

Forord

Denne oppgaven markerer slutten av mitt 5-årige masterstudium i samfunnsøkonomi ved NTNU. Studietiden har vært en spennende og lærerik prosess, hvor jeg har blitt utfordret på mange områder. Jeg har stor tro på at jeg i fremtiden vil trekke mange erfaringer fra min studietid og kan se tilbake på mange gode minner fra tiden som student ved NTNU.

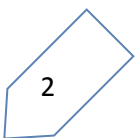
Jeg vil rette en stor takk til min veileder Hans Bonesrønning for hans veiledning, støtte og kommentarer. Når frustrasjonen har vært stor, eller forståelsen liten, har en samtale med Hans alltid ledet meg inn på rett spor. Det har vært en inspirerende og lærerik prosess.

Jeg vil også takke alle venner og kjente som har kommet med innspill til oppgaven, og en spesiell takk til min tante Anja for grundige og gode tilbakemeldinger på oppgaven.

Til slutt vil jeg takke mine foreldre, min bror og mine besteforeldre for all mulig støtte gjennom studietiden.

Trondheim, 29.05.2012

Morten Olimb



Innholdsfortegnelse

1	Innledning	5
1.1	Problemstilling.....	6
1.2	Disposisjon	7
2	Et utvalg tidligere empirisk forskning på arbeidsinnsats og standarder	9
3	Teorigrunnlag	11
3.1	En teoretisk modell for arbeidsinnsats og resultater.....	11
3.2	Underhypotesene	16
4	Datamaterialet	19
4.1	Deskriptiv statistikk.....	19
5	Estimeringsstrategi	23
5.1	Minste kvadraters metode.....	23
5.1.1	Økonometriske utfordringer	24
5.2	Difference in difference	28
5.2.1	Økonometriske utfordringer	29
6	Resultater og diskusjon	31
6.1	Deskriptiv statistikk.....	31
6.2	Estimeringsresultater minste kvadraters metode	32
6.2.1	Effekten av økte standarder på arbeidsinnsats	33
6.2.2	Effekten av økte standarder på resultater	35
6.2.3	Heterogene Effekter	35
6.2.4	Robusthetstester	38
6.3	Estimeringsresultater difference in difference.....	40
6.3.1	Heterogene Effekter	42
6.3.2	Robusthetstester	43
7	Konklusjoner	45

8	Appendiks	47
8.1	Fordelingsfunksjoner	47
8.1.1	Arbeidsinnsats oktober	47
8.1.2	Arbeidsinnsats semesterslutt	47
8.1.3	Δ Innsats	48
8.1.4	Resultat midtsemestereksamen	48
8.1.5	Resultat avsluttende eksamen.....	49
8.2	Tabeller.....	50
9	Referanser.....	53

1 Innledning

I den politiske debatten snakkes det ofte om at Norge må satse på forskning og utdanning for å møte morgendagens utfordringer. Skal vi ha konkurransedyktige kandidater i årene som kommer, må vi utforme utdanningssystemet slik at vi får mest mulig ut av hver enkelt student. Det er mange faktorer som spiller inn på studentenes resultater i høyere utdanning, men mange vil nok hevde at arbeidsinnsatsen til studentene er, om ikke den viktigste, så i hvert fall en av de viktigste innsatsfaktorene når vi skal måle resultater i utdanningen. Selv om de fleste teoretiske modellene innen utdanningsøkonomi gjør antagelser om at arbeidsinnsatsen er en viktig innsatsfaktor, er ikke forskningen på den kausale effekten mellom arbeidsinnsatsen og studentenes resultater entydig. Noen nyere analyser av blant annet Stinebrickner og Stinebrickner (2008) og Bonesrønning og Opstad (2011) har vist at det er en tydelig positiv sammenheng mellom studentenes arbeidsinnsats og resultater. Samtidig utformes nyere reformer med den hensikt at de skal bidra til å øke studentenes incentiver til å yte innsats i studiene (Bonesrønning og Opstad, 2011). Det vil derfor være viktig å kartlegge hvilke faktorer som kan påvirke studentenes arbeidsinnsats.

Når man skal utforme reformer eller tiltak som skal øke studentenes incentiver til å bruke mer tid på studiene er det viktig å ha god kjennskap til effektene av disse. Studentenes arbeidsinnsats bør betraktes som en endogen faktor, som kan påvirkes av institusjonelle karakteristikk, lærerkrefter, hvilke krav som stilles til studentene, personkarakteristikk, peer effekter, kultur og fritidstilbud osv. Det sentrale spørsmålet blir da hvordan utdanningsinstitusjonene kan skape riktige incentiver for å øke studentenes arbeidsinnsats. Et tiltak det er naturlig å tenke seg at studentene vil respondere på er endringer i standardene som settes. Med standarder menes det her hvilke faglige krav som settes til studentene i et emne. Dette kan være så enkelt som flere obligatoriske aktiviteter eller endringer i kravene som settes til de allerede eksisterende obligatoriske aktivitetene. Dersom nye aktiviteter blir innført eller kravene for å få godkjent disse aktivitetene blir høyere, vil en kunne si at dette er en økning i standardene. Økte standarder kan da betraktes som å øke terskelen for oppnå en gitt karakter i et emne. Formålet med denne oppgaven vil være å undersøke hvordan økte standarder påvirker studentenes arbeidsinnsats og der igjennom resultatene.

Mange jobber etter å finne de optimale strategiene for å øke studentenes prestasjoner, og det er en kjensgjerning at standarden har vært fallende ved universiteter og høyskoler verden

over. Det er flere som stiller spørsmål ved om studentene lærer det samme i dag som det forgjengerne gjorde (Figlio og Lucas, 2004). Det har vært lite fokus på hva som påvirker studentenes arbeidsinnsats og hvordan man kan påvirke deres tidsallokering. Det er derfor viktig å belyse hvordan studentenes arbeidsinnsats påvirkes av endringer i standardene. Stinebrickner og Stinebrickner (2004) påpeker viktigheten av å forstå hvilke faktorer som påvirker studentenes arbeidsinnsats etter man har påbegynt sitt studium, fremfor å fokusere på hvilke faktorer som har påvirket studentene i tiden før. Min analyse av hvordan økte standarder påvirker studentenes arbeidsinnsats kan være et bidrag til å belyse dette.

Videre kan man reise spørsmålet om det vil eksistere en optimal standard. Det er naturlig å tenke seg at det er avtakende avkastning av standarder på arbeidsinnsats. Det vil si jo lavere standarden i utgangspunktet er, jo større effekt har en økning i standarden på studentenes arbeidsinnsats. I den motsatte enden av skalaen er det naturlig å tenke seg at ekstremt høye standarder kun fører til at en stor gruppe studenter ikke klarer å fullføre studiene

1.1 Problemstilling

Det er i de senere årene blitt forsket mye på effekten av høyere standarder og hvordan dette påvirker studentenes arbeidsinnsats. Samtidig har det også blitt viet mye oppmerksomhet til hvordan studentenes arbeidsinnsats påvirker sluttkarakteren. Allikevel er det ikke gjort mye arbeid for å se på sammenhengen mellom disse effektene. Dersom studentenes arbeidsinnsats påvirkes av standarden, vil det være meget nyttig å se på hvordan dette påvirker resultatene. Om det er en positiv sammenheng mellom studentenes arbeidsinnsats og sluttkarakteren, vil økte standarder være en god strategi dersom dette øker studentenes arbeidsinnsats. At det har vært gjort lite forskning innen dette området gjør det spennende og utforske videre.

Det er ofte vanlig å anta at studentene er myopiske, noe som vil si at de er nærsynte. Dette kan fremstilles som at de har en høy tidsdiskonteringsfaktor på fremtidige resultater. Det vil si at i emner der det kun er en avsluttende eksamen, vil studentenes subjektive verdsetting av dette resultatet være relativt lav på et tidlig tidspunkt i semesteret. Følgelig vil det også være lite nyttegevinst i å yte en høy arbeidsinnsats i denne perioden. I et slikt perspektiv vil økte standarder i form av obligatoriske aktiviteter som påvirker resultatet, direkte eller indirekte, være med å endre nytteverdien av arbeidsinnsatsen tidlig i semesteret.

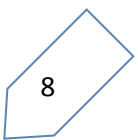
Standarden som betraktes i denne oppgaven er en midtsemestereksamen som ikke påvirker studentenes oppnådde resultat i emnet direkte. Midtsemestereksamen er en flervalgseksamen der standarden betraktes som antallet rette svar som kreves for å bestå denne. Initialt er denne standarden fraværende, men innføres og økes i tidsperioden som analyseres. Studentene som ikke oppnår denne minimumsstandard vil miste retten til å ta den avsluttende eksamen. På denne måten adresseres studentenes kortsynthet ved at de må prestere tilfredsstillende på denne eksamen for i det hele tatt kunne oppnå et resultat i emnet.

Hypotesen jeg ønsker å undersøke i denne oppgaven er at en økning i standardene vil øke studentenes arbeidsinnsats, og derigjennom også de faglige resultatene. Hypotesen skal testes ut fra spørreundersøkelser som er gjort blant studenter i emnet makroøkonomi ved Trondheim Økonomiske Høyskole (TØH). Helt konkret vil den aktuelle problemstillingen i oppgaven være gitt ved:

Hvordan påvirkes studenters arbeidsinnsats og eksamensresultater av økte standarder?

1.2 Disposisjon

I kapittel 2 vil jeg presentere en oversikt over tidligere empirisk forskning på effektene av økte standarder. I kapittel 3 utleder jeg teorigrunnlaget for oppgaven gjennom en modell som viser hvilke faktorer som påvirker studentenes resultater og hvordan de allokere tiden i studiene. Videre vil jeg i kapittel 4 presentere datamaterialet og den deskriptive statistikken for oppgaven. Kapittel 5 gir en presentasjon av estimeringsstrategien jeg benytter for å komme frem til resultatene, og de økonometriske utfordringene knyttet til estimeringen. Disse resultatene presenteres og diskuteres i kapittel 6, hvor jeg også presenterer robusthetstester for analysene. Til slutt gir jeg en oppsummering og konklusjon i kapittel 7.



2 Et utvalg tidligere empirisk forskning på arbeidsinnsats og standarder

Det er ikke et stort utvalg av tidligere empirisk forskning som ser på effekten av standarder i høyere utdanning. Sammenhengen mellom standarder og resultater i tidligere empirisk forskning er i hovedsak gjort på utdanningsnivåer under universitets- og høyskolenivå, og det meste av denne forskningen fokuserer på hvordan en endring i standardene påvirker resultatene. Det er ikke noen tidligere empirisk forskning, meg bekjent, som ser på hvordan studentenes arbeidsinnsats påvirkes av standardene. Den empiriske forskningen på effektene av standarder viser at en økning i standardene har en positiv effekt på resultatene. Bonesrønning (1999) finner at høyere standarder gir bedre elevprestasjoner. Effekten kan sammenlignes med et skift i arbeidstilbudet ved endringer i arbeidsuavhengig inntekt. Også Betts og Grogger (2003) og Figlio og Lucas (2004) finner i sine analyser at resultatene øker med økte standarder. Betts og Grogger (2003) finner at resultatene bedres i skoler med høye standarder, men at denne effekten er sterkest for de elevene som er i den øvre delen av prestasjonssjiktet. Figlio og Lucas (2004) konkluderer med at effekten av den økte standarden er avhengig av elevsammensetningen. De finner at elevene har størst positiv effekt av økte standarder dersom man er en høyt presterende elev i en lavt presterende klasse eller dersom man er en lavt presterende elev i en høyt presterende klasse. Denne forskningen benytter variasjon i standarder mellom lærere, og følgelig er det en utfordring å separere effekten av standarder fra effekten av andre lærekjennetegn.

Bakgrunnen for å analysere effekten av standarder baserer seg på hypotesen om at dette kan være et virkemiddel for å øke studentenes arbeidsinnsats. Det er tidligere gjort mange analyser av hvordan arbeidsinnsats påvirker prestasjonene til studentene, men resultatene på dette området har vært sprikende. Schmidt (1983), Douglas og Sulock (1995), Stinebrickner og Stinebrickner (2008) og Bonesrønning og Opstad (2010) finner en positiv sammenheng mellom arbeidsinnsatsen og resultatene. På motsatt side finner Didia og Hasnat (1998) og Krohn og O'Conner (2005) en negativ sammenheng mellom arbeidsinnsats og resultater.

