaten från regressionsanalysen redovisas i Tabell 3 (på nästa sida). Syftet med den första regressionen är att utan en kontrollvariabel för allmän matematisk förmåga undersöka ifall det finns ett statistiskt samband mellan tentamensresultatet och antal kurser i matematik som lästs utöver behörighetskraven. Som kan utläsas från Kolumn 1 är den estimerade parametern positiv men inte statistiskt signifikant skiljd från noll. Nästa modell adderar gymnasiebetyget på Matematik A som en proxyvariabel för allmän matematisk förmåga. Parameterestimatet för antal kurser minskar i storlek och förblir icke-signifikant (se Kolumn 2). Dock framgår att betyget på Matematik A är signifikant positivt associerat med tentamensresultatet. Ett stegs högre betyg är förknippat med att prestera två poäng mer på tentamen. I Kolumn 3 framgår att resultatet är snarlikt när betyget på Matematik B används istället för betyget på Matematik A.

Tabell 3. Tentamensresultat och matematikkunskaper. Linjär regression.

	(1)	(2)	(3)
Konstant	21.4***	20.5***	20.7***
	[1.28]	[1.35]	[1.42]
Antal kurser utöver behörighetskraven	0.83	0.23	0.43
	[0.57]	[0.64]	[0.68]
Betyg Matematik A	-	2.02***	-
		[0.67]	
Betyg Matematik B	-	-	2.09***
			[0.75]
R^2	0.02	0.09	0.10
Antal studenter	102	102	99

Not: Den beroende variabeln är tentamensresultat. Alla modeller inkluderar tentamensspecifika fixa effekter. Att antalet observationer skiljer sig åt för modell 3 beror på att det förekommer tre individer i data där det saknas uppgift kring betyget på Matematik B. *, **, and *** betecknar signifikansnivåerna tio , fem respektive en procent. Standardfelen (inom parantes) är robusta mot heteroskedasticitet.

I syfte att undersöka ifall antalet kurser i matematik som lästs på gymnasiet utöver behörighetskraven och betyget på Matematik A och B påverkar resultatet på tentamen olika för olika intervall på tentamensresultatsskalan undersöks separat sannolikheten att få betyget G respektive VG på statistiktentamen. Notera att i denna analys inkluderas även individer som inte skrev tentamen alls och därmed inte fått godkänt på kursen, vilket resulterar i totalt 134 observationer. Analysen görs genom att skatta s.k. icke-linjära Probit-modeller och i tabellerna redovisas marginaleffekter.[18]

För att analysera sannolikheten att en student får godkänt eller högre betyg på tentamen konstrueras en dikotomisk variabel som antar värdet noll ifall studenten hade strikt färre än nitton poäng på tentamen – eller inte skrev tentamen alls – och ett ifall studenten hade nitton poäng eller fler.[19] Samma oberoende variabler som tidigare inkluderas i analysen. Den första kolumnen i Tabell 4 (se nästa sida) visar att antalet kurser i matematik utöver behörighetskraven inte är signifikant associerat med sannolikheten att få ett godkänt eller högre betyg på tentamen. Detta resultat ändras inte när kontrollvariabler adderas i form av betyget på Matematik A och B (se Kolumn 2 respektive 3). Det är återigen betygen på dessa matematikkurser som har betydelse: sannolikheten att få godkänt eller bättre resultat på tentamen är signifikant högre ifall studenten har ett högt betyg på Matematik A eller B. För kursen Matematik B gäller t.ex. att ett stegs högre betyg är associerat med 20 procentenheters högre sannolikhet att få godkänt eller högre betyg på kursen i statistik (se Kolumn 3).

18 Marginaleffekterna skattas för stickprovets medelvärde med avseende på de förklarande variablerna och har här samma tolkning som koefficienterna i en regression skattad med minsta-kvadratmetoden. Se Wooldridge (2002) och Green (2008) för ytterligare detaljer.

19 15 individer fick underkänt resultat på tentamen och 28 studenter skrev inte tentamen alls.

Tabell 4. Sannolikheten att få godkänt på tentamen. Marginaleffekter (Probit).

