



SCHOOL OF ECONOMICS AND MANAGEMENT

Nationalekonomiska institutionen

Examensarbete kandidatnivå, NEKH02

Augusti 2025

Kvinnliga förebilder i akademien

En kvantitativ studie av sambandet mellan manliga studenters examensfrekvenser och könsfördelningen i fakulteten i amerikansk högre utbildning

Författare:

Omar Ismailov

Handledare:

Sarah Rosenberg

Abstract

This thesis investigates the relationship between male graduation rates and institutional characteristics in U.S. higher education, focusing on faculty gender composition, the share of students receiving financial aid, and institutional level. Using panel data from 2021-2023, the study applies pooled OLS and fixed effects regression models. No statistically significant relationships were found in the fixed effects model, while pooled OLS indicated negative associations between male graduation rates and both faculty gender composition and financial aid. As the results vary across models, they should be interpreted with caution and may primarily capture differences between institutions rather than underlying relationships.

Innehållsförteckning

Abstract	1
Innehållsförteckning	2
1. Inledning	4
1.1 Introduktion.....	4
1.2 Syfte och frågeställning	4
1.3 Metodöversikt och avgränsningar	5
1.4 Disposition	7
2. Tidigare forskning	8
2.1 Skillnader i utbildningsresultat bland manliga och kvinnliga studenter	8
2.2 Fakultetens könsfördelning och studenters prestation	10
2.3 Kvinnliga förebilder i akademien	11
2.4 Faktorer som påverkar examensfrekvens.....	12
3. Teoretiskt ramverk	14
3.1 Role model theory	14
4. Data och variabler	16
4.1 Dataöversikt och variabelkonstruktion	16
4.2 Beroende variabel	17
4.3 Andel kvinnlig fakultet	17
4.4 Kontrollvariabler	18
4.5 Datamaterialets begränsningar	18
4.6 Deskriptiv statistik	20
5. Metod	22
5.1 Diagnostiska tester	22
5.2 Motivering av metod.....	22
5.3 Regressionsmodeller	24
5.4 Hypotes	24
6. Resultat	26
6.1 Resultat av diagnostiska tester	26
6.1.1 Variance Inflation Factor (VIF)-test	26
6.1.2 Wooldridge-test.....	26
6.1.3 Hausman-test.....	27
6.2 Pooled OLS-modell	28
6.3 Fixed effects-modell	30
6.4 Skillnader mellan pooled OLS och fixed effects	32
6.5 Modellens begränsningar	32
7. Diskussion	34

7.1 Jämförelse med tidigare studier	35
7.2 Begränsningar och framtida forskning.....	36
8. Slutsats	39
Referenser	40
Redogörelse för användning av generativ AI.....	43

1. Inledning

1.1 Introduktion

Andelen män i USA som både påbörjar och avslutar en högskoleutbildning har under en längre tid minskat år efter år (Reeves, 2021). I dagsläget rapporteras det att cirka 40 procent av alla universitetsstudenter är män, och sett till fyraåriga universitet är det 59 procent av dessa som får ut sitt diplom inom sex år, till skillnad från 65 procent av kvinnliga studenter (Wall Street Journal, 2021).

Att män halkar efter i båda måtten får konsekvenser både på individ- och samhällsnivå. Schneider och Yin (2011) uppskattar att de förlorade inkomsterna för bara en årskull som påbörjade studierna hösten 2002 men aldrig tog examen uppgick till 3,8 miljarder dollar. Detta medförde även hundratals miljoner i uteblivna skatteintäkter. Vidare diskuterar The Economist (2024) en växande social och ekonomisk klyfta mellan unga män och kvinnor, där könsskillnader i utbildningsnivå lyfts som en bidragande faktor.

En möjlig, men relativt outforskad, bidragande faktor är hur könsfördelning i fakulteten påverkar examensfrekvensen bland män. Tidigare forskning har visat att kvinnliga förebilder spelar särskild roll för kvinnor (Sav, 2012), men det är mindre känt vilken effekt de har för män. American Association of University Professors (AAUP, 2020) rapporterade att kvinnor 2018 utgjorde 46,7 procent av heltidsanställda fakultetsmedlemmar vid amerikanska universitet, men endast 32,5 procent av professorer. Detta visar att jämställdhet på alla nivåer inom fakulteten ännu inte har uppnåtts.

1.2 Syfte och frågeställning

Syftet med denna uppsats är att undersöka om det finns ett samband mellan examensfrekvens bland manliga studenter och andelen kvinnor i fakulteten vid amerikanska lärosäten. Även institutionernas nivå (tvååriga eller fyraåriga) samt andelen studenter som får statligt ekonomiskt

stöd beaktas. Analysen genomförs med hjälp av en pooled OLS-modell (ordinary least squares) och en fixed effects-modell (FE). Dessa estimeras på data tagen från Integrated Postsecondary Education Data System (IPEDS), som administreras av National Center for Education Statistics (NCES).

Ambitionen med uppsatsen är att bidra till den pågående diskussionen om mäns fallande examensfrekvenser genom att tillföra ett alternativt perspektiv, där även den nuvarande ojämna könsfördelningen inom fakulteten inkluderas. Eftersom det finns en del tidigare forskning inom området bygger uppsatsen på en kombination av tidigare studier och en egen kvantitativ analys. Även om flera studier har belyst betydelsen av kvinnliga förebilder för kvinnliga studenters utfall (se t.ex. Bettinger & Long, 2005; Robst et al. 1998), har få undersökt om en högre andel kvinnor i fakulteten också har en effekt på manliga studenters examensfrekvens vid amerikanska college och universitet på grundutbildningsnivå.

Analysen bygger på paneldata från läsåren 2021-2023. Genom att följa enskilda lärosäten över tre läsår undersöks hur förändringar i fakultetens könssammansättning, samt andra faktorer inom samma lärosäte, över tid relaterar till examensfrekvensen bland män. Utifrån detta formuleras följande forskningsfråga: *I vilken utsträckning hänger examensfrekvensen bland manliga studenter samman med andelen kvinnor i fakulteten, med hänsyn till institutionernas nivå och studenternas tillgång till ekonomiskt stöd?*

1.3 Metodöversikt och avgränsningar

Först genomförs en serie diagnostiska tester för att stärka validiteten i regressionsresultaten. Dessa inkluderar ett Variance Inflation Factor (VIF)-test för att kontrollera multikollinearitet, ett Wooldridge-test för att undersöka autokorrelation samt ett Hausman-test för att avgöra om en fixed effects- eller random effects-modell är mest lämplig.

Därefter inleds analysen med en pooled OLS-modell (ordinary least squares). Denna modell tar inte hänsyn till enhetsspecifika effekter men ger en grundläggande uppskattning av sambanden mellan variablerna. Därmed används resultaten från denna modell som jämförelse med fixed

effects-modellen för att bedöma hur sambanden påverkas när man tar hänsyn till unika egenskaper hos varje lärosäte.

Analysen bygger därefter vidare på en fixed effects-modell där den beroende variabeln är examensfrekvens bland män och den primära oberoende variabeln är andelen av fakulteten som är kvinnor. En ytterligare oberoende variabel inkluderas som kontrollvariabel, denna är andelen studenter som får statliga bidrag. Det finns många olika faktorer som påverkar examensfrekvenser, men analysen fokuserar på ett begränsat antal variabler för att möjliggöra en mer fokuserad undersökning av de centrala sambanden. Urvalet av variabler har även påverkats av datamaterialets begränsningar, där vissa potentiellt relevanta faktorer inte kunnat inkluderas på grund av brist på tillgänglig och konsekvent rapporterad data.

Regressionen genomförs på ett brett urval av amerikanska lärosäten från hela landet eftersom tillgänglig och omfattande data möjliggör en bred kvantitativ analys. USA är även lämpligt då det finns ett stort antal institut med hög variation i både studenter och institutionstyper, samtidigt som de alla verkar inom samma utbildningssystem. Detta skapar förutsättningar för en generell analys på ämnet. Vidare finns det etablerad forskning kring amerikansk högre utbildning vilket underlättar vid jämförande av resultaten med tidigare studier.

Analysen baseras på data från tre läsår, 2021 till 2023, vilket är det senaste tillgängliga i IPEDS-databasen vid uppsatsens skrivande. Fokus ligger uteslutande på grundutbildningsnivån, det vill säga två- och fyåriga examina. Med data som sträcker sig över tre år är ambitionen att undersöka hur förändringar i relevanta faktorer inom skolor kan återspeglas i männens examensfrekvens. Tidigare forskning har tydligt dokumenterat hur både fakultetens könssammansättning och mäns examensfrekvenser har förändrats över tid, vilket motiverar en närmare analys av sambandet mellan dessa faktorer. Utifrån resultatet skulle framtida forskning kunna bygga vidare på detta genom att inkludera flera läsår eller undersöka specifika perioder där det skett mer betydande skiften i de undersökta variablerna.

