

Lånebenägenhet bland studerande med studiemedel

CSN, rapport 2017:4

The logo for CSN (Centrala Studietoget) is located in the bottom right corner. It consists of a red triangle pointing towards the bottom right, with the letters 'CSN' in white, bold, sans-serif font inside it.

CSN

Innehåll

Sammanfattning	3
1. Inledning	4
1.1 Uppdrag och syfte	4
1.2 Avgränsningar och metod	4
1.3 Rapportens struktur	5
2. Andel som tar lån i olika grupper	6
2.1 Kvinnor och män	6
2.2 Ålder	6
2.3 Svensk respektive utländsk bakgrund	7
2.4 Familjesituation	8
2.5 Inkomst	9
2.6 Betyg i gymnasieskolan	10
2.7 Föräldrars utbildningsbakgrund	11
2.8 Föräldrars inkomster	12
2.9 Nivå på studierna	14
2.10 Studiernas omfattning	15
2.11 Generellt respektive högre bidrag	15
2.12 Inriktning på studierna (enbart universitet och högskola)	15
3. Sannolikhet att en studerande med studiemedel även tar studielån	16
3.1 Analysmetod – regression	16
3.2 Inkomst, nivå på studierna, studieomfattning och studerandes bakgrund är viktiga faktorer för att förklara lånebenägenheten	17
Referenser	19
Bilaga	20

Sammanfattning

Benägenheten att ta lån kan påverka studiedeltagandet

Studiestödet ska verka rekryterande för både kvinnor och män och därmed bidra till ett högt deltagande i utbildning. Det ska också utjämna skillnader mellan individer och grupper i befolkningen och i och med det bidra till ökad social rättvisa.

Det är känt att olika individer är olika benägna att ta studielån. Denna benägenhet att ta lån kan påverka vilka och hur många som söker sig till studier. Studielånet är därmed en faktor som kan påverka studiedeltagandet, både totalt och i olika grupper i befolkningen.

I denna registerstudie presenteras andelen av studiemedelstagarna som tar lån, fördelat efter olika bakgrundsvariabler. Det görs även en analys av olika faktorerens inbördes betydelse för sannolikheten att en studerande med studiemedel även tar studielån.

Stora skillnader i lånebenägenhet

Det finns stora skillnader i lånebenägenhet mellan olika studerandegrupper. Olika individers olika benägenhet att ta lån kan ha skilda orsaker. Det kan exempelvis handla om behovet av lån i förhållande till andra inkomstkällor och till de utgifter var och en har. Det kan även handla om en bedömning av framtida arbetsmarknadsutsikter och en förväntad inkomstutveckling, där möjligheten att betala tillbaka lånet kan vägas in. Även traditioner och kulturell bakgrund kan påverka benägenheten att låna.

Män tar oftare lån än kvinnor och studerande med svensk bakgrund har en högre lånebenägenhet än studerande med utländsk bakgrund. Bland de studiemedelstagare som är i åldern 25–34 år är det en högre andel som tar lån än bland yngre och äldre studerande. Heltidsstuderande tar oftare studielån än deltidsstuderande och de som studerar på låga utbildningsnivåer tar mer sällan lån än de som studerar på eftergymnasial nivå. Vidare tar ensamstående förhållandevis ofta lån och de som bor hos sina föräldrar tar relativt sällan lån.

Studerande med lägre inkomster lånar i högre utsträckning än de som har högre inkomster. Detta kan vara en följd av att den som har högre inkomster inte behöver ta lån, men även av att den som inte vill ta lån måste skaffa sig en annan inkomst och då arbetar i högre utsträckning.

Även föräldrars utbildningsbakgrund och inkomstsituation kan påverka studerandes lånebeslut. Lånebenägenheten är högre bland studerande med högutbildade föräldrar än bland dem som har föräldrar med lägre utbildning. Den är vidare högre bland dem som har föräldrar med högre inkomster och lägre bland dem som har föräldrar med lägre inkomster. I dessa fall tycks föräldrarnas utbildningsbakgrund ha större inflytande än deras inkomstsituation.

Studienivå, inkomst och studiernas omfattning viktiga faktorer

Genom regressionsanalys undersöks effekten av flera olika variabler samtidigt. Analysen visar bland annat att studienivå och inkomst är de enskilt viktigaste faktorerna för att förklara lånebenägenheten. Vidare är studiernas omfattning av stor betydelse, liksom studerandes bakgrund. Även familjesituation, ålder och kön spelar roll, men i något mindre utsträckning.

1. Inledning

1.1 Uppdrag och syfte

Studiestödet ska verka rekryterande för både kvinnor och män och därmed bidra till ett högt deltagande i utbildning. Det ska också utjämna skillnader mellan individer och grupper i befolkningen och i och med det bidra till ökad social rättvisa.¹

Studiemedelssystemet är en viktig del av studiestödet. Studiemedlen är avsedda att täcka levnadsomkostnader under studietiden. Studiemedlen består av både bidrag och lån, vilket innebär att levnadsomkostnaderna delas mellan den studerande och samhället. Kostnadsdelningen motiveras i första hand av att både samhället och individen tjänar på utbildning.

Det är känt att olika individer är olika benägna att ta studielån. Denna benägenhet att ta lån kan påverka vilka och hur många som söker sig till studier. Studielånet är därmed en faktor som kan påverka studiedeltagandet, både totalt och i olika grupper i befolkningen.

Av detta följer att det är av intresse att veta vilka studiemedelstagare som tar lån och vilka som inte gör det, samt vilka faktorer som kan påverka lånebenägenheten. Förbättrade kunskaper om lånebenägenhet kan bidra till att öka förståelsen för studiedeltagandet och för studiestödets betydelse och roll för bland annat rekryteringen till studier.

Syftet med denna studie är att bedöma olika faktorer betydelse för viljan och behovet av att ta lån, dvs. för den så kallade lånebenägenheten.

De huvudsakliga frågeställningarna är:

- Hur ser lånebenägenheten ut bland studiemedelstagarna, fördelat efter olika bakgrunds faktorer?
- I vilken utsträckning påverkar olika faktorer lånebenägenheten?

CSN ska genom bevakning och analys av studerandes sociala och ekonomiska situation säkerställa myndighetens kunskap inom området.² Denna rapport har tagits fram inom ramen för detta uppdrag.

1.2 Avgränsningar och metod

Data och avgränsningar

Undersökningen är en registerstudie där populationen i undersökningen utgörs av samtliga studiemedelstagare under 2014. Antalet individer i populationen kan skilja sig något från uppgifter redovisade på andra ställen, dels beroende på tidpunkt för framtagande av uppgifterna, dels mot bakgrund av mindre skillnader i definitioner.

Samtliga uppgifter är registeruppgifter hämtade dels från CSN:s administrativa system, dels från register hos Statistiska Centralbyrån (SCB).

Uppgifterna kommer från följande register.

¹ Prop. 2016/17:1.

² Regleringsbrev för budgetåret 2016 avseende Centrala studiestödsnämnden.

Studiestödets informationssystem (STIS)

Skolverkets elevregister

Universitets- och högskoleregistret

Inkomst- och taxeringsregistret

Registret för totalbefolkningen (RTB)

Totalt omfattar materialet 483 671 studerande med studiemedel. För vissa individer saknas vissa uppgifter vilket gör att det totala antalet individer är färre i redovisningen av vissa bakgrundsvariabler. De variabler som förekommer i rapporten är: kön, ålder, familjesituation, svensk respektive utländsk bakgrund, betyg från gymnasieskolan, inkomst, föräldrars utbildningsbakgrund och inkomster, nivå på studierna, inriktning på studierna, studiernas omfattning och bidragsnivå.

I tabellerna i rapporten redovisas andelen av studiemedelstagarna som förutom bidrag även har haft studielån (så kallat grundlån) under året, fördelat efter de olika variablerna. För att räknas som låntagare räcker det med att lån uppburits under minst en vecka.³

Hur de olika variablerna definieras framgår av respektive avsnitt i kapitel 2.

Undersökningen handlar enbart om studerande som har studiemedel. Studerande som inte har studiemedel omfattas inte.

Arbetsätt och metod

CSN har utarbetat rapporten i samarbete med SCB. CSN har överlämnat uppgifter över studerande med studiemedel till SCB som sedan har lagt på bakgrundsuppgifter hämtade från de register SCB förfogar över. Samtliga uppgifter har sedan, i avidentifierat skick, gjorts tillgängliga för CSN:s bearbetning och analys.

SCB har även, på uppdrag av CSN, genomfört en logistisk regressionsanalys för att bedöma olika faktorerens betydelse för lånebenägenheten. Hur olika faktorer beaktas i denna analys, samt metoden som sådan, diskuteras mer ingående dels i kapitel 3, dels i rapportens bilaga.