	(1)	(2)	(3)
Antal kurser utöver behörighetskraven	-0.05	-0.08	-0.08
	[0.05]	[0.05]	[0.05]
Betyg Matematik A	-	0.13***	-
		[0.06]	
Betyg Matematik B	-	-	0.20***
			[0.079]
Pseudo R^2	0.01	0.03	0.06
Antal studenter	134	134	130

Not: Den beroende variabeln en dummyvariabel som antar värdet ett ifall studenten erhöll 19 poäng eller mer på tentamen, annars värdet noll. Båda modeller inkluderar tentamensspecifika fixa effekter. Att antalet observationer skiljer sig åt för modell 3 beror på att det förekommer fyra individer i data där det saknas uppgift kring betyget på Matematik B. *, **, and *** betecknar signifikansnivåerna tio, fem respektive en procent. Standardfelen (inom parantes) är robusta mot heteroskedasticitet.

I nästa steg upprepas analysen, men nu skattas istället sannolikheten att erhålla VG på tentamen i statistik.[20] På liknande sätt som tidigare skapas en dikotomisk variabel som beroende variabel; variabeln antar värdet noll ifall individen hade strikt färre än 25 poäng på tentamen – eller inte skrev tentamen alls – och ett ifall individen hade 25 poäng eller fler. De oberoende variabler är desamma som tidigare. Kolumn 1 i Tabell 5 visar att det finns en signifikant (på tioprocentsnivån) positiv association mellan sannolikheten att få betyget VG på tentamen och antalet kurer matematik som lästs utöver behörighetskraven. Associationen försvinner dock när kontrollvariabler för gymnasiebetyg på Matematik A respektive B inkluderas i modellen (se Kolumn 2 respektive 3). Istället är det även i detta fall betyget på matematikkurserna på gymnasiet som har betydelse: sannolikheten att en student har VG är signifikant högre – omkring 20 procentenheter – ifall studenten har ett stegs högre betyg på Matematik A eller B.

Tabell 5. Sannolikheten att få väl godkänt på tentamen. Marginaleffekter (Probit).

	(1)	(2)	(3)
Antal kurser utöver behörighetskraven	0.10*	0.06	-0.09
	[0.05]	[0.05]	[0.06]
Betyg Matematik A	-	0.20***	-
		[0.06]	
Betyg Matematik B	-	-	0.19***
			[0.06]
Pseudo R^2	0.03	0.09	0.09
Antal studenter	134	134	130

Not: Den beroende variabeln en dummyvariabel som antar värdet ett ifall studenten erhöll 25 poäng eller mer på tentamen, annars värdet noll. Båda modeller inkluderar tentamensspecifika fixa effekter. Att antalet observationer skiljer sig åt för modell 3 beror på att det förekommer fyra individer i data där det saknas uppgift kring betyget på Matematik B. *, **, and *** betecknar signifikansnivåerna tio, fem respektive en procent. Standardfelen (inom parantes) är robusta mot heteroskedasticitet.

20 42 studenter av 134 erhöll betyget VG på statistiktentamen.

I syfte att undersöka ifall resultaten är robusta har en rad känslighetskontroller genomförts. För det första har analysen upprepats med dummyvariabler istället för diskreta kvantitativa variabler för gymnasiebetyg i matematik. För det andra har analysen även upprepats med dummyvariabler istället för diskreta kvantitativa variabler för antalet matematikkurser utöver behörighetskraven som lästs på gymnasiet (en dummyvariabel som indikerar ifall studenten läst Matematik C; en dummyvariabel som indikerar ifall man har läst Matematik D eller E)[21]. Vidare har i analysen av poängen på statistiktentamen även testats för att inkludera de individer som hade en tentamenspoäng nära noll (extremvärdena). Till sist laborerades i analysen av sannolikheten för att erhålla G eller VG på tentamen med att utesluta 1) individer som inte skrev tentamen alls (och av den anledningen inte kunde få G eller VG) och 2) extremvärdena. I samtliga fall ändrades resultaten bara marginellt.

DISKUSSION

Den första utfallsvariabeln som analyserades var antalet poäng på tentamen i statistik. Det finns inget i resultaten som tyder på att antalet matematikkurser som en student läst utöver behörighetskraven skulle vara viktigt för antalet poäng. Istället framgår att det tycks vara vilket betyg studenten har på matematikkurserna som ingår i behörighetskraven som är viktigt för tentamensresultatet. Detta gäller även för den andra utfallsvariabeln (indikatorn för huruvida man fått betyget godkänt eller bättre). Den tredje utfallsvariabeln (indikatorn för betyget väl godkänt) är å andra sidan signifikant (på tioprocentsnivån) associerat med antalet kurser som lästs utöver behörighetskraven när man inte kontrollerar för betyg på Matematik A eller B. När en sådan kontroll adderas till modellen minskar dock associationen i storlek och upphör att vara signifikant skiljd från noll. Det indikerar – som tidigare diskuterats – att det inte är slumpmässigt vilka studenter som läst fler kurser i matematik på gymnasiet utan att det finns en positiv selektion såtillvida att de med högt betyg på de matematikkurser som ingår i behörighetskraven är mer benägna att läsa fler kurser matematik. Sammanfattningsvis leder analysen av de tre utfallsvariabler till en genensam slutsats: det finns ingen evidens för att antalet kurser matematik studenterna läst utöver behörighetskraven skulle vara viktigt för studieresultatet på en grundläggande kurs i statistik; vad som är viktigt tycks istället vara betyget på de matematikkurser som ingår i behörighetskraven, d.v.s. betygen på Matematik A och B.