1.4 Disposition

Kapitel 1 introducerar ämnet och motiverar studiens relevans. Här formuleras syftet och forskningsfrågan samt en översikt av metoden ges. Kapitel 2 sammanfattar tidigare forskning inom de huvudområden som undersöks. Vidare presenterar kapitel 3 det teoretiska ramverket för uppsatsen, *Role model theory*. Kapitel 4 beskriver datamaterialet som analysen bygger på, där de variabler som används presenteras och definieras tillsammans med deras avgränsningar. Kapitel 5 redogör för metoden och de regressionsmodeller som används, närmare bestämt pooled OLS- och fixed effects-modellen. Kapitel 6 presenterar resultaten från analyserna och sammanfattar de huvudsakliga mönstren från respektive modell. Slutligen diskuteras dessa resultat i kapitel 7 i relation till tidigare forskning och teori. Här lyfts även studiens begränsningar och implikationer. Uppsatsen avslutas med en slutsats som sammanfattar huvudresultaten och ger förslag till framtida forskning. Referenser redovisas i slutet av uppsatsen.

2. Tidigare forskning

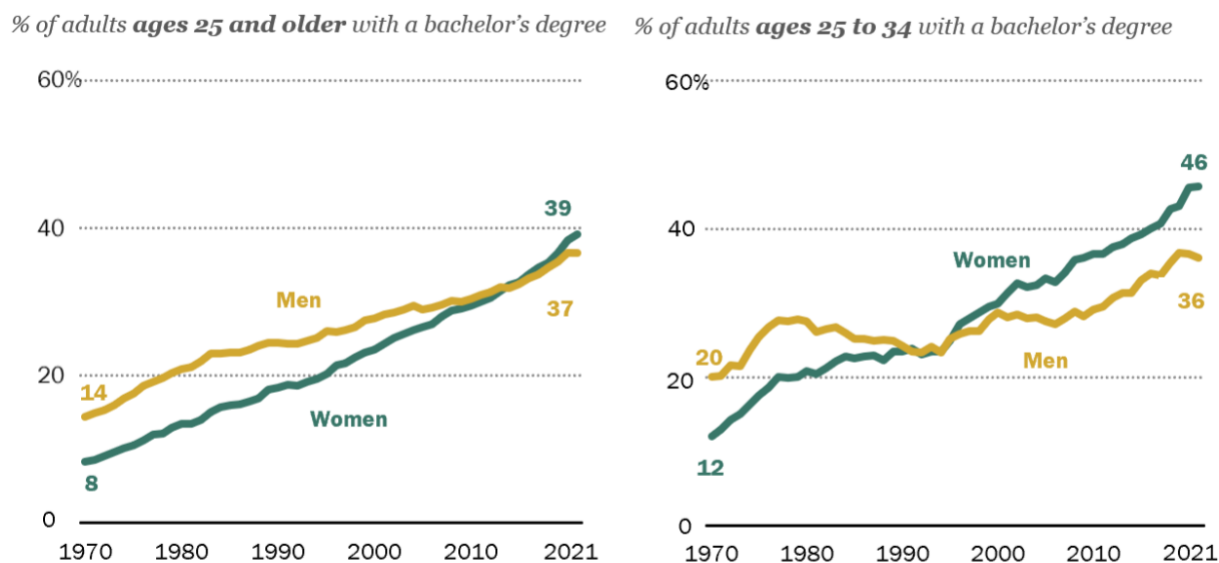
2.1 Skillnader i utbildningsresultat bland manliga och kvinnliga studenter

Efter andra världskriget var kvinnors deltagande och examensfrekvens i högre utbildning avsevärt lägre än männens. Obalansen nådde sin topp 1947 då det gick 2,3 manliga studenter för varje kvinnlig student, och män hade i genomsnitt cirka 10 procent högre examensfrekvens. Från den punkten framåt skiftade trenden till att allt fler kvinnor började studera vid universitet, och denna trend har i stort sett varit obruten sedan dess. Under 1980-talet uppnåddes könsparitet, då kvinnor och män utgjorde ungefär lika stora andelar av alla studenter. Men utvecklingen stannade inte där utan andelen kvinnor inom högre utbildning fortsatte att öka år efter år.

Goldin et al. (2006) analyserar denna utveckling och diskuterar vilka faktorer som bidrog till trendbrottet samt den fortsatta ökningen av kvinnors utbildningsnivå. Bland annat lyfts förändringar på arbetsmarknaden och omställningar i könsroller som centrala orsaker till skiftet som ägde rum. Författarna framhäver två faktorer vilka ledde till att högre utbildning blev mer lönsamt för kvinnor. Kvinnor började arbeta under större delar av sina liv då de i viss mån kunde frigöra sig från rollen som enbart ansvariga för barnomsorg. Dessutom började kvinnor etablera sig i traditionellt manligt dominerade yrken, vilket ytterligare ökade avkastningen på högre utbildning.

Bailey och Dynarski (2011) betonar i sin analys hur inkomstjämlighet har bidragit till könsskillnader i utbildningsresultat. De uppmärksammar samma trender som Goldin et al. (2006), både vad gäller förändringen i könsfördelningen bland universitetsstudenter och skillnader i examensfrekvenser mellan kvinnor och män. Dock lyfter de att samtidigt som detta skifte ägt rum har det totala antalet personer som tar en universitetsexamen ökat betydligt. Även om andelen som tar en universitetsexamen har stigit kraftigt, har inte alla grupper gynnats i samma utsträckning. Ökningen har främst kommit till höginkomstgrupper, medan låginkomsttagare i stor utsträckning har stått utanför denna utveckling.

Mäns utbildningsnivå har visserligen ökat något, trots att deras andel av studentpopulationen minskat. Den största ökningen gäller kvinnor, och enligt Bailey och Dynarski har denna i stor utsträckning drivits av kvinnor från höginkomstfamiljer. Buchmann et al. (2025) tillför ett ytterligare perspektiv på detta förlopp. Deras analys fokuserar på 2000-talet, där de visar hur kvinnors ökade närvaro inom högre utbildning samtidigt har lett till att kvinnor i oproportionerligt hög grad bär på studieskulder. De pekar på att när det är fler kvinnor som utbildar sig, faller också en större del av ansvaret för utbildningens expansion på dem, vilket medför en ökad ekonomisk risk. Genom detta skapas en obalans som stärker argument för att höja männens examensfrekvenser.



Source: Pew Research Center analysis of Current Population Survey Annual Social and Economic Supplement (IPUMS).

PEW RESEARCH CENTER

Figur 1. Andel vuxna i USA med kandidatexamen, 1970–2021, uppdelat på kön.

Figuren visar hur kvinnor gradvis har passerat män i utbildningsnivå. Till vänster visas andelen i befolkningen 25 år och äldre, till höger andelen i åldersgruppen 25–34 år (Pew Research Center, 2021).

2.2 Fakultetens könsfördelning och studenters prestation

Tidigare studier visar att könssammansättningen inom fakulteten inte bara är en fråga om representation utan också påverkar de faktiska utfallen i studenters utbildning. AAUP (2020) ger en överblick av hur demografin inom den amerikanska akademien ser ut idag och hur den har utvecklats över tid. Under de senaste decennierna har andelen kvinnor ökat avsevärt, då de har gått från att utgöra 31,6 procent av de heltidsanställda fakultetsmedlemmarna 1991 till 46,7 procent 2018. Dock framhäver de flera aspekter som komplicerar tolkningen.

I den amerikanska kontexten skiljer man mellan *tenure-track* och *non-tenure-track*-tjänster. *Tenure-track* innebär en karriärväg som kan leda till tillsvidareanställning (*tenure*) och befordran till högre akademiska positioner, inklusive professor. *Non-tenure-track* däremot avser tidsbegränsade och osäkrare anställningar, ofta med fokus på undervisning och med lägre lön och svagare anställningstrygghet. Kvinnor är enligt AAUP fortfarande överrepresenterade i *non-tenure-track*-positioner och deltidsundervisning, vilka kännetecknas av svagare anställningstrygghet och lägre lön. Detta återspeglas i att kvinnor fortfarande har svårare att nå de högsta akademiska positionerna, då enbart 32,5 procent av professorerna är kvinnor. AAUP lyfter även fram löneskillnader mellan könen samt den kraftiga underrepresentationen av kvinnor från minoritetsgrupper. Dessa frågor är viktiga men ligger utanför ramen för denna uppsats.

Hur könsfördelningen påverkar studenters prestation diskuteras av Sav (2012). Med data från 2005 till 2009 finner studien att en högre andel kvinnlig fakultet med *tenure*-anställningar, vilket är den kategorin AAUP visade att kvinnor var underrepresenterade inom, medför en ökning för både manliga och kvinnliga studenters examensfrekvenser. Däremot visades effekten av fler kvinnor i *non-tenure-track*-positioner vara olika för manliga jämfört med kvinnliga studenter. Mäns examensfrekvenser minskade medan kvinnors ökade, författaren betonar dock att detta senare resultat bör tolkas med försiktighet.

I andra mått på studenters prestation undersöker Robst et al (1998) sambandet mellan andelen kvinnliga fakultetsmedlemmar och retention, särskilt för kvinnliga studenter. Studien baseras på data från ett universitet i New York där retentionen ligger långt över det nationella genomsnittet. Deras resultat visar att kvinnliga studenter i högre grad kommer tillbaka till det andra studieåret i

kurser där läraren också är kvinna. Effekten syns särskilt tydligt i de klasser där andelen kvinnliga studenter är betydligt färre, vilket i de flesta fall gäller tekniska utbildningar.

2.3 Kvinnliga förebilder i akademien

Att stärka kvinnors representation i akademien har länge varit ett viktigt mål, drivet av en vilja att bryta de historiska mönster av könsdiskriminering som under lång tid begränsade kvinnors möjligheter till högre utbildning och akademiska positioner. I takt med att samhället rört sig bort från denna exkluderande struktur har forskningen allt mer intresserat sig för vilken roll kvinnor idag har inom akademien och vilka hinder som fortfarande finns kvar.