1.3 Rapportens struktur

I kapitel 2 redovisas lånebenägenheten utifrån olika variabler. Det innebär att andelen studerande som tar lån redovisas fördelat efter kön, ålder, familjesituation, svensk respektive utländsk bakgrund, betyg från gymnasieskolan, inkomst, föräldrars utbildningsbakgrund och inkomster, nivå på studierna, inriktning på studierna, studiernas omfattning och bidragsnivå. Redovisningen är deskriptiv men åtföljs av ett kortare resonemang som kommenterar uppgifterna och som diskuterar eventuella orsaker till att lånebenägenheten skiljer sig mellan olika studerandekategorier.

I kapitel 3 sammanfattas den logistiska regressionsanalys som har genomförts för att bedöma faktorer som kan förutse studiemedelstagares låntagande. Analysen i sin helhet presenteras i bilaga.

³ I normalfallet har den som har rätt till studiemedlens bidragsdel även rätt till studielån. Det finns dock lånerestriktioner fr.o.m. 47 års ålder. Ett mindre antal studiemedelstagare som är 47 år och äldre kan alltså ha en vilja att ta lån, men sakna rätten att få lån.

2. Andel som tar lån i olika grupper

2.1 Kvinnor och män

Män som studerar med studiemedel tar oftare lån än kvinnliga studiemedelstagare. Bland männen är lånebenägenheten 69,9 procent och bland kvinnorna 67,2 procent. Skillnaden i lånebenägenhet mellan kvinnor och män är bestående över tid.

Tabell 1. Lånebenägenhet fördelad efter kön

	Antal med bidrag	Antal med lån	Lånebenägenhet
Kvinnor	286 228	192 262	67,2 %
Män	197 441	138 029	69,9 %
Samtliga	483 671	330 291	68,3 %

En bidragande orsak till att män lånar i lite högre utsträckning än kvinnor är att män mer sällan studerar på deltid och därmed mer sällan har andra inkomster från arbete. En annan orsak kan vara att män oftare än kvinnor lever i ensamhushåll och därmed mer sällan kan få ekonomiskt stöd från en partner. En tidigare undersökning från CSN har också visat att män som studerar och har barn lever i mer inkomstsvaga hushåll än studerande kvinnor som har barn.⁴ Även det skulle kunna föranleda högre lånebenägenhet bland män.

Tabell 2. Lånebenägenhet fördelad efter kön och ålder

	-24 år	25-34 år	35 år-
Kvinnor	66,1 %	74,0 %	54,5 %
Män	63,6 %	80,7 %	70,1 %

De kvinnor som studerar med studiemedel är i genomsnitt äldre än män som studerar med studiemedel. Vid en jämförelse av män och kvinnor i samma åldersgrupper är skillnaderna i lånebenägenhet större än för hela populationen. Störst är skillnaden i den äldsta åldersgruppen, studerande 35 år och äldre, där 54 procent av kvinnorna och 70 procent av männen tar lån.

2.2 Ålder

Lånebenägenheten är högst bland studerande i åldern 25–34 år, lägre bland yngre studerande och lägst i den äldsta åldersgruppen. Detta förhållande har varit oförändrat över tid.

Tabell 3. Lånebenägenhet fördelad efter ålder¹

	Antal med bidrag	Antal med lån	Lånebenägenhet
-24 år	261 758	170 189	65,0 %
25-34 år	165 965	127 310	76,7 %
35 år-	55 948	32 792	58,6 %
Samtliga	483 671	330 291	68,3 %

¹ Ålder vid årets slut.

⁴ CSN (2014).

Att lånebenägenheten varierar mellan studiemedelstagare i olika åldrar beror både på att studerande i olika åldrar har olika inkomskällor och på att de har olika utgifter.⁵ Yngre studerande har ofta lägre utgifter för framför allt boende. De har inte heller barn i lika hög utsträckning som äldre studerande. En relativt hög andel yngre bor också kvar hemma eller får ekonomiskt stöd från sina föräldrar. Samtliga dessa faktorer kan göra att fler klarar sig utan att ta lån.

Studerande i den äldsta åldersgruppen har å sin sida ofta familj och barn, vilket kan innebära högre utgifter. Äldre studerande har dessutom i genomsnitt högre boendekostnader. Trots dessa utgiftsposter lånar äldre alltså i lägre utsträckning. Det kan dels bero på att äldre arbetar i högre utsträckning än yngre, dels på att många äldre får ekonomiskt stöd från en partner. Dessutom kan många lite äldre studerande ha haft studielån sedan tidigare och vara mindre attraherade av att få en ny studieskuld, för vilken återbetalningstiden också blir ganska kort. Ett mindre antal studiemedelstagare som är 47 år eller äldre kan också sakna rätt att låna.

Studerande i den mellanliggande åldersgruppen, där lånebenägenheten är högst, har oftast lämnat föräldrahemmet. Många har också hunnit skaffa familj. Detta innebär att utgifterna har ökat för många samtidigt som stödet från föräldrarna ofta har minskat. I denna åldersgrupp är det inte heller lika vanligt med ekonomiskt stöd från en partner som i den äldsta åldersgruppen. Det kan i sin tur troligen bero på att hushållets samlade inkomster är lägre. Många kan exempelvis leva tillsammans med en annan studerande.

2.3 Svensk respektive utländsk bakgrund

Lånebenägenheten är betydligt lägre bland studiemedelstagare med utländsk bakgrund än bland dem som har svensk bakgrund. Det skiljer hela 24 procentenheter mellan dem som är utrikes födda och dem som är födda i Sverige och har inrikes födda föräldrar.

Tabell 4. Lånebenägenhet fördelad efter bakgrund

	Antal med bidrag	Antal med lån	Lånebenägenhet
Inrikes födda med två inrikes födda föräldrar	316 837	234 999	74,2 %
Inrikes födda med en inrikes och en utrikes född förälder	44 250	31 995	72,3 %
Inrikes födda med två utrikes födda föräldrar	34 046	18 767	55,1 %
Utrikes födda	88 493	44 521	50,3 %
Samtliga	483 626	330 282	68,3 %

En del av skillnaden mellan inrikes och utrikes födda kan antas bero på att studierna bedrivs på delvis olika utbildningsnivåer. Utrikes födda studerar i större utsträckning på grundskolenivå än inrikes födda. Skillnaderna i lånebenägenhet kvarstår dock till stor del även vid en jämförelse mellan studerande på samma utbildningsnivå. Det kan noteras att skillnaden minskar något ju högre utbildningsnivå det handlar om, men att den är markant även vid studier på eftergymnasial nivå. Bland studiemedelstagare på eftergymnasial nivå är det dock inrikes födda med två utrikes födda föräldrar som har den lägsta lånebenägenheten.

⁵ Uppgifterna om studerandes inkomster och utgifter i detta och de två följande styckena är hämtade från enkätundersökningar riktade till studerande, se CSN (2016).

Tabell 5. Lånebenägenhet fördelad efter bakgrund och studienivå

	Grundskolenivå (Sverige)	Gymnasienivå (Sverige)	Eftergymnasial nivå (Sverige)	Gymnasienivå (utland)	Eftergymnasial nivå (utland)
Inrikes födda med två inrikes födda föräldrar	67,9 %	62,9 %	76,9 %	81,2 %	86,3 %
Inrikes födda med en inrikes och en utrikes född förälder	70,9 %	63,0 %	74,8 %	76,4 %	83,8 %
Inrikes födda med två utrikes födda föräldrar	58,1 %	46,3 %	54,6 %	78,7 %	82,7 %
Utrikes födda	31,8 %	41,6 %	64,4 %	74,4 %	81,8 %

Mönstret bland utlandsstuderande liknar förhållandet bland studerande i Sverige. Bland dem som studerar utomlands är dock andelen som tar lån betydligt mer likartad oavsett bakgrund.

Det är oklart vilka orsakerna är till att lånebenägenheten skiljer sig åt mellan utrikes och inrikes födda. En möjlighet är att det finns kulturella skillnader i synen på låntagande i allmänhet och låntagande för studier i synnerhet. I Sverige finns en lång tradition av studielån, vilket saknas i många andra länder. Det kan således finnas en ”vana” av att ta studielån bland personer som har växt upp i landet och som kanske har föräldrar som själva har tagit studielån. En annan möjlighet är att studerande med olika bakgrund värderar sina arbetsmarknadsutsikter olika. Om utrikes födda är mer osäkra på möjligheten att etablera sig på arbetsmarknaden, kan det påverka viljan att skuldsätta sig i samband med studierna. Ytterligare en orsak skulle kunna vara att det bland muslimer i vissa fall kan anses oriktigt att ta lån som ska betalas med ränta, vilket då kan minska låntagandet framför allt bland utrikes födda.⁶

2.4 Familjesituation

De uppgifter om familjesituation som finns tillgängliga är av bristande kvalitet. Den största bristen är att det i registeruppgifterna inte går att skilja mellan ensamstående och sammanboende som inte har barn. Sammanboende utan barn återfinns därför i gruppen ensamstående utan barn. En annan brist är att studerande kan vara folkbokförda på en annan plats än där de faktiskt bor. Det lär till exempel inte vara helt ovanligt att studerande är folkbokförda hos sina föräldrar trots att de bor på sin studieort. Trots dessa brister i kvaliteten på uppgifterna väljer vi att här redovisa lånebenägenheten fördelad efter familjesituation. Mot bakgrund av kvalitetsproblemen ska uppgifterna användas med försiktighet.