En av årsakene til at litteraturen rapporterer sprikende resultater antas å være metodiske problemer knyttet til uobserverbare egenskaper hos studentene. Stinebrickner og Stinebrickner (2007) benytter institusjonelle egenskaper ved Berea College og unike data fra Barrea Panel Study. Deres identifikasjonsstrategi baseres sterkt på det faktum at Berea College tildeler romkamerater for førsteårsstudenter som er betingelsesløst tilfeldig. Dette

gjør at en students karakteristikk ved opptak, som for eksempel akademiske evner, motivasjon etc., er fullstendig ukorrelert med hans eller hennes romkamerat når disse blir tildelt. Denne detaljen gjør at man kan sortere studentene i forskjellige grupper, der hver gruppe er tilnærmet identisk når de begynner sine studier med unntak av at studentene i de forskjellige gruppene blir tildelt romkamerater av forskjellig type. Da kan forskjeller i gjennomsnittlig resultat mellom gruppene bare oppstå dersom noen faktorer, som blir bestemt etter at studentene har begynt sin utdanning, bli påvirket av den observerbare typen romkamerat. De finner i sin analyse en sterk positiv kausal sammenheng mellom arbeidsinnsats og resultater, og hevder på bakgrunn av dette at studentenes evnenivå i stor grad kan påvirkes etter opptak til høyere utdanning.

Det tidligere bidraget som ligger nærmest analyseopplegget i denne oppgaven er Krohn og O'Conner (2005). De analyserer arbeidsinnsats i et system der studentene går opp til to eksamener i samme emne i løpet av et semester, og presenterer en modell der studentene velger arbeidsinnsats før hver eksamen ved å maksimere nytten i en avveining mellom karakteren i emnet og tiden som kan benyttes til andre aktiviteter. I den empiriske analysen finner de en svak negativ sammenheng mellom arbeidsinnsats og prestasjoner. De presenterer også empirisk bevis for å bekrefte sin egen hypotese om at et bedre resultat på midtsemestereksamen vil gjøre at studentene reduserer arbeidsinnsatsen i den resterende delen av emnet.

Analysen i denne oppgaven skal undersøke betydningen standarder har på studentenes arbeidsinnsats. Med bakgrunn i en teoretisk modell basert på den Krohn og O'Conner (2005) benytter i sin analyse vil jeg betrakte hvordan disse standardene påvirker studentenes atferd til en gitt tidsdiskonteringsrate. Videre analyseres sammenhengen mellom arbeidsinnsats og prestasjoner, hvor opplegget i denne oppgaven skiller seg fra Stinebrickner og Stinebrickner (2007) ved at jeg benytter studentfaste effekter for å håndtere problemet med utelatte variabler.

3 Teorigrunnlag

For å undersøke problemstillingen for hvordan standarder, arbeidsinnsats og resultater henger sammen, vil jeg anvende teorien om nyttemaksimerende individer som utgangspunkt for den empiriske analysen. Et naturlig utgangspunkt er at studentene ønsker å maksimere nytten av den tiden de disponerer. En kan tenke seg at studentene får nytte både av resultatet de oppnår i hvert enkelt emne og den tiden de ikke benytter på å studere. Et slikt oppsett vil gjøre det mulig å analysere hvordan en standard påvirker arbeidsinnsatsen til studentene i et emne.

3.1 En teoretisk modell for arbeidsinnsats og resultater

I artikkelen “Student Effort and Performance over the Semester” (2010) utvikler Krohn og O’Conner en modell som beskriver sammenhengen mellom arbeidsinnsatsen til studenter og deres prestasjoner gjennom et semester. Denne modellen benyttes som et rammeverk for modellen jeg presenterer i dette kapittelet. De grunnleggende faktorene er de samme, men modellen i denne oppgaven er mindre omfattende og er justert for de forholdene som gjelder i mitt datamateriale.

Som et utgangspunkt for den teoretiske modellen kan vi tenke oss en student som tar et emne der det gjennomføres to obligatoriske prøver, men der den avsluttende karakteren i sin helhet avgjøres av utfallet på avsluttende eksamen. Resultatene studenten oppnår på de to prøvene antas å være en funksjon av studentens initiale egenskaper (A) (humankapitalen) og arbeidsinnsatsen studenten legger ned i forkant av de to prøvene (h_1, h_2), samt en uobserverbar tilfeldig variabel med gjennomsnittsverdi null (v_1, v_2). Den tilfeldige variabelen kan representere uforutsette faktorer som påvirker studentens prestasjoner. Denne variabelen kan fange opp utenomfaglige aktiviteter, karakteristikk ved den enkelte student eller karakteristikk ved læreren, for å nevne noe. Studentens prestasjoner på de to prøvene kan representeres ved prestasjonsfunksjonene e_t

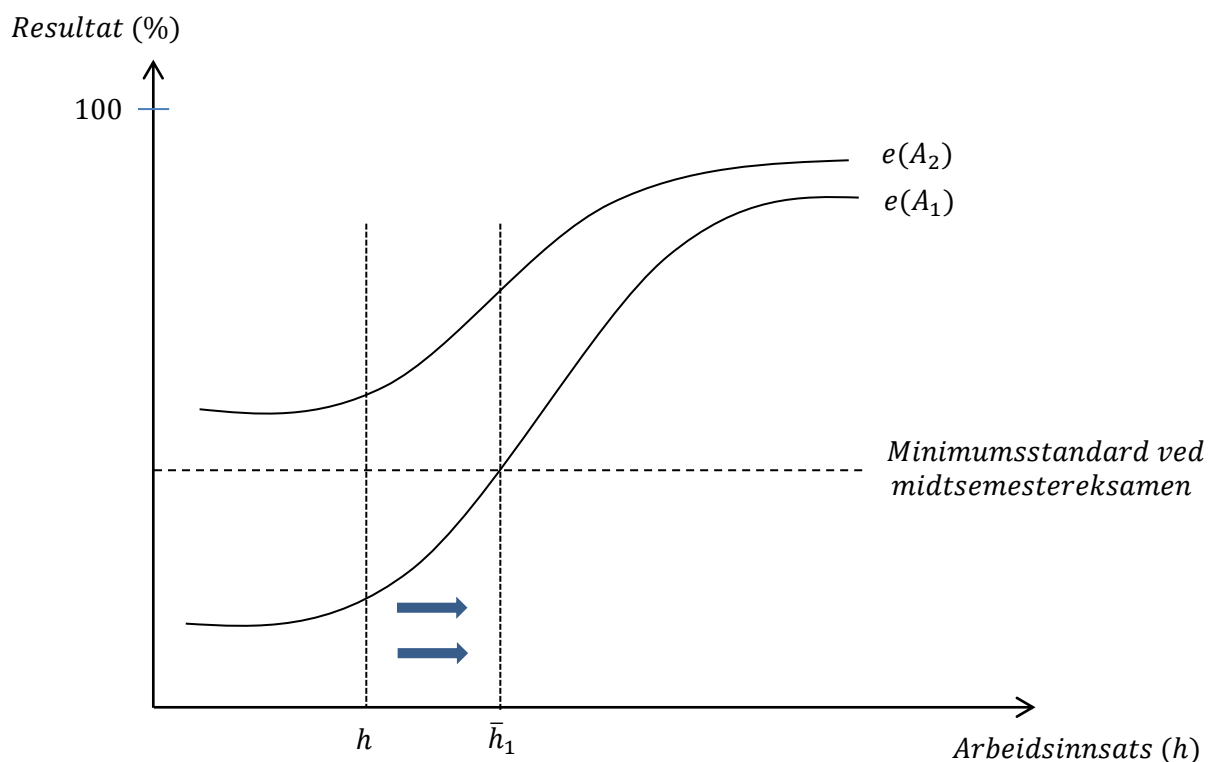
$$e_1 = f_1(A, h_1, v_1) \quad (2.1)$$

$$e_2 = f_2(A, h_1, h_2, v_2) \quad (2.2)$$

hvor e_1 og e_2 representerer henholdsvis resultatene på midtsemestereksamen og avsluttende eksamen. Det er viktig å legge merke til at studentene kan øke egne egenskaper gjennom

semesteret, noe som er representert ved at arbeidsinnsatsen før midtsemestereksamen (h_1) kan påvirke resultatene på den avsluttende eksamen.

Jeg antar her en positiv sammenheng mellom arbeidsinnsatsen og prestasjoner, men den andrederiverte er ikke nødvendigvis positiv over hele "innsatsrommet". Det kan være naturlig å tenke seg at helningen på prestasjonsfunksjonene er stigende ved lave innsatsnivåer, og at denne helningen synker etter hvert som innsatsen øker. En slik illustrasjon av prestasjonsfunksjonen er gitt i figur 3.1. I figuren er det illustrert to forskjellige prestasjonsfunksjoner, hvor hver av funksjonene representerer forskjellige initiale egenskaper hos studentene. Høyere initiale egenskaper gir bedre prestasjoner for ethvert innsatsnivå. En økning i det initiale kunnskapsnivået fra A_1 til A_2 vil gi et bedre resultat for ethvert nivå av arbeidsinnsatsen. Men det er viktig å legge merke til at denne forskjellen er mindre ved høye innsatsnivåer, noe som representeres ved at studenter med et høyt initialt evnenivå har en lavere marginal avkastning av arbeidsinnsatsen, målt i resultater.



Figur 3.1 – Resultater, arbeidsinnsats og initialt evnenivå

Studenten har nytte av resultatet som oppnås i det enkelte emnet (e_t) og tiden som ikke benyttes til arbeidsinnsats i dette emnet (L_t). Denne parameteren inkluderer tiden brukt på andre emner, arbeid, fritid osv. Det er ofte vanlig å anta myopiske studenter, og dermed at studentene har lavere nytte av en fremtidig karakter tidlig i semesteret. Vi definerer derfor to nyttefunksjoner som beskriver studentenes nytte på to ulike tidspunkter i semesteret; før midtsemestereksamen og før avsluttende eksamen

$$u_1 = \frac{u_2}{1 + \delta} + L_1(h_1) + e_1\beta \quad (2.3)$$

$$u_2 = e_1e_2 + L_2(e_1, h_2) \quad (2.4)$$

Der u_1 beskriver nytten til studentene i perioden før midtsemestereksamen og u_2 beskriver studentenes nytte i perioden før avsluttende eksamen. Parameteren e_1 kan ta to verdier; 0 eller 1. Dersom e_1 er slik at $e_1 \geq \bar{e}_1$, så tar denne parameteren verdien 1, og 0 ellers. \bar{e}_1 representerer her en minimumsstandard som må oppnås på midtsemestereksamen for at studentene skal ha muligheten til å ta avsluttende eksamen. Dersom det ikke eksisterer en slik minimumsstandard ville parameteren e_1 ta verdien 1 uansett resultat. Parameteren β representerer her verdien av å få et resultat på den avsluttende eksamen. Dette kan være verdien av å slippe å ta eksamen på nytt, stipendandelen som utløses ved bestått eksamen eller andre faktorer som gjør at nytten øker ved å bestå eksamen. Denne parameteren inngår i første periode for å representere nytteverdien av å bestå midtsemestereksamen. Sagt på en annen måte vil det og ikke bestå midtsemesterprøven gjøre at denne parameteren faller bort, og dette kan da betraktes som en reduksjon i nytten som følge av og ikke kunne ta eksamen. δ forteller oss at vi har myopiske studenter, og jo høyere verdi på denne, jo mindre betyr den avsluttende karakteren i perioden før midtsemestereksamen.

Tiden som benyttes til andre aktiviteter enn emnet er gitt som en funksjon av innsatsen, og vi kan spesifisere de på følgende måte:

$$L_1(h_1) = T_1 - h_1 \quad (2.5)$$

$$L_2(e_1, h_2) = T_2 - e_1h_2 \quad (2.6)$$

hvor T_1 representerer den totale tiden tilgjengelig før midtsemestereksamen og T_2 den totale tiden før avsluttende eksamen. Det at studenten har nytte av tiden som ikke benyttes på emnet impliserer at den marginale negative nytten av arbeidsinnsats er den samme både før og etter midtsemestereksamen.