Carrell et al. (2009) undersöker betydelsen kvinnliga professorer har, särskilt inom naturvetenskapliga ämnen, och finner en tydlig positiv inverkan på kvinnliga studenters studieresultat, men mycket liten påverkan på manliga studenters. För kvinnliga studenter kvarstod sambandet ända fram till examen, då sannolikheten att ta ut examen ökade i närvaro av fler kvinnliga professorer. Denna studie är intressant att ställa i kontrast till föreliggande uppsats som i stället undersöker om andelen kvinnor i fakulteten har någon inverkan på manliga studenters examensfrekvenser. Resultaten hos Carrell et al. indikerar att könssammansättningen kan spela en roll, men att effekterna kan vara könsspecifika. Om det finns samband i denna uppsats kan de därför spegla andra mekanismer än de som identifierats i Carrell et al., vilket väcker frågor om vilka faktorer som faktiskt driver skillnaderna. Detta bör dock beaktas med hänsyn till att studierna genomförts i olika kontexter, både vad gäller ämnesområden och populationer.

Vidare undersökte Bettinger och Long (2005) påverkan av kvinnliga instruktörer på kvinnliga studenters ämnesval och intresse, mätt bland annat genom antalet fortsatta kurser som togs inom ämnet. De fann blandade resultat. Kvinnliga instruktörer ökade intresset för vissa ämnen, särskilt matematik/statistik, geologi och sociologi. I flera manligt dominerade ämnen, till exempel fysik och ekonomi, noterades i stället negativa eller inga effekter. Författarna drog slutsatsen att kvinnliga lärare kan fungera som förebilder för kvinnliga studenter, men även att effekten varierar kraftigt mellan ämnen. Studien inkluderade även en mindre omfattande analys av

manliga studenter i kvinnodominerade ämnen, där resultaten visade att manliga instruktörer hade en stark positiv inverkan inom lärarutbildningar. Däremot undersökte de inte i någon större utsträckning hur instruktörer påverkar studenter av motsatt kön.

2.4 Faktorer som påverkar examensfrekvens

Som tidigare nämnts finns det ett stort antal faktorer som kan påverka om en student fullföljer sin examen. Studenttiden präglas av både akademiska och personliga förändringar som kan påverka studieresultaten på olika sätt. I detta sammanhang blir en av de svåraste frågorna för forskare att avgöra vilka faktorer som ska inkluderas, hur de ska mätas och hur de samverkar med varandra. Flera studier har försökt kartlägga olika aspekter av dessa samband. I det följande presenteras ett urval av forskning som belyser några av de mest framträdande faktorerna bakom variationer i examensfrekvens.

Letkiewicz et al. (2014) genomförde en undersökning på ett nationellt representativt urval av förstagsstudenter som påbörjade sina studier hösten 2003 vid fyraåriga amerikanska lärosäten. Studien omfattade ekonomiska, akademiska, demografiska och institutionella faktorer för att undersöka vilka som påverkade sannolikheten att fullfölja examen. Resultaten visade att ekonomiska faktorer var centrala. Studenter som upplevde ekonomisk stress, oavsett orsak, hade lägre examensfrekvenser. Som väntat var även akademiska faktorer viktiga, där högre genomsnittliga betyg var förknippade med en större sannolikhet att slutföra utbildningen. Demografiska faktorer som kön, ålder och etnicitet hade också betydelse, men effekternas riktning och styrka varierade mellan grupper. Bland de institutionella faktorerna framkom att fyraåriga och mer selektiva universitet uppvisade högre examensfrekvenser, något som är särskilt relevant för denna uppsats.

National Student Clearinghouse Research Center (NSCRC) presenterade i sin nationella rapport från 2018, baserad på data från flera miljoner studenter vid amerikanska lärosäten, en omfattande analys av faktorer som påverkar examensfrekvenser. Särskilt fokus lades på institutionella faktorer, och resultaten överensstämde i stor utsträckning med Letkiewicz et al. (2014), där bland annat fyraåriga och mer selektiva institutioner uppvisade högre examensfrekvenser. Rapporten

betonade även betydelsen av studenters tillgång till resurser och stödprogram, såsom akademisk rådgivning och mentorskap, vilka lyftes fram som viktiga för att öka sannolikheten att fullfölja utbildningen. Vidare identifierades ekonomiska faktorer och studenters socioekonomiska bakgrund som centrala, där resultaten låg i linje med tidigare forskning och visade att studenter från högre socioekonomiska grupper har större sannolikhet att ta examen.

Slutligen undersöker Ewert (2012) ett perspektiv särskilt inriktat på manliga studenters examensfrekvenser. Studien visar att män, oavsett socioekonomisk bakgrund, ras eller etnisk tillhörighet, ligger efter sina kvinnliga motsvarigheter. Analysen pekar på flera faktorer bakom dessa könsskillnader, bland annat att män i genomsnitt rapporterar lägre studieengagemang och uppnår lägre betyg. Män är även överrepresenterade i utbildningar och ämnesinriktningar som generellt uppvisar lägre examensfrekvenser, såsom vissa naturvetenskapliga, tekniska och yrkesinriktade program. Ewert drar slutsatsen att för att minska könsgapet krävs åtgärder som stärker mäns studieengagemang och akademiska prestationer.

3. Teoretiskt ramverk

Role model theory utgör tolkningsramen för att förstå och analysera de samband som observeras i denna uppsats. Teorin används ofta inom utbildningsforskning för att förklara hur lärare och akademisk personal kan fungera som förebilder som påverkar studenters motivation, studieval och prestationer. Eftersom denna studie fokuserar på relationen mellan manliga studenters examensfrekvenser och andelen kvinnor i fakulteten bedöms teorin vara ett lämpligt perspektiv för analysen.

3.1 Role model theory

“Role model theory” utgår från att individer påverkas av personer som de identifierar sig med. I studiesammanhang kan detta ofta handla om lärare eller annan fakultet. En rollmodell behöver inte ha en tidigare relation till studenten, vilket skiljer dem från exempelvis mentorer. Gladstone och Cimpian (2021) identifierar tre huvudegenskaper som särskilt viktiga för att en rollmodell ska vara motiverande:

1. Upplevd kompetens, rollmodellen uppfattas som skicklig och framgångsrik inom sitt område.
2. Upplevd likhet, kan vara demografisk (t.ex. kön, etnicitet) eller psykologisk (t.ex. delade värderingar, intressen).
3. Upplevd möjlighet att nå samma framgång, studenten upplever att rollmodellens framgång ligger inom räckhåll.

Gällande den andra punkten om likhet betonar Gladstone och Cimpian att psykologisk likhet ibland kan vara lika viktig, eller viktigare, än demografisk. Detta gör teorin särskilt relevant för en studie som undersöker potentiella samband mellan kvinnlig fakultet och manliga studenters examensfrekvenser.

Huvudhypotesen i denna uppsats är att det finns ett samband mellan andelen kvinnor i fakulteten och manliga studenters examensfrekvenser, och att detta samband kan vara antingen positivt

eller negativt. Utifrån role model theory kan en ökad andel kvinnor tänkas ha en positiv effekt om psykologisk likhet och upplevd kompetens väger tyngre än demografiska skillnader. Med psykologisk likhet avses likhet i mål och arbetssätt snarare än kön, vilket innebär att en student kan känna igen sig i en lärares ambitionsnivå och sätt att tänka även när förebilden är kvinna. I sådana fall kan manliga studenter inspireras av kvinnliga förebilder och se deras prestationer som möjliga att uppnå.

Samtidigt kan teorin också förutsäga en negativ effekt om demografisk likhet är avgörande för motivationen, vilket kan innebära att en lägre andel eller synlighet av manliga förebilder minskar identifikationen och därmed examensfrekvensen. Denna studie kan inte fullt ut avgöra varför en eventuell effekt uppstår, men resultaten kan indikera vilken av dessa riktningar som tycks dominera i den aktuella kontexten.

Role model theory fångar dock inte alla mekanismer bakom examensfrekvenser. I denna studie beaktas även andra faktorer, såsom ekonomiska förutsättningar och institutionella skillnader. Teorin fungerar därför främst som en grund för att analysera de primära samband som undersöks. Det är värt att notera att tidigare forskning vanligtvis fokuserar på rollmodeller av samma kön, exempelvis kvinnliga förebilders påverkan på kvinnliga studenter och motsvarande för män. Detta gör teorin relevant som utgångspunkt, men lämnar en viss osäkerhet kring om den kan förklara eventuella effekter på män.

4. Data och variabler

4.1 Dataöversikt och variabelkonstruktion

Data för samtliga fyra variabler har hämtats från IPEDS-databaserna. Dessa administreras av NCES, en federal myndighet i USA med huvudansvar att samla in och analysera utbildningsdata nationellt. NCES ska enligt sitt uppdrag rapportera all information på ett objektivt och korrekt sätt, därför anses myndigheten vara en pålitlig källa att grunda denna uppsats på (U.S. Code, Title 20, Chapter 76, Part C, § 9541, 2011).