⁶ Se t.ex. Samuelsson (2000); Sveriges Radio (2005).

Tabell 6. Lånebenägenhet fördelad efter familjesituation¹

	Antal med bidrag	Antal med lån	Lånebenägenhet
Gift, partner utan barn ²	12 168	7 151	58,8 %
Ensamstående utan barn ²	232 526	187 470	80,6 %
Gift, partner, sambo med barn	100 315	55 121	54,9 %
Ensamstående med barn	25 792	16 375	63,5 %
Bor hos föräldrar	106 991	59 634	55,7 %
Samtliga	477 792	325 751	68,2 %

¹ Variabeln familjesituation är skapad av två variabler, familjeställning och familjetyp, hämtade från registret för totalbefolkningen (RTB). Det finns felkällor i uppgifterna, se framför allt fotnot 2. Det kan även finnas fel som beror på felaktig folkbokföring, t.ex. studerande som fortfarande är folkbokförda hos föräldrar, m.m.

² Sammanboende som inte har barn kan inte relateras till varandra och återfinns som ensamstående i statistiken.

Det är troligt att studerande med olika familjesituation har olika behov av att låna och därmed även olika lånebenägenhet. Lånebenägenheten är högst, drygt 80 procent, bland ensamstående utan barn. Det kan möjligen anses logiskt att lånebenägenheten är högst bland ensamstående mot bakgrund av att dessa inte kan få ekonomiskt stöd från en partner och inte har någon att dela hushållsutgifterna med. Något förvånande är dock lånebenägenheten relativt låg, 63,5 procent, bland ensamstående med barn. Lägst är lånebenägenheten bland gifta eller sammanboende med barn och bland studerande som bor hos sina föräldrar. I den förra gruppen kan många sannolikt få stöd från en partner. Den senare gruppen består till stor del av yngre studerande som sannolikt har låga utgifter, inte minst för sitt boende.

2.5 Inkomst

Det är rimligt att anta att den som har högre inkomster har mindre behov av att låna än den som har lägre inkomster. Inkomst kan dock vara en svårtolkad faktor eftersom inkomsten kan påverka lånebenägenheten på detta sätt, men även vara en följd av en ovilja att ta lån. I det senare fallet kan man tänka sig att den som vill avstå från att skuldsätta sig istället måste skaffa sig en inkomst genom arbete. I sammanhanget ska det också nämnas att studiemedlen är behovsprövade och att studerande med höga inkomster inte har rätt till studiemedel under det kalenderhalvår de har sin inkomst. De studerande med höga inkomster som finns med bland studiemedelstagarna har därmed troligen haft merparten av sin inkomst under ett kalenderhalvår då de inte hade studiemedel.

Oavsett åt vilket håll orsakssambandet verkar, är den generella trenden tydlig: Lånebenägenheten minskar med stigande inkomst.

Tabell 7. Lånebenägenhet fördelad efter sammanräknad förvärvsinkomst¹

	Antal med bidrag	Antal med lån	Lånebenägenhet
0-49 999	205 831	137 164	66,6 %
50 000-99 999	104 694	76 128	72,7 %
100 000-149 999	64 580	45 076	69,8 %
150 000-199 999	44 055	29 892	67,9 %
200 000-249 999	24 010	14 589	60,8 %
250 000-299 999	10 702	5 407	50,5 %
300 000-399 999	5 100	2 273	44,6 %
400 000-499 999	609	273	44,8 %
500 000-	208	76	36,5 %
Samtliga ²	459 789	310 878	67,6 %

¹ I sammanräknad förvärvsinkomst ingår löneinkomst inklusive skattepliktiga bidrag och transfereringar (t.ex. pension, arbetsmarknadsstöd och föräldrapenning), samt inkomst av näringsverksamhet.

² Uppgifterna avser inkomster i Sverige. Därför har enbart individer som studerat i Sverige under året tagits med i redovisningen.

2.6 Betyg i gymnasieskolan

Det kan finnas flera tänkbara orsaker till eventuella samband mellan betygsnivå i gymnasieskolan och benägenheten att ta lån vid senare studier med studiemedel. En möjlighet är att studerande med högre betyg studerar på högre utbildningsnivåer och därmed har högre lånebenägenhet (se även avsnitt 2.9 Nivå på studierna). Ett annat samband skulle kunna vara att de med högre betyg har mer högtbildade föräldrar, vilket också kan indikera högre lånebenägenhet (se avsnitt 2.7 Föräldrars utbildningsbakgrund). En annan möjlighet är att de med högre gymnasiebetyg har en bättre ekonomisk situation, antingen genom högre egna inkomster eller genom att det inom denna studerandegrupp är vanligare med stöd från föräldrar eller partner. Det skulle då kunna indikera lägre lånebenägenhet.

Här redovisas betygsnivåer dels i det betygssystem med relativa betyg som gällde fram till och med 1993, dels i de nyare betygssystemen med målrelaterade betyg (inklusive nuvarande betygssystem).

Tabell 8. Lånebenägenhet fördelad efter genomsnittligt gymnasiebetyg (äldre betygssystemet 1-5)

	Antal med bidrag	Antal med lån	Lånebenägenhet
0-1,99	426	333	78,2 %
2,00-2,99	6 511	4 706	72,3 %
3,00-3,99	11 327	6 943	61,3 %
4,00-5,00	2 298	1 097	47,7 %
Samtliga	20 562	13 079	63,6 %

För studerande med betyg från det äldre betygssystemet är lånebenägenheten sjunkande med stigande betygsnivå. Det är dock rimligt att anta att andra faktorer än gymnasiebetyget spelar roll för dessa studerande, vilka är lite äldre, exempelvis hushållsekonomi, besparingar och behov av utbildning. Resultatet, med lägre lånebenägenhet vid högre betyg, kan förefalla rimligt vid ett antagande om att de med högre gymnasiebetyg generellt har haft en bättre inkomstutveckling och mindre utbildningsbehov än de med lägre gymnasiebetyg.

Tabell 9. Lånebenägenhet fördelad efter genomsnittligt gymnasiebetyg (jämförelsetal i målrelaterat betygssystem 0-20)

	Antal med bidrag	Antal med lån	Lånebenägenhet
0-9,99	17 823	12 727	71,4 %
10,00-12,49	47 386	33 532	70,8 %
12,50-14,99	86 958	61 911	71,2 %
15,00-17,49	102 224	73 677	72,1 %
17,50-	84 210	62 921	74,4 %
Samtliga	338 601	244 768	72,3 %

När det gäller studerande med betyg från de målrelaterade betygssystemen är det inte några stora skillnader mellan studiemedelstagare med olika gymnasiebetyg. Den svaga tendensen är dock att de med högre betyg i gymnasieskolan även uppvisar en något större benägenhet att ta studielån. Eftersom utbildningsnivå uppvisar ett relativt starkt samband med lånebenägenhet bör detta dock till stor del kunna förklaras av att de med höga gymnasiebetyg oftare studerar på högre utbildningsnivåer, medan de med låga betyg oftare studerar på lägre utbildningsnivåer (se även avsnitt 2.9 Nivå på studierna).

2.7 Föräldrars utbildningsbakgrund

Tidigare enkätstudier från CSN har visat att studiemedelstagare med högutbildade föräldrar har haft högre lånebenägenhet än studiemedelstagare med föräldrar med lägre utbildningsnivå.⁷

Ett samband mellan föräldrarnas utbildningsnivå och lånebenägenhet skulle kunna ha olika orsaker. En möjlighet är att studerande med högre utbildade föräldrar oftare studerar på eftergymnasial nivå, vilket är förknippat med hög lånebenägenhet (se avsnitt 2.9 Nivå på studierna). Sambandet kan också bero på tradition och synen på utbildning. Det är inte osannolikt att det är mer givet att ta studielån om ens föräldrar har gjort det. I familjer med högre utbildade föräldrar kan också utbildning värdesättas på ett sätt som gör det mer naturligt att ta lån för att ”investera” i utbildning. Tron på den egna studieförmågan kan också vara högre, liksom bedömningen av den framtida löneutvecklingen och därmed möjligheterna att betala tillbaka sitt lån. För den som inte har föräldrar som tagit studielån och kanske inte heller har studerat vidare, kan det vara mer främmande att ta studielån. Å andra sidan bör det vara så att det finns större ekonomiska resurser i familjer där föräldrarna är högutbildade, vilket i sin tur skulle kunna ge dem större möjlighet att understödja studerande barn ekonomiskt. Till frågan om föräldrarnas inkomster återkommer vi i avsnitt 2.8.

I tabell 10 har föräldrarnas utbildning renodlats så att enbart de som har föräldrar med lika hög utbildning finns med. Avsikten med detta är att resultatet ska kunna bli mer tydligt än i det fall även studerande som har föräldrar med olika utbildningsbakgrund tas med. Denna senare grupp finns dock med i tabell 11 där de grupperas tillsammans med dem som har föräldrar med okänd utbildningsnivå.

⁷ CSN (2014).