Ettersom studentenes initiale evnenivå er gitt og variabelen v_t er uobserverbar for studentene vil deres handlingsvariabel være innsatsnivået. I figur 3.1 har jeg tegnet inn en minimumsstandard ved midtsemestereksamen, representert ved den horisontale stiplede linjen. Innsatsnivået som kreves av studentene med et initialt evnenivå gitt ved A_1 for å oppnå denne minimumsstanden er gitt ved \bar{h}_1 , slik at dersom $h_1 \geq \bar{h}_1$ så vil $e_1 \geq \bar{e}_1$. Vi kan da enkelt se fra nyttefunksjonen gitt ved relasjon (2.3) at

$$u_1 = \frac{e_2 + L_2}{1 + \delta} + L_1 + \beta \quad \text{når } h_1 \geq \bar{h}_1 \quad (2.7)$$

$$u_1 = \frac{T_2}{1 + \delta} + L_1 \quad \text{når } h_1 < \bar{h}_1 \quad (2.8)$$

Dermed kan vi konkludere med at studentene vil yte innsats $h_1 \geq \bar{h}_1$ i tiden før midtsemestereksamen dersom $e_2 + L_2 + \beta(1 + \delta) > T_2$. Det er dermed enkelt å se hvordan en innføring av en minimumsstandard påvirker arbeidsinnsatsen til studenter. I fravær av en minimumsstandard vil relasjon (2.8) falle bort og studentene vil da basere sin arbeidsinnsats før midtsemestereksamen utelukkende på resultatet av den avsluttende eksamen og tiden de kan benytte på andre ting i de to periodene. Arbeidsinnsatsen i perioden før midtsemestereksamen baserer seg da på hvor fremsynte studentene er. Det vil si jo høyere verdi på δ , jo lavere arbeidsinnsats vil studentene ha i perioden før midtsemestereksamen. En innføring av en minimumsstandard på midtsemestereksamen er altså mest effektiv på arbeidsinnsatsen i perioden før midtsemestereksamen desto mindre fremsynte studentene er. Denne effekten er relativt åpenbar ettersom en minimumsstandard tvinger studentene til å verdsette den avsluttende eksamen på et tidligere tidspunkt i semesteret. Jo mindre fremsynte studentene er, dess større blir nytteeffekten av å bestå midtsemestereksamen relativt til verdien av eksamensresultatet. Ettersom det er avtakende marginal avkastning av arbeidsinnsatsen på resultatene og avtagende marginal nytte av resultater og fritid, vil det være naturlig at innføring av en relativt lav minimumsstandard vil øke arbeidsinnsatsen for lite

fremtsynte studenter i perioden før midtsemestereksamen. En liten økning i arbeidsinnsatsen vil da føre til en relativt stor endring i resultatene dersom minimumsstandarden i utgangspunktet er null.

Det er vanskeligere å predikere hvilken effekt standarden har på arbeidsinnsatsen i perioden før den avsluttende eksamen, da en slik standard vil endre det optimale innsatsnivået i de to periodene. Hvordan studentene tilpasser seg i perioden før den avsluttende eksamen vil i stor grad avhenge av hvilke antagelser vi gjør om prestasjonsfunksjonene.

Først av alt må en betrakte at studentene før midtsemestereksamen bestemte arbeidsinnsatsen i denne perioden og sin planlagte arbeidsinnsats i den neste perioden, hvor de har antatt $v_1 = v_2 = 0$. Når studentene har gjennomført midtsemestereksamen observerer de verdien på parameteren v_1 , og vil da gjøre endringer i arbeidsinnsatsen i periode 2 basert på $v_2 = E(v_2|v_1)$. Dersom studenten observerer $v_1 > 0$ kan dette betraktes som at studenten scorer høyere enn forventet på midtsemestereksamen, og motsatt dersom $v_1 < 0$. Studenten observerer i et slikt tilfelle at arbeidsinnsatsen som er nødvendig for å oppnå et gitt resultat er mindre enn planlagt, og vil dermed redusere sin innsats i perioden etter midtsemestereksamen.

Videre er det slik at arbeidsinnsatsen i den første perioden påvirker resultatet på den avsluttende eksamen, men det er også naturlig å tenke seg at det påvirker avkastningen av arbeidsinnsatsen i den andre perioden. Dersom det er slik at avkastningen av arbeidsinnsatsen i den andre perioden er økende med arbeidsinnsatsen i den første perioden, vil dette trekke i retning av økt arbeidsinnsats. Krohn og O'Conner (2008) antar at innsatsen som benyttes før midtsemestereksamen i liten grad påvirker avkastningen av arbeidsinnsatsen i perioden før avsluttende eksamen. Gitt denne antagelsen vil det være det totale antallet timer arbeidsinnsats, gitt alt annet konstant, som avgjør resultatet på den avsluttende eksamen. Økt arbeidsinnsats i den første perioden vil da altså føre til redusert arbeidsinnsats i den andre perioden.

Det er også viktig å legge merke til at studentene ikke får høyere nytte av å gjøre det bedre enn minimumsstandarden på midtsemestereksamen. Dette fordi det ikke påvirker den avsluttende karakteren på noen direkte måte. Det kan tenkes at studentene som gjør det godt på denne eksamen får økt nytte av dette, men i denne modellen representeres det kun gjennom at arbeidsinnsatsen i perioden før midtsemestereksamen også påvirker resultatet i den neste

perioden. Dermed kan det være slik at studenter som har et relativt høyt initialt evnenivå ikke responderer på endringer i standardene. Som vi kan se av figur 3.1 vil studenter med et initialt evnenivå gitt ved A_2 oppnå minimumsstandarder selv uten å øke sin arbeidsinnsats.

En utfordring med økte standarder som jeg ikke har diskutert så langt kan være at man kan risikere at standarden settes så høyt at noen av studentene finner det mer attraktivt å droppe emnet. Standarden på midtsemestereksamen kan settes så høyt at arbeidsinnsatsen som må ytes i den første perioden for å bestå midtsemestereksamen er så høy at nytten av å droppe emnet blir høyere. Dette er naturlig ettersom vi har antatt avtagende grensenytte av fritid. Standarden som analyseres i denne i oppgaven er så lav at dette antas og ikke være en utfordring i den videre analysen.

3.2 Underhypotesene

I denne modellen vil en innføring av en minimumsstandard på midtsemestereksamen endre atferden til studentene og konsekvensene av myopia vil bli mindre ettersom studentene nå må ta høyde for at arbeidsinnsatsen i denne perioden kan være avgjørende for om man kan ta den avsluttende eksamen. Standarden fungerer på en slik måte at studentene blir tvunget til å verdsette den avsluttende eksamen høyere i den første perioden. Dette vil resultere i en høyere arbeidsinnsats i perioden før midtsemestereksamen, og isolert sett bidra til å bedre studentenes resultater.

Modellen predikerer ikke tydelig hvilken effekt standardene vil få på arbeidsinnsatsen i perioden før avsluttende eksamen. Som beskrevet over vil det være flere effekter som trekker i ulike retninger, og som baserer seg på ulike antagelser om prestasjonsfunksjonene. Det vil være hvordan studentene tilpasser seg i denne perioden som kan si noe om den totale effekten av standardene på arbeidsinnsatsen. Det er vanskelig å definere hvilken av de ulike effektene som vil dominere, og effektene vil også muligens variere mellom ulike studenter.

Videre har vi sett at et resultat på midtsemestereksamen som er bedre enn forventet vil, isolert sett, føre til at studentene justerer ned arbeidsinnsatsen sin i perioden etter midtsemestereksamen. Altså vil studentene gjennom midtsemestereksamen erfare hvorvidt deres estimerte arbeidsnivå er nok til å oppnå et gitt resultat.

På bakgrunn av den teoretiske modellen kan vi altså forvente å finne at de økte standardene fører til en økning i arbeidsinnsatsen i perioden før midtsemestereksamen. Modellen gir ingen

klar prediksjon på hva som skjer med arbeidsinnsatsen i perioden før avsluttende eksamen, men indikerer at innsatsen reduseres i denne perioden dersom studentene presterer bedre enn forventet på den avsluttende eksamen. Modellen har som forutsetning at økt arbeidsinnsats påvirker resultatene positivt, men også at avkastningen av arbeidsinnsatsen er lavere jo høyere studentenes initiale evnenivå er. Dette indikerer at studentene som har de laveste opptakspoengene vil ha den høyeste avkastningen av arbeidsinnsatsen.

4 Datamaterialet

Denne oppgaven benytter data samlet inn fra Trondheim Økonomiske Høgskole (TØH) i tidsrommet 2007-2010. Datamaterialet er samlet inn for å studere studentenes arbeidsvaner og er hentet inn ved hjelp av spørreundersøkelser blant studenter som begynte sine studier ved TØH i tidsrommet 2006-2009. Disse dataene er sammenstilt med administrative data om studentenes resultater fra videregående skole.

Ved TØH blir studentene vurdert og får sine karakterer basert på en avsluttende eksamen i hvert emne. Det er ikke slik at innsats påvirker karakteren direkte og det er generelt ingen midtsemestereksamen eller obligatoriske innleveringer som påvirker karakteren i et emne. Blant de fire emnene som gjennomføres i høstsemesteret for andreårsstudenter, er Makroøkonomi det eneste emnet med en obligatorisk midtsemesterprøve. Dette er en flervalgsprøve som ble innført i 2006, og som de to første årene ikke hadde noen krav til oppnådd poengsum. I 2008 ble det satt et krav på 10 av 32 riktige for å bestå denne prøven, og kravet ble videre økt til 13 riktige i 2009 og 2010. En del av den avsluttende eksamen i dette emnet består av en lignende flervalgsprøve som inneholder like mange spørsmål som midtsemesterprøven. Når en da sammenligner resultatene på de to tidspunktene er det antall riktige på disse flervalgsoppgavene som undersøkes. Studentene som tar emnet Makroøkonomi blir to ganger i løpet av semesteret spurt om egen arbeidsinnsats; første gang i forbindelse med midtsemesterprøven og andre gang like før avsluttende eksamen.

4.1 Deskriptiv statistikk

Det er totalt 536 observasjoner fra 2007-2010 man har opplysninger om både kjønn og opptakspoeng. Utvalget jeg skal benytte i analysen er de observasjonene der jeg har tilgang på opplysninger om arbeidsinnsatsen i forbindelse med midtsemesterprøven og avsluttende eksamen, samt resultatene på flervalgsspørsmålene på midtsemestereksamen og avsluttende eksamen. Totalt er det i dette utvalget 285 observasjoner. Jeg har undersøkt hvorvidt utvalget jeg benytter i min analyse er representativt ved å se på kjønns sammensetning og opptakspoeng i mitt utvalg i forhold til hele studentmassen på TØH. Vi ser av tabell 1 at det er totalt 47,95 prosent gutter som begynner på TØH i tidsperioden 2006-2009, mens det i utvalget er 44,91 prosent gutter. Studentene på TØH er altså relativt jevnt fordelt mellom de to kjønnene, og vi kan se at der er en noe større andel av jentene som har deltatt i

undersøkelsene, men det er ingen stor forskjell. Heller ikke når vi splitter opp tallmaterialet i forhold til årstall er det store avvik mellom hele studentmassen og utvalget vi analyserer. Kjønnsbalansen er jevn over tid i både utvalget og studentmassen som helhet.

Når vi ser på opptakspoeng ligger gjennomsnittet for studentene som begynte på TØH i perioden 2007-2010 på 53,58 poeng. Studentene som kommer inn på TØH ligger i det øverste sjiktet blant norske videregående elever. Når vi ser på utvalget ligger opptakspoengene i denne gruppen på 53,92, altså noe høyere enn studentgruppen som helhet, men allikevel bare 0,09 standardavvik over.

Generelt gir utvalget som skal analysere et godt bilde av studentmassen, og skiller seg ikke i vesentlig grad ut når det kommer til personkarakteristikk. Utvalget må derfor kunne sies å gi et representativt bilde av studentmassen ved TØH.