För vissa variabler har datan bearbetats i mindre omfattning för att anpassas till studiens syfte och underlätta tolkningen av resultaten. Variabeln för kvinnlig fakultet, som ursprungligen rapporterades i absoluta tal, omvandlades till en andel genom att dividera antalet kvinnliga fakultetsmedlemmar med det totala antalet anställda vid respektive lärosäte. Därefter omvandlades den till procent för att överensstämja med variabeln för studenter som får finansiellt stöd, vilken redan var angiven i procent. Detta gjordes för att standardisera alla mått och underlätta jämförelser med den beroende variabeln, mäns examensfrekvens, som från början också var uttryckt i procent. Vidare kodades även variabeln för institutionernas nivå om till en dummyvariabel då endast tvååriga och fyraåriga institutioner ingick i datamaterialet.

Datamaterialet omfattar 3 285 lärosäten över tre år. Materialet är obalanserat, vilket innebär att det saknas observationer för ett antal lärosäten. Cirka 8,7 procent av alla lärosäten saknar observationer en eller fler oberoende variabler. Detta anses vara acceptabelt för genomförandet av regressionerna, då majoriteten av lärosätena har fullständiga observationer och en fixed effects-modell kan hantera obalanserade dataset. Ursprungligen saknades även data för mäns examensfrekvenser vid 200 lärosäten, men dessa exkluderades från modellerna.

För samtliga dataset gäller att filerna för 2021 och 2022 har uppdaterats av IPEDS med reviderade versioner, medan 2023 års version ännu inte har reviderats vid uppsatsens skrivande. För att öka tillförlitligheten har de reviderade versionerna använts när de funnits tillgängliga.

4.2 Beroende variabel

Den beroende variabeln för denna studie är examensfrekvenser för manliga studenter och betecknas i regressionstabellerna som *grad_rate_men*. IPEDS definierar denna variabel som följande:

$$\frac{\text{Antal som tar examen inom 150\% av normerad tid}}{\text{Justerad årskull}}$$

Ekvation 1: Formel för beräkning av examensfrekvens

Denna omfattar heltidsstuderande, förstagsregistrerade män som tar examen inom 150 procent av normerad tid. Detta motsvarar sex år för fyraåriga program och tre år för tvååriga program. Den justerade årskullen innebär att vissa studenter exkluderas enligt Student Right-to-Know-lagstiftningen. Exempel på detta är studenter som avlidit, blivit permanent funktionshindrade eller som har lämnat studierna för militärtjänst eller andra officiella uppdrag.

För examensåret 2023 avser årskullen män som påbörjade studier vid fyraåriga lärosäten 2017 respektive tvååriga lärosäten 2020. Samma logik gäller för de andra två årgångarna som inkluderas.

4.3 Andel kvinnlig fakultet

Den huvudsakliga oberoende variabeln mäter andelen kvinnliga fakultetsmedlemmar. Denna betecknas *female_faculty_pct* och inkluderar alla kvinnor som räknas som fakultet enligt IPEDS egna definition. Den omfattar kvinnliga professorer, lektorer, instruktörer eller andra kvinnor i ledande positioner om deras huvudsakliga uppgift är undervisning eller forskning.

Forskningsassistenter, doktorander, undervisningsassistenter eller administrativ personal utan undervisnings- eller forskningsuppdrag ingår inte.

I det ursprungliga datasetet fanns ingen variabel för andelen kvinnor, utan denna skapades genom att dividera antalet kvinnliga fakultetsmedlemmar med det totala antalet

fakultetsmedlemmar vid varje lärosäte, detta inkluderar både män och kvinnor enligt samma definition som ovan.

4.4 Kontrollvariabler

Förutom den huvudsakliga oberoende variabeln inkluderas två kontrollvariabler. Den första kontrollvariabeln som inkluderas i modellerna är andel studenter som får ekonomiskt studiestöd och betecknas *pct_with_grants*. Variabeln omfattar bidrag eller stipendier från myndigheter på nationell, statlig eller lokal nivå. Även stöd från institutionen själv inkluderas i statistiken. Denna uttrycktes i sitt ursprung i procent och krävde därmed ingen vidare omvandling.

Den andra kontrollvariabeln som inkluderas i pooled OLS-modellen är institutionernas nivå och är en dummyvariabel som tar värdet 1 för tvååriga institutioner och värdet 0 för fyraåriga institutioner. Den betecknas som *inst_level_dummy*. I det ursprungliga datasetet ingick även institutioner som erbjuder program kortare än två år. Dessa institutioner saknas dock helt i det slutliga datamaterialet. Anledningen är att de inte rapporterar examensfrekvens enligt IPEDS definition för tvååriga och fyraåriga program, utan använder andra mått som inte är kompatibla med denna studie. Då endast två kategorier återstod kodades variabeln om till en dummyvariabel, vilket gör modellen mer tydlig och underlättar tolkningen av koefficienten i regressionsanalysen.

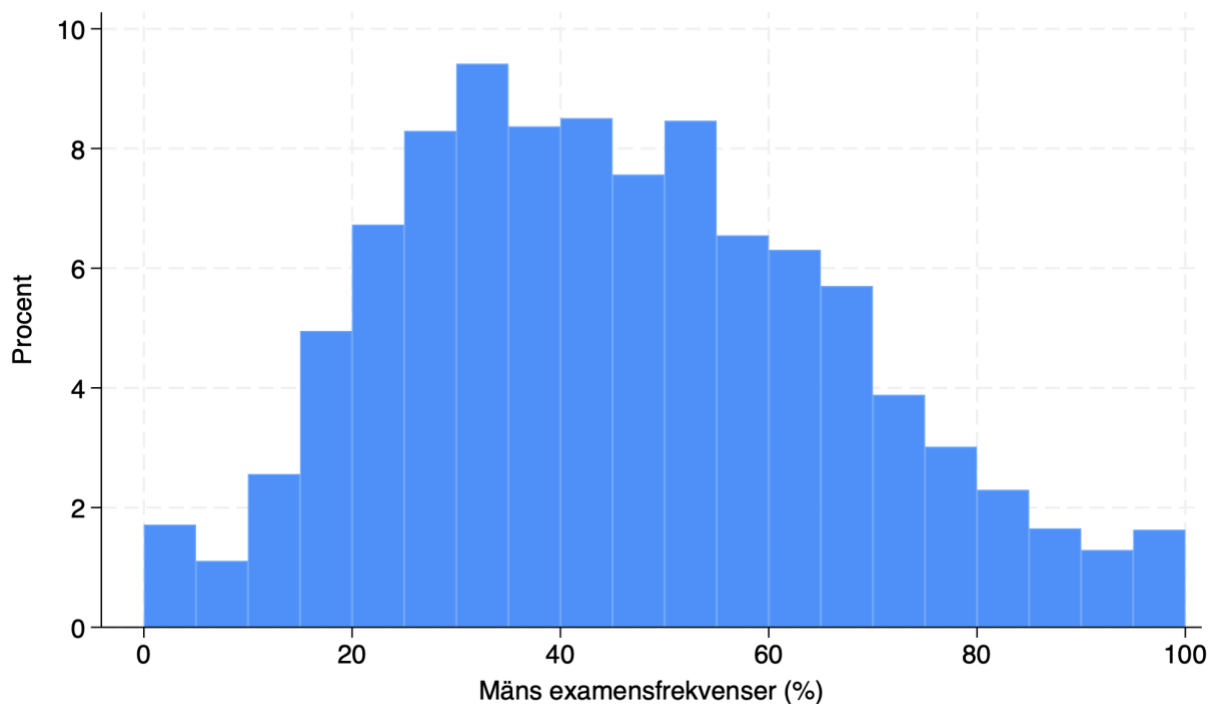
4.5 Datamaterialets begränsningar

Den största begränsningen i det använda datamaterialet är den korta tidsperioden på tre år. Detta begränsar möjligheten att fånga långsiktiga trender, vilket är särskilt relevant i denna studie eftersom fakultetens könsfördelning förändras långsamt. Denna avgränsning beror på en förändring i hur IPEDS rapporterar examensfrekvenser. För åren 2021, 2022 och 2023 rapporteras dessa som andelar i enlighet med Ekvation 1, medan de i tidigare år publicerades i absoluta tal. Det ansågs inte möjligt att på ett tillförlitligt sätt konstruera jämförbara mått för

tidigare år eftersom exakta siffror för varje institutions justerade årskull inte var tillgängliga. Även om den korta tidsperioden är en begränsning finns det också en fördel. Genom att bara använda 2021-2023 undviks de avvikelser som uppstod under 2020 på grund av covid-19-pandemin, då många lärosäten hade ovanliga förutsättningar som kan ha påverkat examensfrekvenserna.

Även om NCES och IPEDS tillhandahåller en av de mest tillförlitliga datakällorna som finns för amerikansk högre utbildning är datamaterialet inte felfritt. Förekomsten av reviderade dataset visar att korrigeringar ibland krävs, vilket innebär att den ursprungliga rapporteringen inte alltid är helt korrekt. För åren 2021 och 2022 har reviderade versioner använts, men för 2023 finns endast de ursprungliga versionerna tillgängliga vid uppsatsens skrivande.

Då en stor del av datan rapporteras direkt av institutionerna finns det en risk för felrapportering. Ett exempel är mäns examensfrekvenser, där en relativt hög andel värden ligger vid extremnivåer. Dessa kan bero på uppskattningar eller avrundningar snarare än exakta mätningar, vilket är viktigt att ha i åtanke vid tolkningen av resultaten.