Hur förhåller sig resultaten i denna studie till de empiriska resultaten i liknande tidigare studier? I motsats till Cohn (1972) finner denna studie inget samband mellan mängden matematik som lästs innan statistikkursen och studieresultatet. En förklaring kan vara att Cohn (1972) i motsats till denna studie inte specifikt tar hänsyn till studentens matematikbetyg. I den mån matematikbetyg och antalet kurspoäng i matematik korrelerar i Cohns (1972) studie finns det risk att skattningen av hur mängden matematik som lästs påverkar studieresultatet på statistikkursen även fångar upp den påverkan som matematikbetyget kan ha. Även Green, Stone, Zegeye & Charles (2007) finner att mängden matematik har betydelse för studieresultatet på en kurs i statistik. Här är det svårare att finna en förklaring till varför resultatet skiljer sig från den nuvarande studien. Att finna en förklaring blir än mer komplicerat av det faktum att man i motsatts till denna studie inte finner att matematikbetyg har betydelse för studieresultatet på en kurs i statistik. Woodward & Galagedera (2006) hittar å andra sidan inte – i överensstämmelse med denna studie – något samband mellan mängden matematik som lästs och studieresultatet på en

21 Det var relativt få studenter som läst Matematik D och E varför de slogs samman till en grupp.

kurs i statistik. Avslutningsvis tycks det inte finnas någon konsensus i den empiriska litteraturen angående hur sambandet ser ut mellan mängden matematik som lästs och studieresultatet på en grundläggande kurs i statistik.

Är det möjligt att på teoretisk väg resonera sig fram till hur sambandet rimligen borde se ut? Baserat på att matematikstudier är en starkt kumulativ process (enligt exempelvis Butterworth, 1999) och att statistik – även på grundläggande nivå – är ett matematiskt ämne kan man hävda att antalet lästa kurser matematik visst borde vara viktigt för studieresultatet på en grundläggande statistikkurs. Å andra sidan kan man inspireras av Woodward & Galagedera (2006) och argumentera för att statistik på grundläggande nivå inte är ett särskilt matematiskt ämne och därför borde antalet kurser matematik som lästs inte ha någon betydelse för studieresultatet på en statistikkurs. Istället skulle det då vara mer allmänna kognitiva förmågor som är viktiga för studieresultatet. Sammanfattningsvis verkar man inte heller på teoretisk väg entydigt kunna avgöra hur antalet kurser i matematik borde påverka studieresultatet på en grundläggande kurs i statistik.

Baserat på de empiriska resultaten i denna studie verkar det ändå rimligt att dra slutsatsen att en grundläggande kurs i statistik inte kräver några särskilda förkunskaper i matematik utan att det istället är allmänna förmågor såsom intuition, logiskt tänkande och problemlösningsförmåga som är viktiga. Det verkar rimligt att studenter med sådana egenskaper både har ett högt betyg på en elementär matematikkurs på gymnasienivå och ett bra studieresultat på en grundläggande kurs i statistik på högskolenivå.

SLUTSATS

Det faktum att statistik på högskolenivå vanligtvis räknas som ett matematiskt ämne motiverar att undersöka sambandet mellan gymnasiala förkunskaper i matematik och studieresultatet på en grundläggande statistikkurs på högskolenivå. Nuvarande studie bidrar till den befintliga litteraturen genom att försöka särskilja betydelsen av hur väl studenterna behärskar behörighetskraven i matematik från betydelsen av antalet matematikkurser man läst utöver behörighetskraven.