Tabell 4.1 – Deskriptiv statistikk – representativitet med hensyn på kjønn og opptakspoeng

	Observasjoner			Utvalg	
	Gj.snitt	Standardavvik		Gj.snitt	Standardavvik
Kjønn (536)	0,4795	0,5000	Kjønn (285)	0,4491	0,4983
2007 (151)	0,4834	0,5014	2007 (79)	0,4177	0,4963
2008 (81)	0,4321	0,4985	2008 (59)	0,4576	0,5025
2009 (169)	0,5089	0,5014	2009 (95)	0,4737	0,5020
2010 (135)	0,4667	0,5007	2010 (52)	0,4423	0,5015
Opptakspoeng	53,588	3,820	Opptakspoeng	53,917	3,818
2007	53,573	4,022	2007	53,381	4,016
2008	54,165	3,612	2008	54,086	3,731
2009	54,844	3,570	2009	55,142	3,426
2010	51,686	3,247	2010	52,302	3,618

Videre vil jeg presentere deskriptiv statistikk for de mest sentrale variablene i analysen. Dette er presentert i tabell 4.2 under, hvor alle variablene også er fordelt på de fire årene jeg analyserer. Dette gir et bilde av hvordan utviklingen har vært blant de mest sentrale variablene i analysen.

Tabell 4.2 – Deskriptiv statistikk

	Gj.snitt	Standardavvik		Gj.snitt	Standardavvik
Innsats oktober			Innsats semesterslutt		
Totalt (285)	3,31	2,19	Totalt (285)	3,48	2,40
2007 (79)	2,66	1,56	2007 (79)	3,34	2,21
2008 (59)	3,69	2,06	2008 (59)	3,96	2,65
2009 (95)	3,40	2,71	2009 (95)	2,95	1,73
2010 (52)	3,74	1,89	2010 (52)	4,13	3,12
Resultat midtsemester			Resultat eksamen		
Totalt (285)	17,74	4,95	Totalt (285)	20,87	4,75
2007 (79)	14,29	5,84	2007 (79)	18,63	4,58
2008 (59)	18,08	3,18	2008 (59)	22,61	4,11
2009 (95)	19,00	3,64	2009 (95)	21,04	4,78
2010 (52)	20,31	4,51	2010 (52)	21,96	4,46
Forventet resultat midtsemester					
Totalt (241)	15,32	4,91			
2007 (69)	13,71	5,81			
2008 (51)	14,79	3,43			
2009 (76)	16,31	4,34			
2010 (45)	16,69	5,11			

Øverst i tabell 4.2 ser en arbeidsinnsatsen i de to periodene. Denne er gitt i antall timer per uke studentene benytter utenom undervisningssituasjoner. Før midtsemestereksamen benytter studentene i gjennomsnitt 3,31 timer per uke, mens de før avsluttende eksamen benytter 3,48 timer per uke. Det er altså ikke stor forskjell mellom arbeidsinnsatsen i de to periodene, men når vi sorterer variablene etter årstall ser vi at den relative forskjellen mellom dem er synkende. Arbeidsinnsatsen før midtsemestereksamen tenderer til å øke i hele tidsrommet,

sett bort fra et lite fall i 2009. Tilsvarende trend kan vi se for arbeidsinnsatsen før avsluttende eksamen.

Resultatene på de to eksamenene er i gjennomsnitt 17,74 riktige på midtsemestereksamen og 20,87 riktige på avsluttende eksamen. For midtsemestereksamen ser vi at antall riktige svar er økende for hele tidsrommet. Resultatene på den avsluttende eksamen følger ikke samme trend, og har et toppunkt 2008. Vi ser allikevel at gjennomsnittlig resultater er bedre på avsluttende eksamen for hele tidsrommet.

Den siste variabelen er forventet resultat på midtsemestereksamen som har et gjennomsnitt på 15,32 riktige svar. Dette er lavere enn det faktiske resultatet både totalt og for hvert enkelt år. Vi kan se at antall forventet rette følger samme utvikling som faktisk antall rette, og at denne er økende i hele tidsrommet.

Det kan også være fordelaktig for deler av analysen senere i oppgaven og dele observasjonene opp basert på kjønn og opptakspoeng. Denne deskriptive statistikken er gjengitt i tabell 8.4 i appendiks og vil for enkelthets skyld bli presentert når den benyttes i analysene.

5 Estimeringsstrategi

Den beste måten å måle effekten av økte standarder ville vært ved å gjennomføre et eksperiment der en randomisert gruppe studenter møtte økte standarder, mens en kontrollgruppe ikke fikk samme behandling. Et slikt eksperiment er meget vanskelig å gjennomføre av både praktiske og etiske grunner. Datasettet jeg benytter kommer fra et naturlig eksperiment (eller kvasieksperiment), og ideen bak et slikt naturlig eksperiment er å utnytte data der ulike individer er stilt ovenfor forskjellige rammebetingelser under sin utdanning (Raaum, 1999). De ulike rammebetingelsene er i mitt datasett endret på bakgrunn av at standarden økes for hvert år, slik som beskrevet tidligere. Et naturlig eksperiment skiller seg fra et vanlig eksperiment ved at gruppene som testes ikke er randomisert.

I et naturlig eksperiment har man en kontrollgruppe som opptrer i samme omgivelsene før rammebetingelsene endres. De blir ikke utsatt for den samme endringen i standardene og vil i tallmaterialet som benyttes i denne oppgaven være studentene fra 2007-kullet. Dette året benyttes da som basisår for analysen, og disse studentene vil da fungere som kontrollgruppe for det naturlige eksperimentet.

5.1 Minste kvadraters metode

For å analysere et kvasieksperiment der jeg undersøker hvordan økte standarder påvirker arbeidsinnsatsen og resultatene over tid vil jeg ta utgangspunkt i to funksjoner. For å estimere hvordan standarden påvirker arbeidsinnsatsen vil jeg benytte følgende funksjon:

$$E_{i,t} = \alpha_0 + \alpha_1 Kj\ddot{o}nn + \alpha_2 Opptaksp + \alpha_3 D_{i,1} + \alpha_4 D_{i,2} + \alpha_5 D_{i,3} + \varepsilon_i \quad (4.1)$$

hvor $E_{i,t}$ er arbeidsinnsatsen til student i på tidspunkt t , hvor $t = 1,2$. Tidspunkt 1 representerer her rapporteringen ved midtsemestereksamen og tidspunkt 2 rapporteringen før avsluttende eksamen. Jeg har innført tidsdummier for hvert av årene, hvor 2007 er basisåret. Dummyene fungerer slik at $D_{i,1}$ tar verdien 1 dersom student i tilhører 2008-kullet og verdien 0 ellers. Tilsvarende fanger $D_{i,2}$ opp studentene i 2009-kullet og $D_{i,3}$ opp studentene i 2010-kullet.

For å estimere effekten av standarden på resultatene vil jeg benytte følgende funksjon:

$$A_{i,t} = \beta_0 + \beta_1 Kj\ddot{o}nn + \beta_2 Opptaksp + \beta_3 D_{i,1} + \beta_4 D_{i,2} + \beta_5 D_{i,3} + \varepsilon_i \quad (4.2)$$

hvor $A_{i,t}$ er resultatet til student i på tidspunkt t , hvor $t = 1,2$. Tidspunkt 1 representerer her midtsemestereksamen og tidspunkt 2 avsluttende eksamen. Tilnærmingen ellers er lik som i relasjon (4.1).

I begge modellene benytter jeg administrative data for å korrelere for kjønn og opptakspoeng. Kjønn er en dummyvariabel som tar verdien 1 dersom det er en gutt og 0 dersom det er en jente. De sentrale variablene i disse modellene er tidsdummyene. De vil i modellen vise effekten av de økte standardene etter basisåret 2007. For at tidsdummyene skal gi en riktig beskrivelse av effektene ved økte standarder må det ligge en antagelse om at det eneste som skiller de ulike studentkullene fra hverandre er kjønn, opptakspoeng og de økte standardene. Jeg forutsetter altså at alt annet er likt i de fire årene jeg analyserer. Dette er en potensielt restriktiv antagelse, da det er naturlig at flere av faktorene som påvirker studentenes arbeidsinnsats og resultater kan endre seg i dette tidsrommet. Det mest åpenbare bruddet med denne forutsetningen er den ulike sammensetningen av kullene, og dermed hvordan peer gruppe effekter kan påvirke resultatene. Samtidig er det andre faktorer utenfor de institusjonelle rammene som kan påvirke studentenes arbeidsinnsats, det være seg kultur- og idrettstilbud eller studentenes økonomiske situasjon. Videre er det også naturlig å tenke seg at endringer i de institusjonelle rammene, slik som endringer i fagsammensetningen, pensum, eller andre administrative endringer fra TØH sin side, også bryter med denne forutsetningen. Nedenfor redegjøres nærmere for potensielle brudd med forutsetningen om at alt annet er likt. Jeg diskuterer hvordan disse bruddene kan tenkes å påvirke resultatene og hvordan vi kan løse disse økonometriske utfordringene.

5.1.1 Økonometriske utfordringer

I modellen er studentenes initiale evnenivå karakterisert ved opptakspoengene fra videregående skole. Det er ingen historiske variable som forklarer hvordan studentene har prestert de to første semestrene på TØH. Dersom disse resultatene påvirker studentenes initiale evnenivå og disse historiske variablene er korrelert med forklaringsvariablene i modellene vil det skape skjevheter i estimatene (Woolridge, 2006). Jeg vil i min analyse anta

at studentenes initiale evnenivå er relativt konstant over tid, slik at opptakspoengene fra videregående skole er representativt for studentenes initiale evnenivå ved starten av emnet makroøkonomi.

Tidsdummyene fanger opp andre effekter enn de økte standardene dersom det foreligger en tidstrend i resultatene eller dersom det foreligger peer gruppe effekter, som beskriver hvordan interaksjon og påvirkning mellom individer innen en gruppe påvirker det enkelte individet i gruppen. Dersom det er slike effekter som fanges opp av tidsdummyene vil det være naturlig at vi vil kunne se den samme trenden i emner studentene har tatt i tidligere semester. For å kontrollere for dette vil jeg teste om vi kan finne samme resultater i et annet samfunnsøkonomisk emne. Jeg har da benyttet samme modell som over for å se om tidsdummyene fanger opp samme trend i emnet mikroøkonomi. Det er viktig å poengtere at disse resultatene kan være vanskelig å sammenlikne med makroøkonomi, ettersom det ikke omhandler antall rette i en flervalgseksamen, men karakteren som studentene har oppnådd i emnet. Det vil allikevel gi meg et bilde av hvorvidt tidsdummyene fanger opp en generell trend eller om det er de økte standardene som påvirker resultatene.

Av de administrative dataene som er samlet inn fra TØH er det enkelte endringer som det er vanskelig å korrelere for. Det er gjort flere strukturelle endringer før høstsemesteret 2009 som kan tenkes å påvirke estimatene; Endring i pensum, endret fagsammensetning og en annerledes gjennomføring av midtsemestereksamen.

En endring i pensumlitteraturen kan være korrelert med både innsatsnivå og resultater. Dersom den nye pensumlitteraturen gjør at studentenes marginale avkastning av arbeidsinnsatsen endres vil den kunne påvirke både studentenes arbeidsinnsats og resultater. Men ettersom læringsmålene og faglæreren i emnet ikke er endret, antar jeg at denne faktoren ikke er korrelert med hverken arbeidsinnsatsen eller resultatene til studentene.

Endring i fagsammensetningen kan tenkes å påvirke studentenes arbeidsinnsats dersom det nye emnet skiller seg fra det forrige med tanke på arbeidsinnsatsen som kreves i de andre emnene. Det vil være vanskelig å finne andre objektive mål på den totale arbeidsinnsatsen i alle emner i løpet av et semester enn antall studiepoeng. Hvordan studentene allokterer sin tid mellom ulike emner og hvordan de vurderer vanskelighetsgraden vil være et subjektivt mål. Dersom fagsammensetningen er endret slik at studentene ser på den nye sammensetningen

som “lettere” enn tidligere kan dette gi en positiv forventningsskjevhet i estimatene. Da det ikke er mulig å finne noe objektivt mål på vanskelighetsgraden av fagsammensetningen og studentene fortsatt må avlegge 30 studiepoeng, vil jeg anta at fagsammensetningen ikke er korrelert med arbeidsinnsatsen.