Figur 2: Histogram över variabeln mäns examensfrekvenser, med 5-procentintervall.

Figur 2 visar att majoriteten av lärosätena ligger mellan cirka 30 och 60 procent, men det finns också ökningarna i rapporteringen vid 0 och 100 procent. Sådana avvikelser kan förklaras av att uppgifterna delvis är självrapporterade, att vissa lärosäten har mycket små årskullar vilket gör andelarna känsliga för enstaka individer, samt att vissa lärosäten i materialet enbart har manliga eller kvinnliga studenter.

Slutligen identifierades en särskild egenskap i variabeln för institutionernas nivå. Den kan uppfattas som tidsinvariant eftersom lärosäten sällan ändrar form mellan tvååriga och fyåriga program, men i materialet förekommer ett mindre antal fall där värdena varierar mellan åren. Vissa lärosäten rapporteras exempelvis som tvååriga ett år men som fyåriga ett annat, vilket antingen kan spegla faktiska förändringar eller bero på brister i rapporteringen. Detta innebär att variabeln inte är lämplig att inkludera i fixed effects-modellen, men den används fortfarande i pooled OLS-modellen, med beaktande av dessa begränsningar.

4.6 Deskriptiv statistik

För att ge en översikt av datamaterialet presenteras i Tabell 1 beskrivande statistik för de variabler som ingår i analysen. Tabellen visar antal observationer, medelvärde, standardavvikelse samt minsta och största värde för respektive variabel. Syftet är att underlätta tolkningen av regressionsresultaten genom att ge en bild av variablernas spridning och nivå.

Tabell 1: Deskriptiv statistik för variablerna i analysen.

Variabel	Antal obs.	Medelvärde	Std.avvikelse	Min	Max
grad_rate_men	9 452	45,2354	21,3255	0	100
female_faculty_pct	9 452	52,0292	16,6086	0	100
pct_with_grants	9 452	72,4758	20,8445	0	100
inst_level_dummy	9 452	0,355	0,4785	0	1

Samtliga variabler omfattar 9 452 observationer, vilket motsvarar data för tre år (2021–2023). De tre första variablerna är uttryckta i procent och har därför värden mellan 0 och 100. Den sista

variabeln, institutionsnivå, är en dummyvariabel som antar värdet 0 för fyraåriga lärosäten och 1 för tvååriga.

Manliga studenters examensfrekvens uppgår i genomsnitt till cirka 45 procent och andelen kvinnlig fakultet till cirka 52 procent. Andelen studenter som får ekonomiska bidrag är högre, med ett genomsnitt på cirka 72 procent. Institutionsnivåns medelvärde på omkring 0,35 indikerar att majoriteten av lärosätena i urvalet är fyraåriga.

5. Metod

5.1 Diagnostiska tester

För att öka modellernas tillförlitlighet genomfördes ett antal diagnostiska tester i statistikprogrammet Stata. Resultaten från dessa tester redovisas i kapitel 6.

För att undersöka förekomst av multikollinearitet genomfördes ett VIF-test mellan de tre oberoende variablerna. Multikollinearitet innebär att två eller flera oberoende variabler är starkt korrelerade, vilket i regressions-sammanhang kan påverka pålitligheten i tolkningen av deras koefficienter.

Ett Wooldridge-test genomfördes för att upptäcka autokorrelation i residualerna. Autokorrelation innebär att feltermerna är korrelerade med varandra över tid. Detta strider mot ett grundläggande antagande i regressionsanalys, vilket är att feltermerna ska vara oberoende. Förekomst av autokorrelation kan därmed påverka modellens precision och leda till missvisande slutsatser. Användningen av robusta standardfel i regressionen bidrar även till att motverka sådana effekter.

Slutligen genomfördes ett Hausman-test för att avgöra om en fixed effects- eller random effects-modell var mest lämplig för genomförandet av regressionen. I en fixed effects-modell antas det att varje enhet, vilket i detta fall är de undersökta skolorna, har egna, unika egenskaper som är konstanta över den undersökta tiden och som kan vara korrelerade med de oberoende variablerna. En random effects-modell utgår däremot från att de skillnader som finns mellan enheter är slumpmässiga och inte har något förutsägbart samband med modellens variabler.

5.2 Motivering av metod

Analysen omfattar både en pooled OLS-modell och en fixed effects-modell med robusta standardfel för att undersöka sambandet mellan mäns examensfrekvenser och andelen kvinnliga fakultetsmedlemmar, andelen studenter som får statliga bidrag samt institutionsnivå. Paneldata möjliggör att följa samma lärosäten över flera år och undersöka hur sambanden mellan de valda

variablerna utvecklas. Denna typ av modell bedöms därför som särskilt lämplig för att genomföra analysen.

Pooled OLS behandlar paneldatan som om alla observationer är oberoende av varandra och som att det inte finns några icke-observerbara enhetsspecifika effekter. I denna uppsats används pooled OLS för att ge en första överblick av sambandet mellan variablerna och som jämförelsepunkt mot fixed effects-modellen, som utgör huvudfokus i denna uppsats.

Fixed effects-modellen används för att kontrollera för egenskaper som inte ändras med tiden. Dessa utomstående faktorer omfattas inte av modellen och kan leda till omitted variable bias, det vill säga när en viktig, utelämnad variabel är korrelerad med någon av de förklarande variablerna i modellen. Vid undersökning av examensfrekvenser finns det en stor mängd faktorer som kan påverka resultatet, varav många är svåra att observera eller mäta. Därför är användningen av en fixed effects-modell ett viktigt steg mot att öka tillförlitligheten i resultaten. För denna uppsats finns faktorer som exempelvis skolors geografiska läge, infrastruktur eller historiska rykte, som inte varierar nämnvärt över tid. Genom att dessa kontrolleras med fixed effects minskar risken för snedvridna resultat.

Robusta standardfel används för att motverka problem med heteroskedasticitet, det vill säga när variansen i feltermerna inte är konstant, vilket är vanligt förekommande i paneldata. Vid undersökning av exempelvis skolor över flera år kan det ske att vissa skolor förändras mycket, medan andra inte förändras alls. Detta kan leda till att residualerna får ojämn spridning och att heteroskedasticitet uppstår. Robusta standardfel utgör då ett enkelt och tillförlitligt sätt att korrigera för detta, utan att påverka modellens struktur.

Robusta standardfel kan även delvis korrigera för autokorrelation i residualerna, om sådan förekommer, utöver att hantera heteroskedasticitet. Dessutom är det i paneldata vanligt att tillämpa robusta eller klustrade standardfel som en försiktighetsåtgärd, oavsett om några specifika problem har identifierats (Schmidheiny, 2024).

5.3 Regressionsmodeller

Följande variabler har presenterats i kapitel 4 för användning i båda regressionerna.

- `grad_rate_men`: Examensfrekvenser för manliga studenter, mätt i procent.
- `female_faculty_pct`: Procenten av fakultet på skolan som görs upp av kvinnor.
- `pct_with_grants`: Procenten av studenter som mottar någon form av ekonomiska bidrag.
- `inst_level_dummy`: Dummyvariabel för om institutionen är tvåårig eller fyraårig.

$$grad_rate_men_{i,t} = \beta_0 + \beta_1 female_faculty_pct_{i,t} + \beta_2 pct_with_grants_{i,t} + \beta_3 inst_level_dummy_{i,t} + \epsilon_{i,t}$$

Ekvation 2: Pooled OLS-modell.

$$grad_rate_men_{i,t} = \beta_0 + \beta_1 female_faculty_pct_{i,t} + \beta_2 pct_with_grants_{i,t} + \alpha_i + \epsilon_{i,t}$$

Ekvation 3: Fixed effects-modell.

Fixed effects-modellen får huvudfokus i denna uppsats eftersom den bättre fångar upp tidsinvarierande skillnader mellan lärosäten. Båda regressioner kontrollerar för variation mellan specifika institut i vid olika tidpunkter t . Eftersom den senare modellen estimeras med fixed effects inkluderas $\alpha_{i,t}$, som fångar upp skolspecifika egenskaper inte mäts av modellen. Slutligen innehåller båda modeller $\epsilon_{i,t}$, vilket är en felterm som fångar upp variation i den beroende variabeln som inte förklaras av de inkluderade oberoende variablerna.

5.4 Hypotes

Denna uppsats prövar nollhypotesen att andelen kvinnor i fakulteten inte har något samband med examensfrekvensen bland manliga studenter. Analysen genomförs med hjälp av en

regressionsanalys med ytterligare oberoende variabler inkluderade som kontrollvariabler. Baserat på tidigare forskning är förväntningen att andelen kvinnor i fakulteten har ett positivt samband med examensfrekvensen bland män, vilket rapporterades i Sav (2012). Men då denna studie publicerades för över tio år sedan och könssammansättningen i fakulteten har skiftat sedan dess, är det inte självklart att samma samband gäller i dagens kontext.

Nollhypotes: Andelen kvinnor i fakulteten har inget samband med examensfrekvensen bland män.

$$H_0: \beta_1 = 0$$

Alternativhypotes: Andelen kvinnor i fakulteten har ett samband med examensfrekvensen bland män.

$$H_A: \beta_1 \neq 0$$

6. Resultat

6.1 Resultat av diagnostiska tester

6.1.1 Variance Inflation Factor (VIF)-test

Ett VIF-test används för att undersöka om det förekommer multikollinearitet mellan de oberoende variablerna. Ett högre VIF värde indikerar högre korrelation med de andra variablerna. Ett VIF värde kan inte vara lägre än 1 och värden över 10 anses indikera potentiella problem med multikollinearitet. De observerade VIF-värdena för denna uppsats redovisas i tabell 2.