Resultaten visar att bra betyg på de behörighetsgivande och mest elementära gymnasiekurserna i matematik är viktiga för studieresultatet på en grundläggande statistikkurs. Något förvånande – och i kontrast till vissa tidigare studier – finns dock inga tecken på att antalet gymnasiekurser i matematik som studerats utöver behörighetskraven i sig skulle förbättra studieresultaten. En tolkning är att en grundläggande statistikkurs inte kräver särskilt djupgående matematiskt kunnande, utan att det är allmänna kognitiva förmågor som intuition, logiskt tänkande och problemlösningsförmåga som är viktiga. En implikation av resultaten tycks vara att ifall man på ett effektivt sätt önskar öka genomströmningen på en grundläggande statistikkurs genom att ändra behörighetskraven ska man anta studenter med högre betyg på de grundläggande gymnasiekurserna i matematik, och inte öka antalet kurser matematik i behörighetskraven. En intervention i statistikundervisning i form av repetition av grundläggande gymnasiematematik verkar inte heller vara ett effektivt sätt att förbättra studieresultaten eftersom vi finner att studenter som läst mer matematik inte presterar bättre på statistikkursen.

REFERENSER

Becker, W. E. (1998). Engaging Students in Quantitative Analysis with Short Case Examples from the Academic and Popular Press. *The American Economic Review*, 88(2), 480–86.

Butterworth, B. (1999). *What Counts: How Every Brain is Hardwired for Math*. Free Press.

Carlson, W. (1999). A Case-Method for Teaching Statistics. *The Journal of Economic Education*, 30(1), 52-58.

Cohn, E. (1972). Students' Characteristics and Performance in Economic Statistics. *Journal of Economic Education*, 3, 106-111.

Garfield, J., Hogg, B., Schau, C. & Whittinghill, D. (2002), First Courses in Statistical Science: The Status of Educational Reform Efforts, *Journal of Statistics Education,* 10(2).

Green, J. J., Stone, C. C., Zegeye, A., & Charles, T. A. (2009). How much math do students need to succeed in business and economics statistics? an ordered probit analysis. *Journal of Statistics Education*, 17(3).

Greene, W. H. (2008). *Econometric Analysis. Sixth Edition*, (Prentice Hall).

Green, J., Stone, C., Zegeye, A. & Charles, T. (2007). Changes in Math Prerequisites and Student Performance in Business Statistics: Do Math Prerequisites Really Matter? *Journal of Economics and Finance Education*, 6(2), 27-38.

Hakeem, S. A. (2001). Effect of Experiential Learning in Business Statistics. *Journal of Education for Business*, 77(2), 95-98.

Hilmer, S. (1996). A Problem-Solving Approach to Teaching Business Statistics. *The American Statistician*, 50, 249-256.

Johnson, M. & Kuennen, E. (2006). Basic math skills and performance in an introductory statistics course. *Journal of Statistics Education*. 14(2), 14.

Krathwohl, D. (2002). A Revision of Bloom's taxonomy: An Overview. *Theory into Practice*, 41(4), 212–218.

Krieg, R. & Uyar, B. (1997). Correlates of Student Performance in Business and Economics Statistics. *Journal of Economics and Finance*, 21, 65-74.

Krieg, R., & Uyar, B. (2001). Student Performance in Business and Economics Statistics: Does Exam Structure Matter?. *Journal of Economics and Finance*, 25, 229-241.

Krutetskii, V.A. (1976). *The Psychology of Mathematical Abilities in Schoolchildren*. Teller, J. (Översatt.). The University of Chicago Press. Chicago. (Originalarbetet publicerades 1968).

Magel, R. (1996). Increasing Student Performance in Large Introductory Statistics Classes. The American Statistician, 50, 51-56.Niss, M. (1994). Mathematics in society. In R. Biehler, R. W. Scholz, R. Strässer, & B. Winkelmann (Eds.). *Didactics of mathematics as a scientific discipline*, 367–378.

Stromberg, A, & Ramanathan, S. (1996). Easy implementation of writing in introductory statistics courses. *The American Statistician*, 50(2), 159-163.

Utts, J., Sommer, B., Acredolo, C., Maher, M., & Matthews, H. (2003). A Study Comparing Traditional and Hybrid Internet-Based Instruction in Introductory Statistics Classes. *Journal of Statistics Education*, 11(3).

Ward, B. (2004). The Best of Both Worlds: A Hybrid Statistics Course. *Journal of Statistics Education*, 12(3).

Wistedt, I. & Lagergren, R. m.fl. (2006). Pedagogik för elever med intresse och fallenhet för matematik. *Nämnaren* 3, 16-21.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge. MA: MIT Press.

Woodward, G. & Galagedera, D. (2006). Does prior mathematics knowledge really lead to variation in elementary statistics performance?: Evidence from a developing country. *International Journal of Educational Development*, 26(6), 631-639.