Høstsemesteret 2009 kan også karakteriseres som annerledes når det kommer til gjennomføringen av midtsemestereksamen. Ifølge eksamensreglementet på TØH har studentene krav på tre forsøk for å bestå midtsemesterprøven. Det betyr at dersom man ikke består på første forsøk, får man muligheten inntil to ganger til. Om man består midtsemesterprøven på første forsøk kan man se bort fra de påfølgende prøvene. I høstsemesteret 2009 var det derimot obligatorisk å delta på to av midtsemesterprøvene. Det vil si at selv om du bestod kravet på den første prøven, var det også obligatorisk å gjennomføre den andre prøven. Det er naturlig å tenke seg at studentene er risikoaverse og dermed ønsker å bestå prøven på det første forsøket, for dermed å slippe å bruke tid på flere forsøk. Når man har en ordning der man må ta midtsemestereksamen to ganger vil risikoen knyttet til ikke å bestå på første forsøk reduseres (da man må ta den to ganger uansett), og studentene har ikke samme incentiv til å yte en like høy arbeidsinnsats. På bakgrunn av dette kan det tenkes at man i 2009 har en negativ forventningsskjevhet i estimatene av effekten av standarden på arbeidsinnsatsen i perioden før midtsemestereksamen.

Høsten 2007 og 2009 ble studentfestivalen UKA arrangert i Trondheim. UKA er en festival som gjennomføres i oktober måned og baserer seg på frivillig engasjement fra nesten 2 000 studenter. Det er naturlig at dette kan ha en negativ påvirkning på arbeidsinnsatsen til de studentene som har engasjert seg i festivalen. Dersom denne festivalen gjør at studentene reduserer sin arbeidsinnsats må en kunne forvente at vi har en negativ forventningsskjevhet i estimatene av effekten av standardene på arbeidsinnsatsen disse årene. Det er allikevel ingen informasjon i tallmaterialet som indikerer hvilke studenter som engasjerer seg i denne festivalen, og derfor er det ikke mulig å korrelere for dette i regresjonene.

I den teoretiske modellen presentert tidligere i oppgaven er det usikkert hvordan standarden påvirker arbeidsinnsatsen i perioden etter midtsemestereksamen. Modellen predikerer allikevel det samme som Krohn og O’Conner (2007) finner i sin analyse, nemlig at dersom studentene gjør det godt på midtsemestereksamen har de en tendens til å redusere arbeidsinnsatsen senere i semesteret. Deres resultat er ikke direkte overførbart til min analyse,

da midtsemestereksamen på TØH ikke teller på den avsluttende karakteren. Men for å sjekke om dette kan være en plausibel forklaring også i min analyse vil jeg undersøke om eventuelle avvik mellom studentenes forventninger til antall rette svar på midtsemestereksamen og antall rette de faktisk oppnådde påvirker resultatene. Overført fra Krohn og O'Conner sine funn kan det muligens være slik at studentene reduserer sin arbeidsinnsats senere i semesteret fordi de gjorde det bedre enn forventet på midtsemesterprøven. Jeg vil derfor legge til en variabel i regresjonen for arbeidsinnsatsen før avsluttende eksamen som viser differansen mellom forventet og faktisk resultat på midtsemestereksamen.

Til slutt må det også nevnes at disse estimatene vil baseres på den rapporterte arbeidsinnsatsen blant studenter, og det er stor sannsynlighet for at det er feilrapporteringer. En måte å løse dette på hadde vært å samle informasjon på flere tidspunkter i løpet av semesteret, men slik informasjon er ikke tilgjengelig. Allikevel er det nærliggende å tro at ingen av de ulike årskullene skiller seg spesielt ut hva kommer til feilrapportering. Så dersom studentene systematiske overrapporterer/underrapporterer sin arbeidsinnsats vil den relative forskjellen mellom kullene være den samme som om rapporteringen hadde vært korrekt. Det er derfor nærliggende å tro at dette ikke vil skape skjevheter i estimatene, men jeg vil allikevel prøve å redusere disse rapporteringsproblemene ved å ekskludere ekstremobservasjoner fra tallmaterialet.

5.2 Difference in difference

Midtsemesterprøven er motivert som tiltak for å forbedre faglige prestasjoner. En viktig del av evalueringen vil derfor være å sjekke om økt arbeidsinnsats slår ut i bedre resultater. For å undersøke dette ønsker jeg å benytte meg av difference in difference tilnærming der jeg undersøker hvordan endringer i resultatene fra midtsemestereksamen til avsluttende eksamen påvirkes av endringen i arbeidsinnsatsen fra perioden før midtsemestereksamen til perioden før avsluttende eksamen.

Når jeg differensierer resultatene og arbeidsinnsatsen gjennom semesteret vil de studentfaste effektene differensieres bort når vi antar at disse er konstante gjennom semesteret. Gitt denne antagelsen vil sannsynligheten for at residualen er korrelert med den individuelle studentens arbeidsinnsats være mye mindre. Denne antagelsen er altså essensiell for at relasjon (4.4) skal gi troverdige estimater på effekten av arbeidsinnsatsen. Likningen jeg ønsker å estimere blir da følgende:

$$A_{i,2} - A_{i,1} = \alpha_0 + \alpha_1(E_{i,2} - E_{i,1}) + \varepsilon_i \quad (4.3)$$

$$\Delta A_i = \alpha_0 + \alpha_1 \Delta E_i + \varepsilon_i \quad (4.4)$$

Det er som beskrevet tidligere ingen grunn til å tro at restleddet er korrelert med arbeidsinnsatsen, men dette gjelder kun når jeg betrakter ett semester. Etersom vi her ser på effektene fra 4 forskjellige semestre kan det være grunn til å tro at restleddet er korrelert med forklaringsvariabelen. Mer konkret er det naturlig å tenke seg at standarden er korrelert med endringen i arbeidsinnsats. Jeg ønsker derfor å estimere regresjonen over der jeg legger til tidsdummyer på samme måte som ved minste kvadraters metode. På denne måten kan vi testa om standarden har andre effekter på resultatene, annet enn gjennom arbeidsinnsatsen. Tidsdummyene vil her vise hvordan standarden påvirker endringer i differansen mellom resultatet på avsluttende eksamen og midtsemestereksamen. Det kan da tenkes at tidsdummyene fanger opp effekten blant de studentene som responderer på standarden ved å jobbe like mye i begge semestre. En naturlig konsekvens av dette vil, dersom det er en positiv sammenheng mellom arbeidsinnsats og resultater, være at differansen mellom resultatene reduseres. Det kan også være slik at tidsdummyene fanger opp andre effekter av standarden enn økt arbeidsinnsats. Dette kan være at studentene får erfaring med flervalgseksamen og

dermed er bedre kjent med eksamensformen, erfaringsoverføring mellom studentene på ulike kull eller lignende faktorer. Jeg vil derfor estimere følgende relasjon:

$$\Delta A_i = \beta_0 + \beta_1 \Delta E_i + \beta_2 D_{i,1} + \beta_3 D_{i,2} + \beta_4 D_{i,3} + v_i \quad (4.5)$$

Til slutt vil jeg også legge til de studentfaste effektene. Dette er for å sikre at disse differensieres bort og dermed ikke påvirker resultatene av de ovennevnte relasjonene.

$$\Delta A_i = \delta_0 + \delta_1 \Delta E_i + \delta_2 Kjønn + \delta_3 Opptaksp + \delta_4 D_{i,1} + \delta_5 D_{i,2} + \delta_6 D_{i,3} + u_i \quad (4.6)$$

Man kan problematisere de resultatene man vil få av denne analysen, og jeg vil derfor nedenfor redegjøre for hvilke utfordringer og potensielle brudd med forutsetningene for modellene ovenfor. Jeg vil diskutere hvordan disse bruddene kan tenkes å påvirke resultatene og hvordan vi kan løse disse økonometriske utfordringene.

5.2.1 Økonometriske utfordringer

Som nevnt er det en viktig antagelse, for at estimeringen skal gi troverdige resultater, at de studentfaste effektene differensieres bort. Dette innebærer også at vi må anta at avkastningen av disse faktorene er konstante i begge periodene. De samme antagelsene må gjelde for forholdet mellom kvaliteten på lærerkreftene/undervisningen og arbeidsinnsatsen til studentene. Det kan tenkes å være slik at disse effektene er komplementære og at avkastningen av arbeidsinnsatsen er økende med kvaliteten på læreren/undervisningen. Men det er ikke sikkert at alle studenter vil respondere likt på en endring i lærerkaraktistikkene, og følgelig klarer ikke relasjon (4.4) å adressere dette problemet. Dermed kan det være et problem knyttet til utelatte variabler. Allikevel vil det i et gitt semester med studentfaste effekter innenfor et gitt emne være slik, under tilfredsstillende forutsetninger, at lærerkaraktistikkene differensieres bort (Bonesrønning og Opstad, 2011).

Det må også være slik at avkastningen av arbeidsinnsatsen må være lik i begge perioder for at estimatene skal være troverdige. Det er i hovedsak to innvendinger mot at dette er tilfelle; oppbygningen av kurset og at arbeidsinnsatsen i den første perioden påvirker avkastningen av arbeidsinnsatsen i den neste perioden. Dersom det er slik at kurset er “vanskeligere” i den

andre perioden vil det være en lavere avkastning av arbeidsinnsatsen i denne perioden, noe som kan gi en negativ forventningsskjevhet i estimatene. Om det er slik at økt arbeidsinnsats i den første perioden gjør avkastningen av arbeidsinnsatsen i den andre perioden høyere, vil de studentene som jobber mye i den første periode (lite myopiske) ha en relativt lavere arbeidsinnsats i perioden etter midtsemestereksamen.

Dersom studentene kun er opptatt av resultatene på den avsluttende eksamen, vil det ikke være noe incentiv til å gjøre det bedre enn minimumsstandarden på midtsemestereksamen. Dette gjør at de to resultatene vi differensierer mellom er forskjellige i sin natur, og kan derfor gi oss en negativ forventningsskjevhet i estimatene. Det er vanskelig å kontrollere for hvordan studentenes opptreden på en slik eksamen påvirkes av disse incentivene. Men jeg vil anta at studentene gitt en eksamenssituasjon ønsker å oppnå flest mulig riktig, uavhengig av resultatets betydning, og dermed at dette ikke vil påvirke resultatene.

Til slutt vil også disse estimatene basere seg på den rapporterte arbeidsinnsatsen blant studenter, og som nevnt tidligere er det fare for at det er feilrapporteringer. En måte å løse dette på hadde vært å samle informasjon på flere tidspunkter i løpet av semesteret, men slik informasjon er ikke tilgjengelig. Det vil allikevel være slik at dersom studentene systematisk overrapporterer/underrapporterer sin arbeidsinnsats vil dette ikke skape skjevheter i en difference in difference tilnærming, da den relative forskjellen vil være konstant. For å prøve å redusere eventuelle rapporteringsproblemer vil jeg gjennomføre samme analyser hvor jeg ekskluderer ekstremobservasjoner fra tallmaterialet.

6 Resultater og diskusjon

I dette kapitelet vil jeg presentere estimeringsresultatene. Først gis en oppsummering av den deskriptive statistikken. Videre gis en presentasjon av resultatene fra minste kvadraters metode og difference in difference, hvor jeg også fremsetter heterogene effekter og robusthetstester.

6.1 Deskriptiv statistikk

Tabell 6.1 viser deskriptiv statistikk for de mest sentrale variablene i analysen. Store deler av denne statistikken er presentert tidligere i oppgaven, men det kan være greit å ha de med i en samlet fremstilling ved begynnelsen av resultatkapitlet. I den videre analysen vil jeg bruke denne statistikken for å vurdere størrelsen på estimatene. Det er viktig å legge merke til at dette er deskriptiv statistikk for hele utvalget, og jeg benytter dette fordi det blir enklere å sammenligne resultater og gir et bedre samlet bilde av effektene

Tabell 6.1 – Deskriptiv statistikk

	Gjennomsnitt	Standardavvik
Andel jenter	0,45	0,50
Opptakspoeng	53,92	3,82
Arbeidsinnsats oktober	3,32	2,19
Arbeidsinnsats semesterslutt	3,48	2,40
Resultat midtsemester	17,74	4,95
Resultat eksamen	20,87	4,75
Forventet-Faktisk	-2,37	4,43
Δ Innsats	0,16	2,50
Δ Resultat	3,12	5,26

6.2 Estimeringsresultater minste kvadraters metode

MKM-estimatene har to nyttige formål. For det første vil det gi et godt bilde av hvordan økte standarder påvirker arbeidsinnsatsen gjennom semesteret. Videre vil det danne et godt grunnlag for videre analyse av difference in difference estimeringen av hvilken effekt arbeidsinnsatsen har på avsluttende karakterer. Tabell 6.2 viser resultatene av OLS-estimatene for relasjon (4.1) og (4.2). De to første kolonnene viser estimatene for relasjon (4.1) på tidspunktene etter midtsemestereksamen og før avsluttende eksamen. Den tredje kolonnen er tilsvarende den andre kolonnen, men her kontrollert for forskjellen mellom studentenes forventninger til og faktiske resultat på midtsemesterprøven. De to neste kolonnene viser estimatene for relasjon (4.2) på de samme tidspunktene.