Tabell 2: Resultat av Variance Inflation Factor (VIF)-testet.

Variabel	VIF	1/VIF
inst_level_dummy	1,13	0,883256
pct_with_grants	1,10	0,909650
female_faculty_pct	1,03	0,967592
Genomsnittligt VIF	1,09	

Samtliga observerade VIF-värden ligger nära ett, vilket tyder på att multikollinearitet inte utgör ett problem i modellerna. Låga VIF-värden utesluter dock inte helt risken för viss multikollinearitet, men nivån bedöms vara tillräcklig för analysens syfte.

6.1.2 Wooldridge-test

Ett Wooldridge-test används för att undersöka om det förekommer autokorrelation i paneldata, det vill säga om feltermen är korrelerade över tid inom samma enhet. Testet fokuserar på autokorrelation av första ordningen, vilket innebär att feltermen för en enhet vid en viss tidpunkt är korrelerade med feltermen från föregående tidpunkt. Testet bygger på följande hypoteser:

Nollhypotes (H_0): Det finns ingen autokorrelation av första ordningen.

Alternativ hypotes (H_1): Det finns autokorrelation av första ordningen.

Wooldridge-testet tolkas utifrån p-värdet. Ett p-värde under 0,05 leder till att nollhypotesen förkastas, vilket indikerar förekomst av autokorrelation.

Testresultatet visar ett p-värde på 0,3528, vilket överstiger signifikansnivån på 5 procent. Detta innebär att nollhypotesen inte förkastas och att det därför inte finns några statistiska belägg för förekomst av autokorrelation av första ordningen i datamaterialet. Detta stärker tillförlitligheten i resultaten från de efterföljande regressionerna.

6.1.3 Hausman-test

Ett Hausman-test används för att avgöra vilken modell mellan en fixed effects-modell och en random effects-modell som är mest lämplig för datamaterialet. Testet jämför de skattade koefficienterna från fixed effects-modellen med dem från random effects-modellen för att undersöka om det finns systematiska skillnader mellan dem. Testet bygger på följande hypoteser:

Nollhypotes (H_0): Det finns ingen systematisk skillnad mellan koefficienterna från fixed effects- och random effects-modellen. Random effects-modellen är mer effektiv och därmed att föredra.

Alternativ hypotes (H_1): Det finns en systematisk skillnad mellan koefficienterna från fixed effects- och random effects-modellen. Fixed effects-modellen är därför att föredra.

Tabell 3: Resultat av Hausman-testet.

Variabel	FE-koefficient (b)	RE-koefficient (B)	Skillnad (b-B)	Std. fel
female_faculty_pct	-0,028342	-0,085367	0,0570147	0,022302
pct_with_grants	-0,0260538	0,0090603	-0,0351141	0,0074525
inst_level_dummy	-2,990997	-7,528473	4,537476	1,623066

Testresultat: $\chi^2(3) = 32,37$, $p = 0.0000$

Resultatet från Hausman-testet visar ett p-värde på 0,0000. Detta understiger signifikansnivån på 5 procent, vilket innebär att nollhypotesen förkastas. Det finns därmed statistiskt stöd för att koefficienterna i de två modellerna skiljer sig åt på ett systematiskt sätt. I detta fall bedöms fixed effects-modellen vara mer tillförlitlig, eftersom den ger konsistenta skattningar även när de individuella effekterna är korrelerade med de oberoende variablerna. fixed effects-modellen utgör därför grunden för den huvudsakliga analysen i denna uppsats.

6.2 Pooled OLS-modell

För att inleda analysen används en pooled OLS-modell över perioden 2021 till 2023. Den beroende variabeln är examensfrekvenser bland manliga studenter. Modellen inkluderar tre oberoende variabler: andelen kvinnliga fakultetsmedlemmar, andelen studenter med ekonomiska bidrag samt en dummyvariabel för institutionsnivå (tvåårig = 1, fyraårig = 0). Som nämnts i avsnitt 4.1 är samtliga variabler, förutom institutionsnivån, uttryckta i procent för att underlätta tolkningen av koefficienterna.

Pooled OLS används som en grundläggande modell för att ge en översiktlig bild av sambanden mellan variablerna och fungerar även som jämförelsepunkt inför den efterföljande fixed effects-modellen. För att öka tillförlitligheten i skattningarna har modellen estimerats med robusta standardfel.

Tabell 4: Resultat från pooled OLS-modellen.

VARIABLES	grad_rate_men
female_faculty_pct	-0,120*** (0,0161)
pct_with_grants	-0,0654*** (0,0115)
inst_level_dummy	-7,960*** (0,481)
Constant	49,54*** (1,265)
Antal observationer	9 452
R-kvadrat	0,058
p-värde (modellens F-test):	<0,001
Robusta standardfel inom parentes	
*** p<0,01, ** p<0,05, * p<0,1	

Modellen som helhet är statistiskt signifikant ($p < 0,001$), vilket innebär att minst en av de inkluderade variablerna har ett samband med den beroende variabeln. Samtliga koefficienter är dessutom signifikanta på enprocentsnivån, vilket tyder på att varje förklarande variabel bidrar till att förklara variationen i examensfrekvenser bland manliga studenter.

Andelen kvinnliga fakultetsmedlemmar har en koefficient på -0,120. Detta innebär att en ökning med en procentenhet i andelen kvinnor i fakulteten hänger samman med en genomsnittlig minskning i männens examensfrekvenser med cirka 0,12 procentenheter.

För andelen studenter som får ekonomiska bidrag är koefficienten -0,0654, vilket tyder på ett negativt samband med den beroende variabeln. Det vill säga att om andelen studenter med ekonomiska bidrag ökar med en procentenhet, är det förknippat med en minskning i männens examensfrekvenser med cirka 0,07 procentenheter.

Dummyvariabeln för institutionernas nivå (tvååriga eller fyraåriga) har en koefficient på -7,960. Variabeln är kodad så att fyraåriga lärosäten motsvaras av siffran 0 och tvååriga av 1. Detta innebär att tillhöra en tvåårig institution är förknippat med en i genomsnitt 7,96 procentenheter lägre examensfrekvens för manliga studenter, jämfört med att tillhöra en fyraårig institution.

6.3 Fixed effects-modell

Analysen fortsätter med en fixed effects-modell för att kontrollera för tidsinvarianta faktorer som inte direkt observeras men som kan påverka resultaten. Modellen estimeras med robusta standardfel och inkluderar två oberoende variabler, andelen kvinnliga fakultetsmedlemmar samt andelen studenter med ekonomiska bidrag. Variabeln för institutionernas nivå exkluderas eftersom den i grunden kan uppfattas som tidsinvariant och inte meningsfull i en fixed effects-modell, samt på grund av de dataproblem som tidigare diskuterats. Valet av modell stöds av Hausman-testets resultat, som indikerar att fixed effects ger mer tillförlitliga skattningar i detta fall.

Tabell 5: Resultat från fixed effects-modellen.

VARIABLES	grad_rate_men
female_faculty_pct	-0,0290 (0,0497)
pct_with_grants	-0,0271 (0,0164)
Constant	48,70*** (2,985)
Antal observationer	9 452
Antal enheter	3 285
R-kvadrat (overall)	0,001
p-värde (modellens F-test):	0,242
Robusta standardfel inom parentes	
*** p<0,01, ** p<0,05, * p<0,1	

Sett till signifikansnivån för modellen som helhet framgår det att den inte är statistiskt signifikant ($p = 0,242$). Detta innebär att det inte finns tillräckligt med statistiskt stöd för att de förklarande variablerna har ett samband med examensfrekvensen bland män.

Koefficienten för andelen kvinnliga fakultetsmedlemmar visar ett negativt samband på $-0,0290$, vilket innebär att en ökning av denna variabel med en procentenhet hänger samman med en genomsnittlig minskning av mäns examensfrekvenser med cirka $0,03$ procentenheter.

Resultatet är dock inte statistiskt signifikant, då p-värdet ($0,560$) överstiger signifikansnivån på 5 procent.

Även andelen studenter med ekonomiska bidrag uppvisar ett negativt samband, med en koefficient på $-0,0271$. Det innebär att en ökning med en procentenhet i andelen studenter med ekonomiska bidrag är förknippat med en minskning i mäns examensfrekvenser med ungefär $0,03$ procentenheter. Inte heller detta resultat är statistiskt signifikant ($p = 0,100$).

6.4 Skillnader mellan pooled OLS och fixed effects

Till skillnad från pooled OLS-modellen, där de inkluderade koefficienterna var statistiskt signifikanta på enprocentsnivån, visar fixed effects-modellen inga signifikanta samband mellan de oberoende variablerna och examensfrekvenser bland manliga studenter. En möjlig förklaring är att sambanden som observerades i pooled OLS-modellen drevs av skillnader mellan lärosäten, exempelvis strukturella faktorer som inte förändras över tid. När dessa tidsinvarianta egenskaper kontrolleras bort i fixed effects-modellen, försvinner sambanden, vilket tyder på att de tidigare resultaten delvis kan ha varit ett resultat av sådana underliggande skillnader snarare än verkliga förändringar inom lärosätena över tid.