Tabell 10. Lånebenägenhet fördelad efter föräldrars utbildningsbakgrund (båda föräldrarna har samma utbildningsnivå)¹

	Antal med bidrag	Antal med lån	Lånebenägenhet
Förgymnasial	13 202	7 990	60,5 %
Gymnasial	87 399	60 805	69,6 %
Eftergymnasial	111 400	83 579	75,0 %
Samtliga	212 001	152 374	71,9 %

¹ Det saknas uppgift om utbildningsbakgrund för 96 000 fäder och 69 300 mödrar. För 56 600 studerande saknas uppgift om utbildningsbakgrund för båda föräldrarna.

Registerstudien verifierar det tidigare uppmärksammade sambandet mellan föräldrars utbildningsbakgrund och lånebenägenhet. Lånebenägenheten ökar med stigande utbildningsnivå hos föräldrarna. Det finns en skillnad på 9 procentenheter mellan studerande med föräldrar med förgymnasial utbildning och gymnasial utbildning. Lånebenägenheten är sedan ytterligare 5 procentenheter högre bland dem som har föräldrar med eftergymnasial utbildning.

Tabell 11. Lånebenägenhet fördelad efter föräldrars utbildningsbakgrund (föräldrarna har högst gymnasial utbildning, högst eftergymnasial utbildning respektive övriga)¹

	Antal med bidrag	Antal med lån	Lånebenägenhet
Högst gymnasial	144 529	98 437	68,1 %
Eftergymnasial	111 400	83 579	75,0 %
Övriga	227 742	148 275	65,1 %
Samtliga	483 671	330 291	68,3 %

¹ Det saknas uppgift om utbildningsbakgrund för 96 000 fäder och 69 300 mödrar. För 56 600 studerande saknas uppgift om utbildningsbakgrund för båda föräldrarna.

Vid en jämförelse mellan studerande med föräldrar med högst gymnasial utbildning och studerande vars båda föräldrar har eftergymnasial utbildning, är skillnaden i lånebenägenhet 7 procentenheter. De som har en förälder med eftergymnasial utbildning och en med högst gymnasial utbildning, samt de som har minst en förälder med okänd utbildning, har den lägsta lånebenägenheten. I denna grupp finns många studerande med föräldrar födda i utlandet, vilket kan påverka utfallet (se även avsnitt 2.3 Svensk respektive utländsk bakgrund).

2.8 Föräldrars inkomster

Föräldrars inkomster skulle kunna påverka lånebenägenheten bland studerande genom att föräldrar med högre inkomster normalt har större ekonomiska möjligheter att bidra ekonomiskt, vilket i så fall skulle minska lånebenägenheten. Det ska noteras att uppgifterna avser studiemedelstagare oavsett ålder. Det är sannolikt att äldre studerande mer sällan får ekonomiskt stöd från sina föräldrar, oavsett föräldrarnas inkomstsituation.

Tabell 12. Lånebenägenhet fördelad efter faderns sammanräknade förvärvsinkomst¹

	Antal med bidrag	Antal med lån	Lånebenägenhet
0-49 999	21 248	14 054	66,1 %
50 000-99 999	9 921	6 542	65,9 %
100 000-149 999	16 655	11 701	70,3 %
150 000-199 999	24 872	17 423	70,1 %
200 000-249 999	26 382	18 455	70,0 %
250 000-299 999	30 338	21 493	70,8 %
300 000-399 999	90 836	64 779	71,3 %
400 000-499 999	67 993	48 834	71,8 %
500 000-	105 494	76 864	72,9 %
Samtliga ²	393 739	280 145	71,1 %

¹ I sammanräknad förvärvsinkomst ingår löneinkomst inklusive skattepliktiga bidrag och transfereringar (t.ex. pension, arbetsmarknadsstöd och föräldrapenning), samt inkomst av näringsverksamhet.

² Uppgifterna avser inkomster i Sverige. Därför har enbart individer som studerat i Sverige under året tagits med i redovisningen.

Tabell 13. Lånebenägenhet fördelad efter moderns sammanräknade förvärvsinkomst¹

	Antal med bidrag	Antal med lån	Lånebenägenhet
0-49 999	23 349	14 375	61,6 %
50 000-99 999	12 059	7 887	65,4 %
100 000-149 999	26 935	18 795	69,8 %
150 000-199 999	32 174	22 287	69,3 %
200 000-249 999	40 764	28 375	69,6 %
250 000-299 999	60 085	42 137	70,1 %
300 000-399 999	120 604	87 589	72,6 %
400 000-499 999	55 442	41 013	74,0 %
500 000-	47 639	35 026	73,5 %
Samtliga ²	419 051	297 484	71,0 %

¹ I sammanräknad förvärvsinkomst ingår löneinkomst inklusive skattepliktiga bidrag och transfereringar (t.ex. pension, arbetsmarknadsstöd och föräldrapenning), samt inkomst av näringsverksamhet.

² Uppgifterna avser inkomster i Sverige. Därför har enbart individer som studerat i Sverige under året tagits med i redovisningen.

Resultaten kan möjligen anses något förvånande eftersom den tydliga trenden är att de som har föräldrar med högre inkomster oftare tar lån än de som har föräldrar med lägre inkomster. Bättre föräldraekonomi innebär alltså inte att studerande avstår från att ta lån.

Dessa resultat tyder på att andra faktorer än familjens ekonomi är viktigare för valet att låna. Det finns exempelvis ett mycket tydligt samband mellan föräldrarnas inkomster och deras utbildningsnivå, där högutbildade föräldrar även har höga inkomster. I dessa fall tycks alltså utbildningsbakgrunden ha större inflytande på barnens val att ta lån. Med andra ord kan man möjligen dra slutsatsen att icke-monetära aspekter, såsom tradition och värderingar, väger tyngre än rent ekonomiska faktorer, åtminstone när det gäller ekonomin i familjekretsen.

2.9 Nivå på studierna

Lånebenägenheten varierar stort mellan studerande på olika utbildningsnivåer och har gjort så under lång tid. Ju högre nivå på studierna, desto högre lånebenägenhet. Det kan finnas olika förklaringar till att studerande på skilda nivåer i olika hög utsträckning väljer att ta lån. Valet skulle kunna ha med arbetsmarknadsutsikter att göra. De som studerar på högre nivåer kan ha en mer positiv förväntan på kommande arbete och inkomster än studerande på lägre nivåer. Även ålder kan spela roll, där studerande på lägre nivåer i genomsnitt är äldre än studerande på högre nivåer. Bakgrund kan vara en annan orsak, där framför allt studerande på grundskolenivå ofta har utländsk bakgrund.

Tabell 14. Lånebenägenhet fördelad efter studienivå

	Antal med bidrag	Antal med lån	Lånebenägenhet
Grundskolenivå (Sverige)	22 328	8 041	36,0 %
Gymnasienivå (Sverige)	118 609	65 211	55,0 %
Eftergymnasial nivå (Sverige)	340 483	251 109	73,8 %
Gymnasienivå (utland)	744	587	78,9 %
Eftergymnasial nivå (utland)	33 537	28 464	84,9 %
Samtliga ¹	483 671	330 291	68,3 %

¹ Nettosummering. En individ kan förekomma i flera grupper vilket medför att kategorin samtliga innehåller färre individer än summan av delkategorierna.

Det skiljer hela 38 procentenheter i lånebenägenhet mellan studerande med studiemedel på grundskolenivå och studerande på eftergymnasial nivå. Bland studerande på gymnasienivå är det lite mer än hälften som tar lån, vilket placerar studerandegruppen mittemellan studerande på grundskolenivå och studerande på eftergymnasial nivå.

En annan orsak till att en lägre andel tar lån på grundskole- och gymnasienivåerna skulle kunna vara att dessa studerande har möjlighet att få studiemedel med den högre bidragsnivån. Den som får mer bidrag kan ha mindre behov av lån. Något mer än hälften av studiemedelstagarna på grundskolenivå och ungefär en tredjedel av dem på gymnasienivå får studiemedel med högre bidrag (se även avsnitt 2.11 Bidragsnivå).

Tabell 15. Lånebenägenhet fördelad efter studie- och bidragsnivå

	Generellt bidrag	Högre bidrag
Grundskolenivå (Sverige)	40,3 %	33,9 %
Gymnasienivå (Sverige)	54,9 %	56,3 %

På grundskolenivå är lånebenägenheten högre bland dem som har högre bidrag än bland dem med generellt bidrag. På gymnasienivån är förhållandet dock det omvända. Lånebenägenheten för studerande på grundskolenivå med högre bidrag är också markant lägre än lånebenägenheten bland gymnasiestuderande. Bidragsnivån tycks alltså mindre betydelsefull för lånebenägenheten än utbildningsnivån.

Utlandsstuderande är de som oftast tar lån. Utlandsstuderande på gymnasial nivå tar oftare lån än studerande på eftergymnasial nivå i Sverige. När det gäller utlandsstudier kan det vara så att dessa ofta medför högre omkostnader som förutsätter låntagande. Möjligheterna att arbeta parallellt med studierna kan också vara mer begränsade än i Sverige.