Tabell 6.2 – Resultater minste kvadraters metode

	Arbeidsinnsats			Resultater	
	Oktober	Semesterslutt		Midtsemester	Eksamen
Kjønn	-0,580* (0,256)	-0,417 (0,282)	-0,407 (0,305)	0,418 (0,521)	1,389** (0,520)
Opptakspoeng	0,053 (0,035)	0,029 (0,038)	0,022 (0,042)	0,251*** (0,071)	0,341*** (0,070)
Konstant	0,0786 (1,879)	1,939 (2,073)	2,264 (2,275)	0,700 (3,823)	-0,157 (3,817)
2008	1,019** (0,368)	0,618 (0,406)	1,051* (0,444)	3,600*** (0,751)	3,681*** (0,748)
2009	0,678* (0,331)	-0,417 (0,366)	-0,291 (0,404)	4,243*** (0,675)	1,731* (0,673)
2010	1,150** (0,383)	0,833* (0,423)	1,254** (0,463)	6,278*** (0,781)	3,663*** (0,778)
Forventet-Faktisk			0,084* (0,036)		
N	285	285	241	285	285

* p<0.05; ** p<0.01; *** p<0.001

6.2.1 Effekten av økte standarder på arbeidsinnsats

Av tabell 6.2 ser en at arbeidsinnsatsen som rapporteres i forbindelse med midtsemestereksamen er tydelig positivt korrelert med tidsdummyene. Ettersom tidsdummyene i denne modellen fanger opp effekten av de økte standardene sammenlignet med basisåret 2007, ser vi fra kolonne 1 at det er en positiv korrelasjon mellom de økte standardene og arbeidsinnsatsen tidlig i semesteret. En ser at arbeidsinnsatsen til studentene øker med 1,019 timer per uke i perioden før midtsemestereksamen fra basisåret 2007 til 2008, noe som tilsvarer en økning på 0,46 standardavvik. Vi ser videre at arbeidsinnsatsen øker noe mindre i 2009 når vi sammenlikner med basisåret 2007. Dette tilsvarer en økning i arbeidsinnsatsen på 0,31 standardavvik. Den største effekten av de økte standardene finner vi i studieåret 2010, hvor vi har en økning på 0,53 standardavvik i arbeidsinnsatsen. Det er altså tydelig at standardene har en signifikant effekt på studentenes arbeidsinnsats i perioden før midtsemestereksamen. Det kan også være verdt å legge merke til at gutter har en lavere arbeidsinnsats enn jenter i dette tidsrommet.

Fra kolonne 2 ser en at det ikke er noen klare signifikante resultater på den totale arbeidsmengden til studentene som rapporteres før avsluttende eksamen. Det eneste året vi får en signifikant effekt er i 2010, hvor økningen i standardene gir en økning på 0,35 standardavvik i arbeidsinnsatsen.

Kolonne 3 viser de samme resultatene som jeg har diskutert over, men hvor jeg her inkluderer som forklaringsvariabel et mål på forskjellen mellom studentenes forventninger til antall rette på midtsemestereksamen og det faktiske resultatet. En ser da at de estimerte koeffisientene for tidsdummyene endres noe. Jeg får et statistisk signifikant resultat for 2008. Resultatene samsvarer da også tydeligere med resultatene fra kolonne 1, og vi finner nå at de økte standardene gir en økning i arbeidsinnsatsen på 0,44 standardavvik i 2008 og 0,52 standardavvik i 2010.

Man ser også at en prestasjon på midtsemestereksamen som er bedre enn forventet gir en svak negativ signifikant effekt på arbeidsinnsatsen. Dersom man scorer 1 poeng høyere på midtsemestereksamen enn forventet vil dette gi 0,04 standardavvik lavere arbeidsinnsats i perioden før avsluttende eksamen. De studentene som har høye forventninger relativt til faktisk prestasjon vil respondere med å øke sin arbeidsinnsats i neste periode. Altså vil studenter som har prestert høyere enn egne forventninger redusere sin innsats i den neste

perioden. Dette resultatet er tilsvarende det Krohn og O'Conner (2010) finner i sin analyse. Selv om midtsemesterprøven i Krohn og O'Conner sin analyse er annerledes i natur, da den teller på den avsluttende eksamen, er det fortsatt slik at en prestasjon over egne forventninger gjør at arbeidsinnsatsen justeres ned. Dette er også i samsvar med prediksjonene fra den teoretiske modellen presentert tidligere i oppgaven. En så da at denne effekten kan skyldes at studentene observerer hvilket nivå på arbeidsinnsatsen som er nødvendig for å oppnå et gitt resultat.

Disse resultatene gir en indikasjon på at innføringen av et minimumskrav til antall riktige svar for å bestå midtsemestereksamen bidrar til at studentene øker sin arbeidsinnsats. Det er ikke like tydelig hvilken effekt det har å øke dette minimumskravet videre når vi ser hvordan effektene varierer mellom de tre årene. En mulig forklaring på dette kan være at økningen i kravet til antall rette mellom 2008 og 2009 er så liten at det ikke slår ut i økt arbeidsmengde hos studentene. Det er naturlig å tenke seg at effekten ved å innføre en standard er sterkere enn å øke en allerede eksisterende standard. I dette tilfellet vil det si at man har en større effekt på arbeidsinnsatsen av å innføre et krav om minimum 10 riktige, enn man vil ha av å øke dette kravet fra 10 til 13 riktige.

Det kan være flere grunner til at man ikke finner noen klar sammenheng i perioden etter midtsemesterprøven, slik vi også fant i den teoretiske modellen. Studentenes arbeidsinnsats i denne perioden er avhengige av flere faktorer, slik som diskutert tidligere. Det kan også tenkes at studentene er myopiske også innenfor en gitt periode, slik at de øker sin arbeidsinnsats betraktelig inn mot slutten av perioden. Denne hypotesen kan forsterkes når en vet at studentene rapporterer i den kommende uken før eksamen planlegger å benytte i gjennomsnitt 53,30 timer på eksamensforberedelser. Disse timene er ikke en del av studentenes rapporterte arbeidsinnsats ved semesterlutt, noe som er i motsetningen til rapporteringen ved midtsemestereksamen hvor man blir spurt om arbeidsinnsatsen etter at eksamen er gjennomført. En annen mulig forklaring på de uklare resultatene i perioden før avsluttende eksamen kan være at det er vanskelig å vite eksakt hvorvidt studentene rapporterer sin egen arbeidsinnsats kun for perioden før avsluttende eksamen eller for semesteret som helhet.

6.2.2 Effekten av økte standarder på resultater

Fra fjerde kolonne i tabell 6.2 ser vi at det er en tydelig positiv korrelasjon mellom antallet riktige svar på midtsemestereksamen og de økte standardene målt gjennom tidsdummyene. Resultatene på midtsemestereksamen øker for hvert år som går sammenlignet med basisåret 2007. Den økte standarden gir i 2008 en økning i antall riktige på midtsemestereksamen på 0,73 standardavvik. Videre ser en at siden økte standarden gir en økning i resultatene på 0,86 standardavvik i 2009 og 1,27 standardavvik i 2010. Vi ser altså en tydelig positiv sammenheng mellom de økte standardene og studentenes resultater. Det kan være naturlig å tenke at dette henger sammen med de tidligere presenterte resultatene om at den økte standarden gir økt arbeidsinnsats hos studentene. En grunn til at vi må være forsiktige med å konkludere er at midtsemestereksamen for eksempel har gitt en skarpere og smalere definisjon av relevant pensum, som har muliggjort mer strategisk tilpasning fra studentenes side. Men hvorvidt det er en sammenheng mellom de to, kommer jeg tilbake til i de videre analysene av difference in difference estimeringen.

Fra kolonne 5 ser vi også en positiv effekt av de økte standardene, men vi kan se at effekten ikke er like sterk som ved midtsemestereksamen. Den økte standarden gir en økning i resultatene på 0,77 standardavvik i 2008, mens den i de to påfølgende årene fører til en økning på henholdsvis 0,36 standardavvik og 0,77 standardavvik.

Det er også interessant å merke seg at samtidig som studentenes resultater på midtsemesterprøven øker for hvert år, ser en også at forventningene om antall rette er økende. Studentenes oppfattelse av eget kunnskapsnivå følger altså samme utvikling som antall rette på midtsemestereksamen og antallet timer arbeidsinnsats i emnet.

En midtsemestereksamen som kun har obligatorisk oppmøte uten noen standard, slik som i 2007, vil i sin natur være mindre forpliktende ovenfor studentene. Dette kan være noe av forklaringen på hvorfor de økte standardene har en sterkere effekt på resultatene ved midtsemestereksamen, sammenlignet med avsluttende eksamen.

6.2.3 Heterogene Effekter

Selv om utvalget av studenter er jevnt fordelt mellom kjønnene og studentenes initiale evnenivå generelt ligger på et høyt nivå, er det av interesse å undersøke hvorvidt det finnes noen heterogene effekter i resultatene jeg har presentert over. Jeg har derfor gjennomført de

samme regresjonene som illustrert over, men hvor jeg har delt utvalget på bakgrunn av henholdsvis kjønn og opptakspoeng. Jeg karakteriserer lave opptakspoeng som de studentene som har en lavere poengsum enn gjennomsnittsverdien 53,92 og høye opptakspoeng for de som ligger over denne verdien. Resultatene av disse regresjonene er gjengitt i tabell 6.3 og 6.4, hvor tabell 6.3 viser heterogene effekter ved estimering av arbeidsinnsats og tabell 6.4 viser heterogene effekter ved estimering av resultater.

I den videre analysen av heterogene effekter, både i dette delkapittelet og de kommende, vil beskrivelse av forskjeller i effektene mellom kjønn og initialt evnenivå være eksplorativt. Mitt formål har ikke vært å gjøre noen stringent statistisk testing og jeg har derfor ikke testet hvorvidt det er en statistisk signifikant forskjell mellom kjønnene eller ulike initiale evnenivå.

Tabell 6.3 – Heterogene effekter Arbeidsinnsats

	Arbeidsinnsats							
	Kjønn				Opptakspoeng			
	Oktober		Semesterslutt		Oktober		Semesterslutt	
	Gutt	Jente	Gutt	Jente	Høy	Lav	Høy	Lav
Kjønn					-1,184*	-0,141	-0,968*	-0,055
					(0,505)	(0,269)	(0,447)	(0,368)
Opptakspoeng	-0,784	0,117*	-0,014	0,046	0,066	0,017	0,023	-0,014
	(0,046)	(0,048)	(0,067)	(0,046)	(0,077)	(0,082)	(0,069)	(0,113)
Konstant	6,668**	-3,511	3,956	0,990	-0,644	1,876	2,472	4,072
	(2,478)	(2,604)	(3,601)	(2,479)	(4,489)	(4,197)	(3,979)	(5,736)
2008	1,335**	0,802	1,039	0,276	0,828	1,091**	0,689	0,558
	(0,440)	(0,550)	(0,639)	(0,524)	(0,734)	(0,390)	(0,650)	(0,533)
2009	0,267	1,045*	-0,643	-0,211	1,133	0,223	-0,483	-0,317
	(0,392)	(0,498)	(0,570)	(0,474)	(0,603)	(0,386)	(0,534)	(0,528)
2010	0,840	1,275*	0,264	1,237*	1,409	0,959*	1,818*	0,464
	(0,463)	(0,567)	(0,673)	(0,540)	(0,893)	(0,372)	(0,791)	(0,509)
N	128	157	128	157	114	171	114	171

* p<0.05; ** p<0.01; *** p<0.001

I tabell 6.3 fremstilles de heterogene effektene av MKM-resultatene på arbeidsinnsats. Av de 4 første kolonnene kan vi se at de økte standarden ikke påvirker det ene kjønn mer enn det andre. En ser at den økte standardens effekt på arbeidsinnsatsen i perioden før midtsemestereksamen er sterkest blant gutter i 2008, mens utslaget er sterkest blant jentene i 2009 og 2010. Tilsvarende resultater kan en se når jeg estimerer effekten av standardene på arbeidsinnsatsen i perioden før avsluttende eksamen. Det er også få av resultatene som er statistisk signifikante. Det er heller ingen systematisk forskjell mellom de studentene som har lave opptakspoeng og de som har høye opptakspoeng. Det tydeligste resultatet en kan lese ut av disse estimatene er at det blant studentene som har høye opptakspoeng er slik at jentene yter en større arbeidsinnsats enn guttene.