Minskningen i förklaringsgrad mellan de två modellerna kan förklaras av skillnader i hur variation hanteras. Fixed effects-modellen har i praktiken mindre information att arbeta med än pooled OLS, eftersom den endast utnyttjar variation inom varje lärosäte över tid, samt att en variabel färre ingår i modellen. Pooled OLS använder däremot all tillgänglig variation, både mellan lärosäten och inom samma lärosäte, vilket generellt leder till en högre förklaringsgrad. När fixed effects-modellen istället kontrollerar för alla tidsinvarianta faktorer genom att eliminera variationen mellan lärosäten, återstår mindre variation för modellen att förklara. Detta blir särskilt tydligt med en kort tidsperiod, såsom de tre år som denna analys omfattar. Trots detta är användningen av fixed effects metodologiskt motiverad, och den lägre förklaringsgraden är en naturlig konsekvens av att kontrollera för potentiellt snedvridande, konstanta skillnader.

6.5 Modellens begränsningar

Fixed effects-modellen påverkas av flera begränsningar. Den korta tidsperioden i analysen innebär att möjligheten att identifiera variation över tid är begränsad. Detta förstärks av att den centrala förklaringsvariabeln, andelen kvinnlig fakultet, förändras långsamt och därmed uppvisar liten variation mellan åren. Kombinationen av kort tidsperiod och begränsad tidsvariation gör att den variation inom lärosäten som fixed effects-modellen kan utnyttja blir mycket liten. Modellen omfattar dessutom ett begränsat antal variabler, vilket innebär att en stor del av variationen i

utfallet sannolikt förklaras av faktorer som inte ingår i analysen. Sammantaget bidrar dessa begränsningar till den låga förklaringsgrad som observeras i modellen.

7. Diskussion

Denna uppsats grundade sig i frågeställningen om i vilken utsträckning manliga studenters examensfrekvenser var relaterade till andelen kvinnor i fakulteten, med hänsyn till institutionernas nivå och studenternas tillgång till ekonomiskt stöd. Analysen ger två olika bilder beroende på modellval. Pooled OLS indikerar signifikanta samband mellan de undersökta variablerna och männens examensfrekvenser, medan fixed effects inte visar några statistiskt säkerställda samband. Detta pekar på att resultaten är känsliga för metodval, vilket i sig är en viktig iakttagelse.

Baserat enbart på resultaten från fixed effects-modellen bör eventuella samband tolkas med försiktighet, då inget statistiskt säkerställt samband kunde påvisas mellan de undersökta variablerna och manliga studenters examensfrekvenser. Detta utesluter dock inte att samband kan förekomma. Tidigare forskning är fortfarande begränsad men pekar på att relationer mellan könssammansättning i fakulteten och studenters utfall kan finnas. I denna studie gav pooled OLS-modellen dessutom signifikanta resultat, vilket stärker möjligheten att samband existerar.

Skillnaden mellan modellerna kan förklaras av hur de hanterar variation i datamaterialet. Pooled OLS använder både variationen mellan och inom lärosäten, vilket gör att modellen kan upptäcka mönster som delvis kan drivas av tidsinvarianta faktorer som inte observeras. Fixed effects-modellen eliminerar denna variation och kontrollerar därmed för konstanta skillnader mellan lärosäten, vilket ger mer tillförlitliga skattningar men också innebär att mycket av variationen försvinner. Detta blir särskilt tydligt när den centrala förklaringsvariabeln, andelen kvinnor i fakulteten, förändras långsamt samtidigt som analysen endast omfattar tre år.

Sammantaget visar pooled OLS-modellen på samband som kan bero på skillnader mellan lärosäten, medan fixed effects-modellen inte finner några signifikanta effekter inom lärosäten över tid. Tillsammans pekar resultaten på att frågan är komplex och att vidare analyser med längre tidsperioder och fler variabler skulle behövas för att kunna dra starkare slutsatser.

7.1 Jämförelse med tidigare studier

Givet att fixed effects-modellen inte gav signifikanta resultat medan pooled OLS gjorde det, blir frågan om hur resultaten ska relateras till tidigare forskning mer komplex. Att den mer rigorösa modellen tappat signifikans väcker viss tvekan om resultaten från pooled OLS är fullt tillförlitliga. De icke-signifikanta resultaten kan varken bekräfta eller motsäga tidigare studier, och de signifikanta resultaten bör därför också tolkas med försiktighet. Med detta sagt går det fortfarande att göra vissa jämförelser för att identifiera möjliga mönster.

För variabeln kvinnlig fakultet visade fixed effects ett svagt negativt men icke-signifikant samband, medan pooled OLS gav ett signifikant negativt samband. Detta tyder på att skillnader mellan lärosäten kan driva resultaten i pooled OLS, medan variationen inom samma lärosäte över tid är för liten för att fångas upp i fixed effects. Båda dessa resultat skiljer sig från Sav (2012), som fann att en högre andel kvinnor i fakulteten var associerad med högre examensfrekvenser. Som tidigare diskuterats skrevs Savs studie i en annan kontext, och könssammansättningen inom akademien har förändrats sedan dess. AAUP (2020) visar exempelvis att andelen kvinnor i fakulteter har ökat markant under senare år även om skillnader kvarstår, särskilt inom professorstjänster och *non-tenure-track*-positioner. Det är möjligt att de positiva effekter som tidigare kunde observeras har minskat i takt med att könsfördelningen närmat sig paritet. Detta är dock spekulativt och bör undersökas vidare i framtida forskning.

För andelen studenter med ekonomiskt stöd visade fixed effects en koefficient nära noll och utan statistisk signifikans. Koefficientens riktning pekade ändå mot ett svagt negativt samband, vilket indikerar att en större andel ekonomiskt utsatta studenter kan vara associerad med något lägre examensfrekvenser bland män. I pooled OLS var detta samband signifikant, vilket stärker tolkningen att ekonomiska förutsättningar kan spela roll, även om detta inte kunde fastställas när tidsinvarianta faktorer kontrollerades bort. Liknande mönster framkom i Sav (2012), som visade ett negativt samband mellan ekonomiskt stöd och mäns examensfrekvenser. Även Letkiewicz et al. (2014) fann att studenter med begränsade ekonomiska resurser i högre grad riskerar att inte ta examen i tid, bland annat på grund av stress och konkurrerande åtaganden vid sidan av studierna. Mot denna bakgrund framstår resultatet från pooled OLS-modellen som mer i linje med tidigare

forskning, även om skillnaden mot fixed effects visar att sambandet inte kan fastställas med säkerhet.

Institutionernas nivå kunde inte inkluderas i fixed effects eftersom den i praktiken inte förändras över tid inom lärosäten. De få fall där variabeln varierar mellan åren bedöms snarare spegla inkonsekvent rapportering än faktiska omstruktureringar. Detta är dock inget problem för analysen, eftersom fixed effects-modellen ändå kontrollerar för de strukturella skillnader mellan två- och fyrfåriga lärosäten som variabeln avser att fånga. I pooled OLS framkom däremot ett signifikant negativt samband, där tvååriga lärosäten i genomsnitt hade lägre examensfrekvenser bland män än fyrfåriga. Detta ligger i linje med tidigare studier, exempelvis NSCRC (2024), som visar att tvååriga institutioner generellt har lägre examensfrekvenser. Möjliga förklaringar är bland annat mindre resurser, lägre selektivitet vid antagning och att många studenter inte har examen som primärt mål utan planerar att övergå till fyrfåriga utbildningar eller etablera sig snabbt på arbetsmarknaden.

Sammanfattningsvis visade fixed effects inga statistiskt säkerställda samband, medan pooled OLS gav flera signifikanta resultat som i stor utsträckning överensstämmer med tidigare forskning. Skillnaderna mellan modellerna understryker att resultaten är känsliga för metodval och att tolkningen därför bör ske med viss försiktighet.

7.2 Begränsningar och framtida forskning

Begränsningar har diskuterats löpande genom uppsatsen, men det är särskilt viktigt att belysa dessa med tanke på de icke-signifikanta resultaten och den genomgående låga förklaringsgraden. En tydlig genomgång av begränsningarna kan underlätta för framtida studier att minska de felkällor som kan ha bidragit till de svaga resultaten. De faktorer som bedöms ha haft störst inverkan i denna studie är den korta tidsperioden på endast tre år och det begränsade antalet oberoende variabler i modellerna.

Tidsperioden blev kort på grund av databegränsningar där rapporteringen av den beroende variabeln, examensfrekvenser bland manliga studenter, har förändrats över tid och därför inte varit konsekvent tillgänglig. Denna begränsning innebar att analysen endast kunde omfatta tre år,

vilket minskade möjligheten att upptäcka långsiktiga trender och variationer inom lärosäten. En längre tidsperiod hade sannolikt varit särskilt värdefull för variabler som könssammansättning, där förändringarna troligen är små från år till år.