2.10 Studiernas omfattning

Det är rimligt att anta att deltidsstuderande har större möjligheter än heltidsstuderande att arbeta parallellt med studierna. En del deltidsstuderande kan också ha arbetet som sin huvudsakliga sysselsättning och studierna som bisyssla. Deltidsstuderande kan därmed ha högre inkomster och mindre behov av lån.⁸

Tabell 16. Lånebenägenhet fördelad efter studieomfattning

	Antal med bidrag	Antal med lån	Lånebenägenhet
Deltid	65 272	35 401	54,2 %
Heltid	451 162	317 326	70,3 %
Samtliga ¹	483 671	330 291	68,3 %

¹ Nettosummering. En individ kan förekomma i båda grupperna vilket medför att kategorin samtliga innehåller färre individer än summan av delkategorierna.

Heltidsstuderande har en lånebenägenhet på 70,3 procent, vilket är 16 procentenheter högre än deltidsstuderande. Detta förhållande har i princip varit bestående över tid. Resultatet bekräftar således hypotesen att deltidsstuderande har mindre behov av att låna.

2.11 Generellt respektive högre bidrag

Studiemedlen har två möjliga bidragsnivåer, en generell och en högre. Studiemedel med den högre bidragsnivån kan ges till studerande på grundskole- eller gymnasienivå som saknar grundskole- eller gymnasiekompetens. Även studerande som repeterar på grundskolenivån kan få högre bidrag.⁹ Under 2017 motsvarar det generella bidraget 3 085 kronor i månaden och det högre bidraget 7 180 kronor i månaden. Ett högre bidragsbelopp kan ge vissa studerande större möjligheter att klara sig utan att ta lån.

Tabell 17. Lånebenägenhet fördelad efter bidragsnivå

	Antal med bidrag	Antal med lån	Lånebenägenhet
Generellt bidrag	444 199	310 636	69,9 %
Högre bidrag	43 967	22 699	51,6 %
Samtliga ¹	483 671	330 291	68,3 %

¹ Nettosummering. En individ kan förekomma i båda grupperna vilket medför att kategorin samtliga innehåller färre individer än summan av delkategorierna.

Studerande med det högre bidraget tar lån i lägre utsträckning än studerande som har det generella bidraget. En väsentlig del av skillnaden i lånebenägenhet tycks dock kunna förklaras av nivån på studierna snarare än bidragsnivån (se avsnitt 2.9 Nivå på studierna).

2.12 Inriktning på studierna (enbart universitet och högskola)

För studerande vid universitet och högskola finns uppgifter om inriktning på studierna samt om studierna bedrivs på ett program eller på en fristående kurs.

⁸ Deltidsstuderande har en genomsnittlig sammanräknad inkomst som är 20 000 kronor högre än heltidsstuderande.

⁹ Högre bidrag kan ges till studerande som är 25 år eller äldre och till vissa arbetslösa ungdomar i åldern 20-24 år.

Tabell 18. Lånebenägenhet fördelad efter fristående kurser eller programstudier

	Antal med bidrag	Antal med lån	Lånebenägenhet
Kurser	43 638	27 403	62,8 %
Program	254 805	192 992	75,7 %
Samtliga	298 443	220 395	73,8 %

Programstuderande tar oftare lån än studerande på fristående kurser. Lånebenägenheten på program överstiger lånebenägenheten på fristående kurser med 13 procentenheter. En bidragande orsak till denna skillnad är att kursstuderande i betydligt högre utsträckning än programstuderande bedriver sina studier på deltid (se även avsnitt 2.10 Studiernas omfattning).

Tabell 19. Lånebenägenhet fördelad efter inriktning på programstudier

	Antal med bidrag	Antal med lån	Lånebenägenhet
Allmän	203	143	70,4 %
Pedagogik och lärarutbildning	35 155	25 197	71,7 %
Humaniora och konst	13 567	10 955	80,7 %
Samhällsvetenskap, juridik, handel, administration	67 784	52 608	77,6 %
Naturvetenskap, matematik, data	18 027	13 305	73,8 %
Teknik, tillverkning	58 328	44 470	76,2 %
Lant- och skogsbruk, djursjukvård	2 890	2 345	81,1 %
Hälso- och sjukvård, social omsorg	53 317	39 788	74,6 %
Tjänster	5 534	4 181	75,6 %
Samtliga	254 805	192 992	75,7 %

Programstudier kan även fördelas efter inriktning på programmet. Denna fördelning visar att lånebenägenheten är högst, 81,1 procent, på utbildningar inom lantbruk, skogsbruk och djursjukvård. Lägst lånebenägenhet, 70,4 procent, finns på allmänna program. Det är svårt att säga vad skillnaderna mellan olika programinriktningar beror på. Exempelvis är andelen deltidstuderande relativt likartad för flera olika inriktningar, trots att lånebenägenheten varierar.

3. Sannolikhet att en studerande med studiemedel även tar studielån

3.1 Analysmetod – regression

För att kunna bedöma sannolikheten att en studerande med studiemedel även väljer att ta lån har CSN låtit Statistiska Centralbyrån (SCB) genomföra en analys av detta. SCB har genomfört analysen genom så kallad logistisk regressionsanalys. Förenklat uttryckt visar en sådan analys effekten av en förklarande variabel på sannolikheten att en viss händelse inträffar, i detta fall sannolikheten att en individ som har studiemedel även tar studielån. SCB:s analys presenteras i sin helhet i bilaga.

En av fördelarna med regressionsanalys är möjligheten att undersöka effekten av flera olika variabler samtidigt. I grunden handlar denna ansats om att försöka isolera effekten av en variabel, vilket blir möjligt om andra relevanta variabler kan tas i beaktande.

Analysmodellen byggdes upp genom att variabler lades till och togs bort tills den bästa modellen nåddes. De variabler som slutligen valdes att ingå i modellen var kön, ålder, svensk eller utländsk bakgrund, inkomst, familjesituation¹⁰, studiernas omfattning (heltid/deltid) och studienivå, inklusive om studierna bedrivs i Sverige eller utomlands.

Alla de faktorer som beskrivs i föregående kapitel finns således inte med i analysmodellen. Det främsta skälet till att variabler inte tagits med är att det saknas uppgifter för många individer i populationen. En sådan variabel är föräldrars utbildningsbakgrund.

Det var på förhand känt att kvinnor tenderar att studera i högre åldrar än män, att personer med svensk bakgrund tenderar att studera på högre utbildningsnivå än personer med utländsk bakgrund och att yngre personer tenderar att bo tillsammans med sina föräldrar i högre utsträckning än äldre. I dessa fall kan det uppstå så kallade interaktionseffekter där de olika variablerna interagerar med varandra. Därför används så kallade interaktionstermer i modellen. Det innebär att kombinationer av variabler analyseras tillsammans och att resultaten avser dessa kombinationer. För jämförelse används också en modell utan interaktioner.

3.2 Inkomst, nivå på studierna, studieomfattning och studerandes bakgrund är viktiga faktorer för att förklara lånebenägenheten

Analysen visar att ju lägre inkomstklass den studerande befinner sig i, desto högre är lånebenägenheten. Detta är förväntat eftersom högre inkomst normalt medför ett lägre behov av lån. När det gäller inkomst kan dock resultatet även vara en effekt av en strävan att slippa ta lån, där den som inte vill låna blir tvungen att arbeta för att skaffa sig inkomster. Detta bör föranleda en viss försiktighet i tolkningen av inkomstvariabeln.

Vidare är lånebenägenheten lägre för deltids- än för heltidsstuderande. Även detta är förväntat eftersom en studerande med lägre studietakt har mer tid att arbeta och skaffa sig sidoinkomster. Många som studerar på deltid har sannolikt studierna som bisyssla och arbete som huvudsysselsättning.

För de interagerande variablerna blir tolkningen mer komplicerad. För interaktionen mellan kön, åldersgrupp och familjesituation gäller bland annat att sammanboende män med barn är minst lånebenägna, följt av unga personer som bor hos sina föräldrar. Mest lånebenägna är ensamstående män över 35 år med barn, följt av ensamstående män mellan 25 och 34 år utan barn. Kvinnor är i regel mindre lånebenägna än män, med undantag för unga kvinnor som inte har barn, som är mer lånebenägna än unga män som inte har barn.

Störst könsskillnad märks i grupperna gifta/partners äldre än 35 år utan barn och sammanboende mellan 25 och 34 år med barn. I båda dessa grupper är kvinnor mycket mindre lånebenägna än män.

¹⁰ Det finns kvalitetsbrister när det gäller uppgift om familjesituation, se avsnitt 2.4. Det kunde därför ha varit en bra idé att stryka denna variabel ur analysen. Då den ser ut att bidra en del till att förklara utfallet har den dock inkluderats. Se bilaga för mer information.

För interaktionen mellan bakgrund och utbildningsnivå gäller att lånebenägenheten vid lika bakgrund i regel är högre ju högre utbildningsnivå det rör sig om. Vid lika utbildningsnivå är i regel lånebenägenheten högre bland inrikes än bland utrikes födda. Vidare gäller att personer som studerar i utlandet är mer lånebenägna än personer som studerar i Sverige.