Tabell 6.4 – Heterogene effekter Resultater

	Resultater							
	Kjønn				Opptakspoeng			
	Midtsemester		Eksamen		Midtsemester		Eksamen	
	Gutt	Jente	Gutt	Jente	Høy	Lav	Høy	Lav
Kjønn					0,586 (0,805)	0,387 (0,680)	0,339 (0,877)	2,121** (0,642)
Opptakspoeng	0,274* (0,113)	0,231* (0,092)	0,232 (0,121)	0,402*** (0,087)	0,207 (0,123)	0,147 (0,208)	0,384** (0,134)	0,033 (0,196)
cons	-0,162 (6,036)	1,843 (4,968)	6,552 (6,462)	-3,091 (4,692)	2,869 (7,156)	6,194 (10,595)	-2,111 (7,795)	15,196 (10,008)
2008	4,410*** (1,072)	2,921** (1,049)	4,453*** (1,147)	3,146** (,991)	3,200** (1,169)	3,954*** (0,984)	2,922* (1,274)	4,312*** (0,930)
2009	3,803*** (0,955)	4,671*** (0,950)	2,235* (1,023)	1,373 (0,897)	4,595*** (0,962)	4,196*** (0,975)	1,662 (1,048)	2,098* (0,921)
2010	6,694*** (1,128)	5,966*** (1,082)	4,640*** (1,208)	2,839*** (1,021)	9,951*** (1,424)	5,074*** (0,940)	5,493*** (1,551)	3,030*** (0,888)
N	128	157	128	157	114	171	114	171

* p<0.05; ** p<0.01; *** p<0.001

I tabell 6.4 ser vi heterogene effekter av MKM-estimatene basert på resultatene. I den første delen av tabellen har jeg delt opp resultatene etter kjønn. En ser av de to første kolonnene at det er lite som tyder på at det er noen systematisk forskjell mellom kjønnene i hvordan standardene påvirker resultatene på midtsemestereksamen. Av de to neste kolonnene kan en se at de økte standardene påvirker guttenes resultater på den avsluttende eksamen i noe sterkere grad enn jentenes resultater når vi sammenligner størrelsene på estimatene. Samtidig ser en at jenters resultater påvirkes i sterkere grad av opptakspoengene enn hva som er tilfelle for guttene. Dette kan henge sammen med gjennomsnittet av opptakspoengene for jenter ligger høyere enn gjennomsnittet av opptakspoengene for gutter, og dermed at standardene som tiltak for å øke arbeidsinnsatsen har en sterkere effekt blant guttene enn hos jentene. Denne hypotesen vil jeg komme nærmere tilbake til under difference in difference analysen senere i oppgaven.

Når jeg deler gruppene inn etter opptakspoeng, gitt av de fire siste kolonnene, ser en ingen systematisk forskjell mellom gruppene i hvordan standarden påvirker resultatene. Vi ser allikevel fra kolonne 8 at det blant gruppen som har under gjennomsnittlig opptakspoeng, er slik at gutter scorer 2,12 poeng bedre enn jenter i samme kategori på den avsluttende eksamen.

6.2.4 Robusthetstester

For at vi skal være sikre på at de resultatene jeg har presentert er robuste vil jeg her presentere de robusthetstestene jeg har foretatt.

Antagelsen om at studentenes omgivelser, sett bort fra standardene, er like i hele tidsrommet er en streng antagelse. Tidsdummyene kan også fange opp en generell trend eller forskjeller mellom kullene. Dersom det er en slik trend som fanges opp av tidsdummyene så vil man kunne observere samme utvikling i andre samfunnsøkonomiske emner som de samme studentene har tatt. For å sjekke om tidsdummyene fanger opp en tidstrend eller peer gruppe effekter har jeg derfor gjennomført en tilsvarende analyse for emnet mikroøkonomi. I tabell 8.2 i appendiks har jeg presentert resultatene av disse MKM-estimatene. Vi ser her at det kun er i 2010 at vi får et statistisk signifikant resultat, og det er lite som tyder på at vi ser den samme utviklingen i dette emnet. Det er derfor lite som tyder på at resultatene fra tidligere er drevet av en generell trend eller peer gruppe effekter.

Det er viktig å kontrollere fordelingen av svarene i undersøkelsen for å sjekke om de er noenlunde normalfordelt. Når vi måler rapportert innsats på de to tidspunktene kan man se av

fordelingsfunksjonen i appendiks at det eksisterer noen “uteliggere” i tallmaterialet. Før midtsemestereksamen rapporterer studentene at de bruker i gjennomsnitt 3,32 timer per uke egeninnsats i Makroøkonomi. Den laveste rapporterte arbeidsinnsatsen er 0 timer og det høyeste 20 timer. Kun 2 personer har rapportert at de jobber mer enn 10 timer i emnet makroøkonomi i perioden før midtsemestereksamen. Før den avsluttende eksamen jobber studentene i gjennomsnitt 3,48 timer i uken, hvor det laveste som rapporteres er 0 timer og den høyest rapporterte arbeidsinnsatsen er 15 timer. Det er her 4 personer som rapporterer en arbeidsinnsats på over 10 timer per uke. Ved å utelukke disse ekstremobservasjonene får vi en bedre normalfordeling av den innrapporterte arbeidsinnsatsen. Utvalget reduseres da til 279 observasjoner og jeg har testet i hvilken grad MKM-estimatene påvirkes av disse ekstremobservasjonene i tallmaterialet. Resultatene av denne estimeringen finnes i tabell 8.1 i appendiks.

Som tabell 8.1 viser, vil ikke resultatene påvirkes i nevneverdig grad, og naturlig nok er det resultatene knyttet til arbeidsinnsatsen som påvirkes i størst grad. Den største forskjellen finner vi i kolonne 3 der den økte standardens effekt på arbeidsinnsatsen ikke lenger statistisk signifikant for noen av de tre årene. En ser også at det ikke lenger er noen statistisk signifikant sammenheng mellom arbeidsinnsatsen og differansen i forventet og faktisk resultat. Det kan være verdt å legge merke til at effektene av standardene i denne situasjonen har en sterkere effekt på arbeidsinnsatsen i perioden før midtsemestereksamen sammenlignet med resultatene presentert over.

6.3 Estimeringsresultater difference in difference

Estimeringsresultatene fra difference in difference er rapportert i tabell 6.5. Den første kolonnen fra tabellen viser resultatene fra estimeringen av relasjon (4.4) der det ikke inkluderes noen kontrollvariabler. I de to påfølgende kolonnene inkluderes tidsdummyene, kjønn og opptakspoeng, og gir resultatene fra estimeringen av henholdsvis relasjon (4.5) og (4.6). De fire siste kolonnene i tabellen viser de heterogene effektene når vi estimerer relasjon (4.6) og deler utvalget opp etter henholdsvis kjønn og opptakspoeng.

Tabell 6.5 – Difference in difference

	ΔResultater			Kjønn		Opptakspoeng	
				Gutter	Jenter	Høy	Lav
ΔInnsats	0,401** (0,123)	0,354** (0,122)	0,353** (0,122)	0,612** (0,200)	0,241 (0,156)	0,256 (0,161)	0,504** (0,185)
2008		0,329 (0,872)	0,223 (0,873)	0,224 (1,288)	0,351 (1,183)	-0,243 (1,351)	0,627 (1,166)
2009		-1,901* (0,783)	-2,126** (0,795)	-1,010 (1,161)	-2,997** (1,087)	-2,519* (1,141)	-1,826 (1,156)
2010		-2,586** (0,904)	-2,586** (0,907)	-1,703 (1,360)	-3,117* (1,217)	-4,562** (1,646)	-1,795 (1,114)
Kjønn			0,914 (0,605)			-3,022 (0,930)	-1,691* (0,803)
Opptakspoeng			0,098 (0,082)	-0,081 (0,136)	0,188 (0,104)	0,190 (0,143)	-0,099 (0,245)
N	285	285	285	128	157	114	171

* p<0.05; ** p<0.01; *** p<0.001

Av den første kolonnen i tabell 6.5 ser vi at det på individnivå er en positivt og signifikant effekt av økt arbeidsinnsats på resultatene. Denne effekten er signifikant på 1 prosentnivå, og viser at det er en klar sammenheng mellom arbeidsinnsatsen og resultatene. Effekten reduseres noe i overgangen til kolonne 2, noe som skyldes at tidsdummyene blir lagt til som kontrollvariabel. Når vi legger til tidsdummyene reduseres estimatet fra 0,401 til 0,354. Vi ser at effekten fortsatt er positiv og signifikant, ved 1 prosent signifikansnivå. I kolonne 3 har jeg

også korrelert for kjønn og opptakspoeng. Dette påvirker som forventet ikke resultatene i nevneverdig grad, da disse variablene ikke endres mellom de to periodene. Resultatet står seg fortsatt og vi ser at en økning i arbeidsinnsatsen på 1 time per uke vil gi en økning 0,35 riktige svar på den avsluttende eksamen i forhold til midtsemestereksamen.

Basert på estimatet i kolonne 2 vil et standardavviks endring i innsatsen fra perioden før midtsemestereksamen til perioden før avsluttende eksamen gi en økning i resultatene på 0,17 standardavvik. Dette er en noe svakere effekt enn det som er funnet av Bonesrønning og Opstad (2010), hvor en økning i arbeidsinnsatsen tilsvarende et standardavvik ga en økning i 0,25 standardavvik på resultatene. Sammenlignet med Stinebrickner og Stinebrickner (2007) er resultatet en del svakere, da de finner at en økning i arbeidsinnsatsen tilsvarende et standardavvik gir en økning i resultatene på 0,9 standardavvik.

Tidsdummyene fanger også her opp effektene av de økte standardene, og viser hvordan forskjellen mellom resultatene utvikler seg over tid. De forteller oss at det er negativ sammenheng mellom de økte standardene og forskjellene i resultatene. Det vil si at de økte standardene gjør at differansen mellom resultatene på midtsemestereksamen og avsluttende eksamen blir mindre. Dette er ikke et overraskende resultat, da en innføring av et minimumskrav vil endre betydningen av midtsemestereksamen. Midtsemestereksamen hadde i basisåret 2007 ikke noen påvirkning på den avsluttende karakteren, men fra og med 2008 har den fått en indirekte påvirkning gjennom et minimumskrav som må oppnås for og i det hele tatt kunne oppnå en karakter. Dette gjør noe med prøvens natur og forskjellene i resultatene vil reduseres med en høyere minimumsstandard, gitt utgangspunktets forskjeller.

Tidsdummyene kan også her fange opp de studentene som ikke varierer sin innsats mellom de to periodene. Ettersom tidsdummyene viser effekten av standarden sammenlignet med basisåret 2007, kan det være slik at de studentene som responderer på standarden med å yte tilnærmet samme arbeidsinnsats i de to periodene også oppnår jevnere resultater. Sagt på en annen måte, de som øker arbeidsinnsatsen betraktelig fra den først perioden til den andre perioden vil også få et markant bedre resultat på den avsluttende eksamen.

6.3.1 Heterogene Effekter

De fire siste kolonnene i tabell 6.5 viser resultatene når vi deler utvalget inn etter henholdsvis kjønn og opptakspoeng. En ser at det er blant gutter at effekten av arbeidsinnsatsen på resultatene er størst, og det er også kun blant gutter at resultatet er statistisk signifikant. Effekten av økt arbeidsinnsats blant gutter er mye sterkere enn for gruppen som helhet, og viser at dersom man øker arbeidsinnsatsen tilsvarende et standardavvik, vil prestasjonene på avsluttende eksamen øke med 0,29 standardavvik.