När det gäller antalet oberoende variabler omfattade fixed effects-modellen endast två, andelen kvinnor i fakulteten och andelen studenter med ekonomiskt stöd. Institutionernas nivå kunde inte inkluderas i denna modell eftersom den kan uppfattas som tidsinvariant, men analyserades i pooled OLS, vilket innebar att tre variabler kunde beaktas där. Begränsningen berodde delvis på svårigheter att hitta lämplig data, men också på att modellerna konstruerades på ett sätt som i efterhand visade sig vara mindre fördelaktigt. Avsikten var att minska antalet variabler för att snäva in analysen och fokusera på de mest centrala frågeställningarna, men detta innebar i praktiken en svaghet. Variabler som exempelvis antagningskrav och selektivitet eller lönenivåer inom fakulteten hade sannolikt inte varit särskilt svåra att inkludera och kunde ha bidragit till en mer välgrundad modell och en högre förklaringsgrad.

Andra potentiella variabler, såsom skolors geografiska läge eller historiska rykte, är mer problematiska att inkludera då de kan vara svåra att kvantifiera på ett tillförlitligt sätt. Även om mätningar skulle finnas riskerar de att vara subjektiva och svåra att jämföra mellan lärosäten. Till exempel kan en bedömning av ett lärosätes rykte variera kraftigt beroende på vem som gör bedömningen och vilken metod som används, vilket gör att resultaten kan bli svårtolkade och mindre tillförlitliga.

En fördel med att enbart använda data från NCES var att materialet är välstrukturerat och jämförbart mellan lärosäten, vilket möjliggjorde en smidig sammanslagning av dataset. NCES tillämpar en standardiserad rapportering, vilket minimerar problem med datakompatibilitet. Samtidigt innebar detta att andra statistikdatabaser inte användes, eftersom deras data ofta rapporteras på andra sätt och skulle kräva omfattande bearbetning för att bli jämförbar. Det är möjligt att alternativa källor hade kunnat ge fler år av data eller fler variabler, men de föll bort på grund av dessa begränsningar i tillämpningen.

Genom att bemöta framför allt två begränsningar, den korta tidsperioden och det begränsade antalet oberoende variabler, skulle framtida forskning kunna genomföras med större precision och tillförlitlighet. I synnerhet kan längre tidsserier ge bättre möjligheter att fånga effekter av

långsamt föränderliga variabler, medan fler variabler skulle kunna minska risken för omitted variable bias.

Mäns examensfrekvenser är en aktuell och relevant fråga, då de under en längre tid legat efter kvinnors och inte visat några tydliga förbättringar. Detta gör ämnet viktigt att fortsätta undersöka och diskutera, och alla nya perspektiv som kan bidra till en djupare förståelse bör därför välkomnas.

8. Slutsats

Syftet med denna uppsats var att undersöka i vilken utsträckning examensfrekvenser bland manliga studenter hade ett samband med andelen kvinnor i fakulteten, andelen studenter med ekonomiskt stöd och institutionernas nivå. Inga av sambanden kunde statistiskt säkerställas i fixed effects-modellen, och de resultat som framkom i pooled OLS bör därför tolkas med försiktighet. Dessutom uppvisade analyserna en låg förklaringsgrad och omfattade en kort tidsperiod, vilket ytterligare begränsar styrkan i de slutsatser som kan dras.

Resultaten från pooled OLS indikerade dock vissa mönster. En högre andel kvinnor i fakulteten var förknippad med lägre examensfrekvenser för män. Detta samband är dock inte entydigt i tidigare forskning, och eftersom det endast framkom i pooled OLS bör det tolkas med stor försiktighet då det kan spegla skillnader mellan lärosäten snarare än underliggande mönster.

Även om sambanden inte kunde fastställas med säkerhet, ger resultaten en viss indikation på att institutionella faktorer kan påverka manliga studenters examensfrekvenser. Detta är särskilt relevant i en tid då skillnaderna i utbildningsnivå mellan män och kvinnor kvarstår, och där mäns examensfrekvenser under lång tid har legat lägre än kvinnors. Uppsatsen utgör därmed en relevant kvantitativ analys som, om begränsningarna åtgärdas, kan ge värdefulla bidrag till framtida forskning inom ett område som fortfarande är relativt outforskat.

Referenser

Bailey, M.J. & Dynarski, S.M. (2011). Gains and Gaps: Changing Inequality in U.S. College Entry and Completion, *National Bureau of Economic Research (Working Paper Series)*. Hämtad från: <https://www.nber.org/papers/w17633>

Belkin, D. (2021). A Generation of American Men Give Up on College: 'I Just Feel Lost', *Wall Street Journal*. Hämtad från: <https://www.wsj.com/us-news/education/college-university-fall-higher-education-men-women-enrollment-admissions-back-to-school-11630948233>

Bettinger, E.P. and Long, B.T. (2005). Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students, *American Economic Review*. Hämtad från: <https://doi.org/10.1257/000282805774670149>

Buchmann, C., Dwyer, R.E. & Yao, M. (2025) The Deepening Gender Divide in Credentials, 2000–2020: Continuity, Change, and Implications, *RSF: The Russell Sage Foundation Journal of the Social Sciences*. Hämtad från: <https://doi.org/10.7758/RSF.2025.11.1.08>

Carrell, S.E., Page, M.E. & West, J.E. (2009). Sex and Science: How Professor Gender Perpetuates the Gender Gap, *National Bureau of Economic Research (Working Paper Series)*. Hämtad från: <https://doi.org/10.3386/w14959>

Colby, G. & Chelsea, F. (2020). Data Snapshot: Full-Time Women Faculty and Faculty of Color, *American Association of University Professors*. Hämtad från: <https://www.aaup.org/news/data-snapshot-full-time-women-faculty-and-faculty-color>

Ewert, S. (2012). Fewer Diplomas for Men: The Influence of College Experiences on the Gender Gap in College Graduation, *The Journal of Higher Education*. Hämtad från: <http://www.jstor.org/stable/23324237>

Gladstone, J.R. & Cimpian, A. (2021). Which role models are effective for which students? A systematic review and four recommendations for maximizing the effectiveness of role models in STEM, *International Journal of STEM Education*. Hämtad från: <https://doi.org/10.1186/s40594-021-00315-x>

Goldin, C., Katz, L.F. & Kuziemko, I. (2006). The Homecoming of American College Women: The Reversal of the College Gender Gap, *Journal of Economic Perspectives*. Hämtad från: <https://doi.org/10.1257/jep.20.4.133>

Letkiewicz, J., Lim, H., Heckman, S., Bartholomae, S. & Fox, J. (2014). The Path to Graduation: Factors Predicting On-Time Graduation Rates, *Journal of College Student Retention: Research, Theory and Practice*. Hämtad från: https://www.researchgate.net/publication/271727666_The_Path_to_Graduation_Factors_Predicting_On-Time_Graduation_Rates

National Center for Education Statistics (NCES) (n.d.). Integrated Postsecondary Education Data System (IPEDS), *U.S. Department of Education*. Hämtad från: <https://nces.ed.gov/ipeds/datacenter/DataFiles.aspx?year=-1&sid=4c9ba82d-323e-444a-96cb-804d9d3235c7&rtid=7>

Parker, K. (2021). Women in the U.S. are outpacing men in college graduation, *Pew Research Center*. Hämtad från: https://www.pewresearch.org/short-reads/2021/11/08/whats-behind-the-growing-gap-between-men-and-women-in-college-completion/ft_2021-11-08_highered_01-png/

Reeves, R.V. & Smith, E. (2021). The male college crisis is not just in enrollment, but completion, *Brookings*. Hämtad från: <https://www.brookings.edu/articles/the-male-college-crisis-is-not-just-in-enrollment-but-completion/>

Robst, J., Keil, J. & Russo, D. (1998) The effect of gender composition of faculty on student retention, *Economics of Education Review*. Hämtad från: [https://doi.org/10.1016/S0272-7757\(97\)00049-6](https://doi.org/10.1016/S0272-7757(97)00049-6)

Sav, G. T. (2012). Female Faculty, Tenure, and Student Graduation Success: Efficiency Implications for University Funding, *International Journal of Business Management & Economic Research*. Hämtad från: <https://corescholar.libraries.wright.edu/econ/86>

Schmidheiny, K. (2024). Panel Data: Fixed and Random Effects, *Short Guides to Microeconometrics*. Hämtad från: <https://www.schmidheiny.name/teaching/panel2up.pdf>

Schneider, M. & Yin, L. (2011). The High Cost of Low Graduation Rates: How Much Does Dropping Out of College Really Cost?, *American Institutes for Research*. Hämtad från: <https://eric.ed.gov/?id=ED523102>

Shapiro, D., Dundar, A., Huie, F., Wakhungu, P.K., Bhimdiwala, A. & Wilson, S. E. (2018). Completing College: A National View of Student Completion Rates – Fall 2012 Cohort (Signature Report No. 16), *National Student Clearinghouse Research Center*. Hämtad från: <https://nscresearchcenter.org/signaturereport16/>

The Economist (2024). Why young men and women are drifting apart. Hämtad från: <https://www.economist.com/international/2024/03/13/why-the-growing-gulf-between-young-men-and-women>

United States Code (2011). 20 U.S.C. § 9541, Establishment. Washington, DC: *U.S. Government Publishing Office*. Hämtad från: <https://www.govinfo.gov/content/pkg/USCODE-2011-title20/html/USCODE-2011-title20-chap76-subchapI-partC-sec9541.htm>

Redogörelse för användning av generativ AI

I denna uppsats har OpenAI:s AI-modell ChatGPT använts som stöd för språklig bearbetning av text samt för praktisk hjälp i statistikprogrammet Stata, bland annat vid genomförande av regressioner och framställning av tabeller.