Skattningarna visar att den enskilt viktigaste variabeln för att förklara lånebenägenhet är inkomst: Ju lägre inkomst desto högre lånebenägenhet. Den näst viktigaste variabeln är interaktionen mellan bakgrund och utbildningsnivå och den tredje viktigaste variabeln är huruvida det rör sig om heltids- eller deltidstudier. Interaktionen mellan kön, ålder och familjesituation bidrar till att förklara lånebenägenheten, men inte lika mycket som övriga studerade variabler. Resultaten för samtliga grupper i modellen med interaktionsvariabler återfinns i bilaga.

För jämförelse har en modell även tagits fram utan interaktionsvariabler, dvs. där varje variabel analyseras separat. Denna modell bedöms ha lite lägre förklaringsvärde än den modell som innehåller interaktionsvariabler.

När det gäller de variabler som i modellen med interaktionsvariabler inte ingick i någon interaktionsvariabel (inkomst och studietakt), är utfallet detsamma i denna alternativa modell. Vidare kan sägas att lånebenägenheten är högre ju högre utbildningsnivå det gäller, högre för inrikes än för utrikes födda, och aningen högre för män än för kvinnor. För ålder gäller att lånebenägenheten är högst i åldersintervallet 25–34 år och lägst bland dem som är yngre än 25 år. För familjesituation gäller att lånebenägenheten är högst för ensamstående utan barn och lägst för personer som bor hos sina föräldrar.

Tabell 20. Faktorer som påverkar lånebenägenhet, sorterade efter betydelse (modell utan interaktionsvariabler)

	Hög lånebenägenhet	Låg lånebenägenhet
1.	Hög utbildningsnivå (särskilt utlandsstudier)	Låg utbildningsnivå (särskilt studier i Sverige)
2.	Låg inkomst	Hög inkomst
3.	Heltidsstudier	Deltidsstudier
4.	Svensk bakgrund	Utländsk bakgrund
5.	Ensamstående utan barn	Bor hos föräldrar
6.	25-34 år	-24 år
7.	Män	Kvinnor

I den senare modellen, där variablerna redovisas separat, är utbildningsnivå den viktigaste variabeln för att förklara lånebenägenhet, följt av inkomst, studiernas omfattning, bakgrund, familjesituation, ålder och slutligen kön. Det går dock inte att säga något om lånebenägenheten hos exempelvis ensamstående unga män. Viss information går således förlorad om interaktioner inte tas i beaktande.

Referenser

CSN (2014). *Den ekonomiska situationen för studerande med barn*. CSN, rapport 2014:3. Dnr 2014-219-3658.

CSN (2014). *Skilda studieförutsättningar. En analys av studier, studieekonomi och hälsa utifrån föräldrarnas utbildningsbakgrund*. CSN, rapport 2014:8. Dnr 2014-219-6424.

CSN (2016). *Studerandes ekonomiska och sociala situation 2015*. CSN, rapport 2016:2. Dnr 2016-219-1333.

Prop. 2016/17:1. *Budgetpropositionen för 2017*. Utgiftsområde 15 Studiestöd.

Regleringsbrev för budgetåret 2016 avseende Centrala studiestödsnämnden.

Samuelsson, Jan (2000). *Islamisk ekonomi*. Studentlitteratur.

Sveriges Radio (2005). *Ränta på studielån problem för muslimer*.

<http://sverigesradio.se/sida/artikel.aspx?programid=83&artikel=747284> [2017-07-27]

Bilaga

Regressionsanalys CSN



Statistiska centralbyrån
Statistics Sweden

Regressionsanalys CSN

Bakgrund

Vi vill modellera sannolikheten att en individ som tar studiemedel även tar studielån. Vi kallar denna sannolikhet för p , och variabler som kan påverka den för X_k . Vi antar att sambandet mellan p och X_k är logistiskt, dvs. kan skrivas på formen

$$\log \frac{p}{1-p} = \beta_0 + \sum_k \beta_k X_k$$

där parametern β_k avgör hur mycket variabel k påverkar p .

Samspel mellan variabler modelleras genom att lägga till en interaktionsterm $\beta_{kl} X_k X_l$, där parametern β_{kl} avgör hur mycket samspelet mellan variablerna k och l påverkar p .

Valda variabler

Faktorerna inkomstklass, heltid, utländsk/svensk bakgrund, utbildningsnivå, kön, åldersintervall och familjesituation används.

Variabeln `csfvi_int` har nio nivåer, motsvarande olika storlek på sammanräknad förvärvsinkomst. Denna inkomstvariabel valdes för att den inte innehåller studiemedel, och därför inte skevas av den variabel vi vill förklara.

Nivåerna är numrerade 1 till 9, och motsvarar inkomsterna 0-49 999, 50 000-99 999, 100 000-149 999, 150 000-199 999, 200 000-249 999, 250 000-299 999, 300 000-399 999, 400 000-499 999 respektive 500 000 och uppåt.

Variabeln `Heltid` har nivån 1 vid heltids- och 0 vid deltidstudier.

Variabeln `UtlSvBakgAlt` har nivåerna 11, 12, 21, 22 och 23, som betecknar utrikes födda med två utrikes födda föräldrar, inrikes födda med två utrikes födda föräldrar, utrikes födda med minst en inrikes född förälder, inrikes födda med en inrikes och en utrikes född förälder, respektive inrikes födda med två inrikes födda föräldrar.

Variabeln `UtbNiva` har nivån 11 vid grundskolenivå i Sverige, 12 vid gymnasiala studier i Sverige, 13 vid eftergymnasiala studier i Sverige, 22 vid gymnasiala studier i utlandet, samt 23 vid eftergymnasiala studier i utlandet. Notera att samma individ kan förekomma i fler än en av dessa grupper, men inkluderas här enbart i den högsta utbildningsnivå som individen förekommer i.

Variabeln `kon` har nivån 1 för män och 2 för kvinnor.

Variabeln Aldersintervall har nivåerna 1, 2 och 3, som betecknar åldern upp till 24 år, 25 till 34 år, resp. 35 år och uppåt.

Variabeln Familjesituation har nivåerna 1 till 5, som betecknar gift, partner, sambo utan barn; ensamstående utan barn; bor hos föräldrar; gift, partner, sambo med barn; samt ensamstående med barn. Ett problem med denna variabel är dock att sammanboende personer utan barn inte kan relateras till varandra och därför står som ensamstående. Det kunde därför ha varit en bra idé att stryka denna variabel ur analysen, men då den ser ut att bidra en del till att förklara utfallet (se avsnittet Utvärdering av modell) inkluderades den.

Interaktionstermer som testas är UtlSvBakgAlt tillsammans med UtbNiva, samt kon tillsammans med Aldersintervall och Familjesituation. Detta för att det är känt på förhand att:

- 1) Kvinnor tenderar att studera i högre åldrar än män,
- 2) Personer med svensk bakgrund tenderar att studera på högre nivå i större utsträckning än personer med utländsk bakgrund, och
- 3) Yngre personer tenderar att bo hos sina föräldrar i högre utsträckning än äldre.

Två variabler som betecknar respektive föräldrars utbildningsbakgrund övervägdes för inklusion i analysen, men ströks då data om detta enbart finns bland personer med svensk bakgrund och därför skulle en stor andel individer såväl som en potentiellt viktig källa till information falla bort om dessa variabler inkluderades.

Analys

Modellen byggs upp med så kallad stegvis regression, där variabler läggs till och tas bort en i taget tills den bästa, enligt ett valt kriterium, modellen har skapats. Kriteriet som används för att välja modell är Akaike's informationskriterium (Akaike, 1974).

För att välja vilka variabler som läggs till eller tas bort används signifikansnivå, vilket innebär att den mest variabel som är mest signifikant för att förklara utfallet läggs till, och om någon variabel inte längre är signifikant efter att en annan variabel har lagts till, så tas denna bort.

Analysen utförs i SAS med metoden `proc hplogistic`.

Slutlig modell

Variablerna som analysen valde ut var inkomst, heltid, interaktionen mellan kön, åldersintervall och familjesituation, samt interaktionen mellan utländsk/svensk bakgrund och utbildningsnivå. Skattningar av parametrar tillsammans med medelfel och p -värden visas i Appendix A1.

För jämförelse används också en modell med samma variabler som den beskriven ovan, men utan interaktioner. Skattningar av parametrar tillsammans med medelfel och p -värden för denna modell visas i Appendix A2.

Diskussion

En sak är viktig att notera: När interaktionstermer har valts så är inte de enskilda variablerna som interagerar med i modellen. Ej heller är tvåvariabelinteraktioner med i det fall en trevariabelinteraktion är inkluderad. Detta kan bero på att interaktionen är så pass stark att de interagerande variablerna inte ensamt förklarar något utöver vad interaktionen förklarar, givet att en interaktion är med.

I övrigt kan vi se från parameterskattningarna i Appendix 1 att ju lägre inkomstklass desto högre lånebenägenhet. Vidare är lånebenägenheten lägre för deltids- än för heltidsstudenter. Detta är rimligt: Ju högre inkomst desto lägre behov av lån, och ju lägre studietakt desto mer tid åt sidoinkomster. Möjligen hade därför en interaktion mellan dessa två variabler kunnat tas med, men det hade inte tillfört något ytterligare (se avsnittet Utvärdering av modell).

För de interagerande variablerna blir tolkningen mer komplicerad än för de två ovan nämnda. För interaktionen mellan kön, åldersgrupp och familjesituation gäller bland annat att minst lånebenägna är sammanboende män med barn, följt av unga personer som bor hos sina föräldrar. Mest lånebenägna är ensamstående män över 35 år med barn, följt av ensamstående män mellan 25 och 34 år utan barn. Kvinnor är i regel mindre lånebenägna än män, med undantag för unga kvinnor som inte har barn, som är mer lånebenägna än unga män som inte har barn.

Störst könsskillnad märks i grupperna sammanboende personer över 35 år utan barn och sammanboende personer mellan 25 och 34 år med barn. I båda dessa grupper är kvinnor mycket mindre lånebenägna än män.

För interaktionen mellan bakgrund och utbildningsnivå gäller att för en fix bakgrund är lånebenägenheten i regel högre ju högre utbildningsnivå det rör sig om. För en fix utbildningsnivå är i regel lånebenägenheten högre bland inrikes än bland utrikes födda. Vidare gäller att personer som studerar i utlandet är mer lånebenägna än personer som studerar i Sverige. Detta oavsett bakgrund; personer med utländsk bakgrund som studerar i utlandet är i regel mer lånebenägna än personer utan utländsk bakgrund som studerar på motsvarande nivå i Sverige.

Allt detta följer av storleken och ordningen på parameterskattningarna. Slutligen kan vi se från skattningarna att den enskilt viktigaste variabeln för att förklara lånebenägenhet är inkomst: Ju lägre inkomst desto högre lånebenägenhet. Den näst viktigaste variabeln är interaktionen mellan bakgrund och utbildningsnivå, och den tredje viktigaste variabeln är huruvida det rör sig om heltids- eller deltidsstudier. Interaktionen mellan kön, ålder och familjesituation bidrar men inte lika mycket som övriga studerade variabler.

För modellen utan interaktioner kan konstateras samma saker som för den föregående modellen vad gäller de variabler som inte ingick i någon interaktion. Vidare kan sägas att lånebenägenheten är högre ju högre utbildningsnivå det gäller, högre för inrikes än för utrikes födda, och aningen högre för män än för kvinnor. För ålder gäller att lånebenägenheten är högst för det mellersta intervallet och lägst för det lägsta, och för familjesituation gäller att lånebenägenheten är högst för ensamstående utan barn och lägst för personer som bor hos sina föräldrar.

För den senare modellen är utbildningsnivå den viktigaste variabeln för att förklara lånebenägenhet, följt av inkomst, studiernas omfattning, bakgrund, familjesituation, ålder och

slutligen kön. Det går dock inte att säga något om lånebenägenheten hos exempelvis ensamstående unga män, och viss information går således förlorad om interaktioner inte tas i beaktande.

Utvärdering av modell

Ett vanligt sätt att utvärdera en modell är den omskalade bestämningskoefficienten, \tilde{R}^2 , som avgör hur stor andel av den beroende variabelns varians som förklaras av de oberoende variablerna. Den här modellen har ett värde på \tilde{R}^2 ungefär lika med 0,25. Detta innebär att cirka 25 % av variansen hos lånebenägenheten förklaras av de studerade variablerna, och det finns en del varians kvar att redogöra för. Dock kan en andel av den återstående variansen mycket väl bero enbart på slump, och det är svårt att avgöra hur mycket varians som det är önskvärt att redogöra för.

I en enklare jämförelsemodell där familjesituation inte inkluderades förklarades bara 18 % av variansen, och i en mer komplicerad jämförelsemodell där föräldrarnas utbildningsnivå inkluderades förklarades bara 23 % av variansen, varför den valda modellen är att föredra. I ytterligare en jämförelsemodell lades interaktion mellan studietakt och inkomst till. Detta höjde inte \tilde{R}^2 märkbart, på bekostnad av ökad komplexitet i modellen. Denna interaktion är därför inte nödvändig eller för den delen rekommenderad att ta med.

Slutligen, i modellen utan interaktioner förklarades drygt 23 % av variansen. Denna modell hade därför till synes kunnat räcka, men är ofullständig då på förhand kända interaktioner inte tas hänsyn till i den, och otillräcklig av anledningar nämnda under Diskussion.

Måttet på förklaringsgrad som har betraktats i det här avsnittet kan förstås som ett mått på hur stor andel av variationen hos lånebenägenheten som förklaras av de variabler som ingår i modellerna. Detta mått är det visserligen bra att inte stirra sig alltför blind på, men det har ett värde i synnerhet när det gäller jämförelser mellan modeller: Ett högre värde är i regel bättre än ett lägre. Å andra sidan kan en alltför komplicerad modell ”råka” förklara variation som egentligen bara beror på slump, och det är inte alltid säkert att ett högt värde är önskvärt. Däremot kan det nog vara säkert att säga att ett värde på 25 % är aningen bättre än ett på 23 %, då 25 % inte är särskilt misstänkt högt.

Referenser

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 716-723.

Appendix: Output från SAS

I huvudtabellerna nedan visas bland annat parameterskattningar med motsvarande medelfel och p -värden för de olika kategoriska variablerna som är inkluderade i modellerna.

Parameterskattningarna ska förstås relativt mot varandra och basnivån, som alltid anges som 0 och utan medelfel eller p -värde. En negativ skattning för en nivå på en variabel innebär lägre lånebenägenhet för nivån än för basnivån, och en positiv skattning innebär högre lånebenägenhet. Om en nivå har en högre skattning än en annan innebär detta att den första nivån har högre lånebenägenhet än den andra. Basnivån är godtyckligt vald och spelar i princip bara roll för jämförelse.

Medelfelet anger osäkerheten i skattningen, och p -värdet anger sannolikheten att parametern egentligen har värdet 0, givet skattning och medelfel. Låga p -värden är önskvärda och indikerar att nivån kan urskiljas från basnivån och därför även att variabeln bidrar till utfallet.

Utöver detta visas i tabellerna även antal frihetsgrader och t -värde. Att antalet frihetsgrader är oändligt innebär att variablerna antas vara normalfördelade, och t -värdet är värdet på normalfördelningens täthetsfunktion för den aktuella variabelnivån.

A1: Modell med interaktioner

Selection Information			
Selection Method	Stepwise		
Select Criterion	Significance Level		
Stop Criterion	Significance Level		
Choose Criterion	AIC		
Effect Hierarchy Enforced	None		
Entry Significance Level (SLE)	0.05		
Stay Significance Level (SLS)	0.05		
Stop Horizon	1		

Selection Summary			
Step	Effect Entered	Number Effects In	AICp Value
0	Intercept	1	596916 .
1	kon*Aldersi*Familjes	2	552913 <.0001
2	UtlSvBakgAlt*UtbNiva	3	517919 <.0001
3	csfvi_int	4	509877 <.0001
4	Heltid		5504283* <.0001

Fit Statistics	
-2 Log Likelihood	504157
AIC (smaller is better)	504283
AICC (smaller is better)	504283
BIC (smaller is better)	504981
R-Square	0.1766
Max-rescaled R-Square	0.2475

Parameter Estimates					
Parameter	Estimate	Standard Error	DF	t Value	Pr > t
Intercept	1.0923	0.1536	Infty	7.11	<.0001
csfvi_int 1	1.8869	0.1501	Infty	12.57	<.0001
csfvi_int 2	1.9257	0.1502	Infty	12.82	<.0001
csfvi_int 3	1.7503	0.1503	Infty	11.65	<.0001
csfvi_int 4	1.5602	0.1504	Infty	10.37	<.0001
csfvi_int 5	1.2279	0.1506	Infty	8.15	<.0001
csfvi_int 6	0.8076	0.1514	Infty	5.34	<.0001
csfvi_int 7	0.5633	0.1529	Infty	3.68	0.0002
csfvi_int 8	0.5036	0.1734	Infty	2.90	0.0037
csfvi_int 9	0
Heltid 0	-0.9967	0.0134	Infty	-74.30	<.0001
Heltid 1	0
UtlSvBakgAlt*UtbNiva 11 11	-3.5693	0.0329	Infty	-108.27	<.0001
UtlSvBakgAlt*UtbNiva 11 12	-2.7166	0.0256	Infty	-105.79	<.0001
UtlSvBakgAlt*UtbNiva 11 13	-1.7219	0.0256	Infty	-67.16	<.0001
UtlSvBakgAlt*UtbNiva 11 22	-1.1106	0.3010	Infty	-3.69	0.0002
UtlSvBakgAlt*UtbNiva 11 23	-0.3877	0.0517	Infty	-7.49	<.0001
UtlSvBakgAlt*UtbNiva 12 11	-1.6841	0.1651	Infty	-10.20	<.0001
UtlSvBakgAlt*UtbNiva 12 12	-2.0909	0.0351	Infty	-59.42	<.0001
UtlSvBakgAlt*UtbNiva 12 13	-1.8042	0.0263	Infty	-68.49	<.0001
UtlSvBakgAlt*UtbNiva 12 22	-0.7519	0.3248	Infty	-2.31	0.0206
UtlSvBakgAlt*UtbNiva 12 23	-0.0419	0.0524	Infty	-0.80	0.4237
UtlSvBakgAlt*UtbNiva 21 11	-1.7171	0.5390	Infty	-3.19	0.0014
UtlSvBakgAlt*UtbNiva 21 12	-1.8145	0.1034	Infty	-17.55	<.0001
UtlSvBakgAlt*UtbNiva 21 13	-1.1938	0.0585	Infty	-20.39	<.0001
UtlSvBakgAlt*UtbNiva 21 22	-0.6581	1.1683	Infty	-0.56	0.5733
UtlSvBakgAlt*UtbNiva 21 23	-0.5342	0.1193	Infty	-4.48	<.0001
UtlSvBakgAlt*UtbNiva 22 11	-1.3734	0.1862	Infty	-7.37	<.0001
UtlSvBakgAlt*UtbNiva 22 12	-1.5384	0.0322	Infty	-47.65	<.0001
UtlSvBakgAlt*UtbNiva 22 13	-1.1042	0.0263	Infty	-41.91	<.0001
UtlSvBakgAlt*UtbNiva 22 22	-0.9416	0.2745	Infty	-3.43	0.0006
UtlSvBakgAlt*UtbNiva 22 23	-0.1214	0.0519	Infty	-2.34	0.0194
UtlSvBakgAlt*UtbNiva 23 11	-1.4008	0.0901	Infty	-15.54	<.0001
UtlSvBakgAlt*UtbNiva 23 12	-1.5412	0.0243	Infty	-63.36	<.0001
UtlSvBakgAlt*UtbNiva 23 13	-1.0168	0.0230	Infty	-44.21	<.0001
UtlSvBakgAlt*UtbNiva 23 22	-0.5356	0.1534	Infty	-3.49	0.0005
UtlSvBakgAlt*UtbNiva 23 23	0
kon*Aldersi*Familjes 1 1 1	-0.3016	0.1085	Infty	-2.78	0.0054
kon*Aldersi*Familjes 1 1 2	-0.1320	0.0293	Infty	-4.50	<.0001
kon*Aldersi*Familjes 1 1 3	-1.6820	0.0291	Infty	-57.62	<.0001
kon*Aldersi*Familjes 1 1 4	-1.7348	0.0321	Infty	-53.89	<.0001
kon*Aldersi*Familjes 1 1 5	-1.4938	0.0436	Infty	-34.22	<.0001

Parameter Estimates					
Parameter	Estimate	Standard Error	DF	t Value	Pr > t
kon*Aldersi*Familjes 1 2 1	-0.04507	0.05915	Infty	-0.760	.4461
kon*Aldersi*Familjes 1 2 2	0.3388	0.03027	Infty	11.19	<.0001
kon*Aldersi*Familjes 1 2 3	-0.6501	0.03446	Infty	-18.87	<.0001
kon*Aldersi*Familjes 1 2 4	0.2055	0.03839	Infty	5.35	<.0001
kon*Aldersi*Familjes 1 2 5	0.2713	0.1042	Infty	2.600	.0092
kon*Aldersi*Familjes 1 3 1	-0.05846	0.08514	Infty	-0.690	.4923
kon*Aldersi*Familjes 1 3 2	0.2297	0.04247	Infty	5.41	<.0001
kon*Aldersi*Familjes 1 3 3	-0.3804	0.1098	Infty	-3.460	.0005
kon*Aldersi*Familjes 1 3 4	-0.08638	0.03845	Infty	-2.250	.0247
kon*Aldersi*Familjes 1 3 5	0.3975	0.1011	Infty	3.93	<.0001
kon*Aldersi*Familjes 2 1 1	-0.8604	0.05631	Infty	-15.28	<.0001
kon*Aldersi*Familjes 2 1 2	-0.2588	0.02866	Infty	-9.03	<.0001
kon*Aldersi*Familjes 2 1 3	-1.6570	0.02915	Infty	-56.84	<.0001
kon*Aldersi*Familjes 2 1 4	-1.6239	0.03139	Infty	-51.74	<.0001
kon*Aldersi*Familjes 2 1 5	-1.0612	0.03900	Infty	-27.21	<.0001
kon*Aldersi*Familjes 2 2 1	-0.5905	0.04349	Infty	-13.58	<.0001
kon*Aldersi*Familjes 2 2 2	0.1380	0.02965	Infty	4.66	<.0001
kon*Aldersi*Familjes 2 2 3	-0.6473	0.03612	Infty	-17.92	<.0001
kon*Aldersi*Familjes 2 2 4	-0.4952	0.02997	Infty	-16.53	<.0001
kon*Aldersi*Familjes 2 2 5	0.1771	0.03950	Infty	4.48	<.0001
kon*Aldersi*Familjes 2 3 1	-0.9502	0.05174	Infty	-18.36	<.0001
kon*Aldersi*Familjes 2 3 2	-0.2811	0.03933	Infty	-7.15	<.0001
kon*Aldersi*Familjes 2 3 3	-0.4497	0.1308	Infty	-3.440	.0006
kon*Aldersi*Familjes 2 3 4	-0.8293	0.03007	Infty	-27.58	<.0001
kon*Aldersi*Familjes 2 3 5	0

A2: Modell utan interaktioner

Selection Information	
Selection Method	Stepwise
Select Criterion	Significance Level
Stop Criterion	Significance Level
Choose Criterion	AIC
Effect Hierarchy Enforced	None
Entry Significance Level (SLE)	0.05
Stay Significance Level (SLS)	0.05
Stop Horizon	1

Selection Summary			
Step	Effect Entered	Number Effects In	AIC p Value
0	Intercept	1	596916 .
1	Familjesituation	2	563222 <.0001
2	UtbNiva	3	539071 <.0001
3	UtlSvBakgAlt	4	530492 <.0001
4	Heltid	5	523671 <.0001
5	Aldersintervall	6	515600 <.0001
6	csfvi_int	7	510275 <.0001
7	kon	8	510054* <.0001

Fit Statistics	
-2 Log Likelihood	510004
AIC (smaller is better)	510054
AICC (smaller is better)	510054
BIC (smaller is better)	510331
R-Square	0.1664
Max-rescaled R-Square	0.2332

Parameter Estimates				
Parameter	Estimate	Standard Error	DF	t Value Pr > t
Intercept	1.0269	0.1521	Infty	6.75 <.0001
csfvi_int 1	1.7991	0.1507	Infty	11.94 <.0001
csfvi_int 2	1.8554	0.1508	Infty	12.31 <.0001
csfvi_int 3	1.6919	0.1508	Infty	11.22 <.0001
csfvi_int 4	1.5099	0.1509	Infty	10.00 <.0001
csfvi_int 5	1.1792	0.1512	Infty	7.80 <.0001
csfvi_int 6	0.7670	0.1520	Infty	5.05 <.0001
csfvi_int 7	0.5324	0.1535	Infty	3.47 0.0005
csfvi_int 8	0.5015	0.1744	Infty	2.88 0.0040

Parameter Estimates				
Parameter	Estimate	Standard Error	DFt	Value Pr > t
csfvi_int 9	0	.	.	.
Heltid 0	-1.0023	0.01338	Infty	-74.90<.0001
Heltid 1	0	.	.	.
UtbNiva 11	-2.5966	0.02859	Infty	-90.82<.0001
UtbNiva 12	-1.7373	0.01873	Infty	-92.75<.0001
UtbNiva 13	-1.1231	0.01791	Infty	-62.70<.0001
UtbNiva 22	-0.6269	0.1143	Infty	-5.49<.0001
UtbNiva 23	0	.	.	.
UtlSvBakgAlt 11	-0.8936	0.00971	Infty	-92.01<.0001
UtlSvBakgAlt 12	-0.7156	0.01273	Infty	-56.22<.0001
UtlSvBakgAlt 21	-0.2517	0.04419	Infty	-5.70<.0001
UtlSvBakgAlt 22	-0.06967	0.01235	Infty	-5.64<.0001
UtlSvBakgAlt 23	0	.	.	.
kon 1	0.1070	0.00717	Infty	14.91<.0001
kon 2	0	.	.	.
Aldersintervall 1	-0.6523	0.01234	Infty	-52.87<.0001
Aldersintervall 2	0.1665	0.01208	Infty	13.78<.0001
Aldersintervall 3	0	.	.	.
Familjesituation 1	-0.3143	0.02465	Infty	-12.75<.0001
Familjesituation 2	0.4992	0.01557	Infty	32.07<.0001
Familjesituation 3	-0.7559	0.01627	Infty	-46.46<.0001
Familjesituation 4	-0.5278	0.01563	Infty	-33.77<.0001
Familjesituation 5	0	.	.	.