Videre ser vi at det fortsatt er en negativ sammenheng mellom tidsdummyene og endringen i resultatet, men det er nå kun blant jentene at denne effekten er statistisk signifikant.. Det er altså slik at den økte standarden i emnet gjør forskjellen mellom midtsemestereksamen og avsluttende eksamen mindre blant jentene. Denne forskjellen i resultatene mellom kjønnene stemmer godt med forklaringen av tidsdummyene over. Fra tabell 8.4 i appendiks ser en at det er blant guttene at forskjellen i arbeidsinnsatsen er størst. Det er altså ikke overraskende slik at effekten av arbeidsinnsatsen er størst blant de som har størst skille mellom innsatsen i de to periodene. For jentene er det en mer jevn fordeling av arbeidsinnsatsen som kan tenkes å bidra til at differansen i resultatene på de to prøvene reduseres.

I de to siste kolonnene ser en effektene når vi fordeler utvalget etter opptakspoeng. En ser at det er gruppen med lave opptakspoeng som har en positiv statistisk signifikant effekt av arbeidsinnsatsen og at det er blant de med høye opptakspoeng at vi ser en statistisk signifikant effekt av tidsdummyene. Av tabell 8.4 kan vi se at samme forklaring som over vil gjelde her. Det er mindre differanse i arbeidsinnsatsen blant de med høye opptakspoeng og effektene av endret arbeidsinnsats vil da være lavere. Det er verdt å merke seg at det er slik at jentene har høyere gjennomsnittlig opptakspoeng enn guttene. Det er derfor ikke overraskende at vi ser noen av den samme effekten blant gutter som blant de med lave opptakspoeng. Det er derfor også vanskelig å konkludere med hvorvidt de heterogene effektene skyldes kjønnsmessige forskjeller eller forskjeller knyttet til studentenes initiale evnenivå.

6.3.2 Robusthetstester

På samme måte som for MKM-resultatene må jeg sjekke om difference in difference estimatene er robuste. Jeg vil derfor i denne delen gjennomføre den samme robusthetstesten som jeg gjorde i analysen tidligere, hvor jeg fjerner ekstremobservasjonene.

Jeg begynner med å fjerne ekstremobservasjonene for arbeidsinnsatsen på tilsvarende måte som tidligere. I resultatene presentert over er den gjennomsnittlige differansen i arbeidsinnsats 0,161 med et standardavvik på 2,501. Den største differansen i negativ forstand, det vil si de som har en høyere arbeidsinnsats i perioden før midtsemestereksamen sammenlignet med perioden etter, er på -16 timer. Den største differansen i positiv forstand er gitt er på 13 timer. I appendiks er det gjengitt en fordelingsfunksjon for denne variabelen, og en ser at den er relativt normalfordelt, men med lange haler i begge ender. Når jeg utelukker de 6 ekstremobservasjonen som beskrevet tidligere reduseres maksimums og minimumsverdiene seg til henholdsvis 7,25 og -9. Vi kan se av fordelingsfunksjonen for den nye variabelen at den nå har en bedre normalfordeling, hvor halene er redusert. Gjennomsnittsverdien er nå gitt ved 0,128 og standardavviket er redusert til 1.797. Ved å kjøre samme regresjon med den nye variabelen kan vi se på hvilken måte disse ekstremobservasjonene påvirker resultatene. Estimeringsresultatene er presentert i tabell 8.3 i appendiks.

Fra tabell 8.3 kan vi se at det ikke er store endringer i estimatene. Det er en positiv effekt av økt arbeidsinnsats på resultatene, og effekten er fortsatt statistisk signifikant på 1 prosentnivå. En ser også at effekten fortsatt er sterkest blant gutter og de med lave inntakspoeng. Det kan umiddelbart se ut som at arbeidsinnsatsen får en større betydning når vi fjerner ekstremobservasjonene, men det er også slik at standardavviket er redusert. En økning i arbeidsinnsatsen tilsvarende et standardavvik vil ha omtrent akkurat samme effekt på resultatene som tidligere.

7 Konklusjoner

I denne oppgaven analyseres effekten av økte standarder på arbeidsinnsats og resultater ved Trondheim Økonomiske Høyskole i perioden 2007-2010. Resultatene tyder på at det er en positiv statistisk signifikant effekt av standarden på studentenes arbeidsinnsats. Standarden som innføres er et minimumskrav for å bestå en midtsemestereksamen, og resultatene forteller at det er i perioden før denne midtsemestereksamen at vi får en signifikant effekt på arbeidsinnsatsen. Dette tyder på at en innføring av en minimumsstandard på denne prøven fungerer godt som et virkemiddel for å øke studentenes arbeidsinnsats tidlig i semesteret, og dermed bidrar til å endre studentenes adferd og redusere betydningen av myopia.

I tråd med den teoretiske modellen er det vanskelig å si noe om effekten av standardene på arbeidsinnsatsen i perioden før den avsluttende eksamen. Men vi finner en positiv statistisk signifikant effekt i 2008 og 2010 når vi korrigerer for studentenes avvik mellom forventet og faktisk resultat på midtsemestereksamen. Samtidig finner vi at de studentene som oppnår et resultat over forventet nivå på midtsemestereksamen vil redusere sin arbeidsinnsats i perioden før avsluttende eksamen. Dette resultatet overlever ikke robusthetstestene, men er allikevel i tråd med det Krohn og O'Connor (2005) finner i sin analyse. Det er ingen ting som tyder på at standarden har en negativ effekt på arbeidsinnsatsen i perioden før avsluttende eksamen og man kan se at den totale arbeidsinnsatsen til studentene har økt når man betrakter semesteret som helhet.

MKM-resultatene gir også signifikante resultater når vi ser på effekten av standarden på studentenes resultater. At standardene direkte påvirker resultatene er lite trolig, men det kan tenkes at det eksisterer en læringseffekt av midtsemestereksamen som det er vanskelig å ta høyde for i analysen. Disse resultatene kan derfor ha en positiv forventningsskjevhet. Men det er liten grunn til å tro at det er denne læringseffekten som driver resultatene, da midtsemestereksamen har vært avholdt siden 2006. Min hypotese har vært at de økte resultatene kommer på bakgrunn av den økte arbeidsinnsatsen til studentene.

Jeg finner bekreftende resultater for denne hypotesen i min difference in difference estimering. Det er en positiv kausal sammenheng mellom arbeidsinnsatsen og resultatene. Når studentene øker arbeidsinnsatsen fra den ene perioden til den andre tilsvarende et standardavvik vil dette gi en økning i resultatene på 0,17 standardavvik. Resultatet er noe svakere enn det som er funnet av Bonesrønning og Opstad (2011) og Stinebrickner og

Stinebrickner (2008), men bekrefter allikevel hypotesen om at standarder vil øke studentenes arbeidsinnsats, og derigjennom resultater.

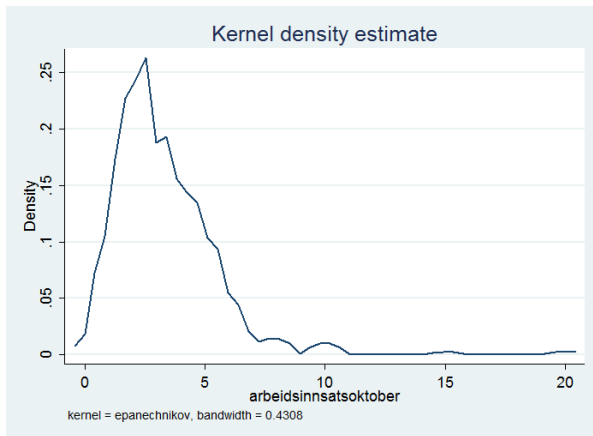
Det er flere indikasjoner på at underhypotesene kan stemme, men det er ingen klare bevis som gjør at vi kan konkludere. Det er tydelig at det er studentene som har et lavt initialt evnenivå som har størst effekt av den økte arbeidsinnsatsen. Men som vi så er det også slik at det er blant denne gruppen at differansen i arbeidsinnsats mellom de to periodene er størst. Det gjør det vanskelig å si noe sikkert om hvilke av gruppene som har størst avkastning av arbeidsinnsatsen. Effekten av arbeidsinnsatsen er også størst blant gutter når vi deler opp utvalget etter kjønn, og også her er det slik at det er denne gruppen som har størst differanse i arbeidsinnsatsen mellom de periodene. Dette gjør at det er vanskelig å konkludere hvorvidt dette er et resultat av forskjeller mellom kjønn eller forskjeller mellom grupper med ulikt initialt evnenivå.

Det finnes mange ulike typer standarder, og det er ikke mulig ut fra denne oppgaven å konkludere på hvorvidt alle former for økte standarder vil gi en høyere arbeidsinnsats og økte resultater. Det er også begrenset hvor mye man kan generalisere resultatene ettersom analysen baseres på et emne, med konstante lærekjennetegn og en eksamensform. Men ettersom analysen viser at det er mulig å endre studentenes atferd ved øke standardene og på denne måten begrense betydningen av myopia, bør dette område utforskes videre.

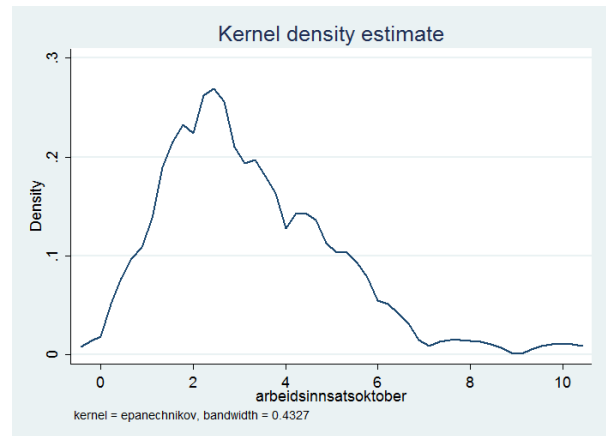
8 Appendiks

8.1 Fordelingsfunksjoner

8.1.1 Arbeidsinnsats oktober

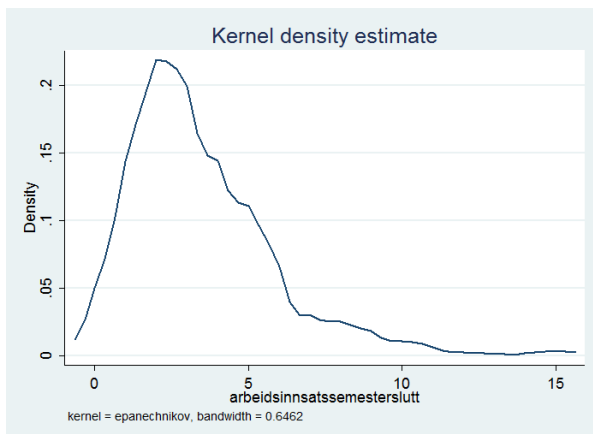


Figur 8.1
– Variasjon i arbeidsinnsats i oktober for hele utvalget

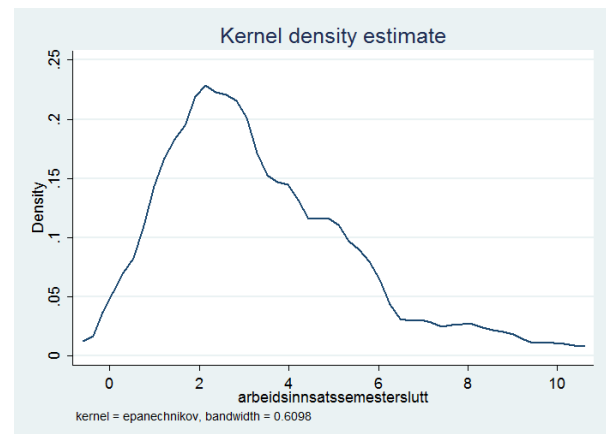


Figur 8.2
– Variasjon i arbeidsinnsats i oktober for utvalget uten ekstremobservasjoner

8.1.2 Arbeidsinnsats semesterslutt

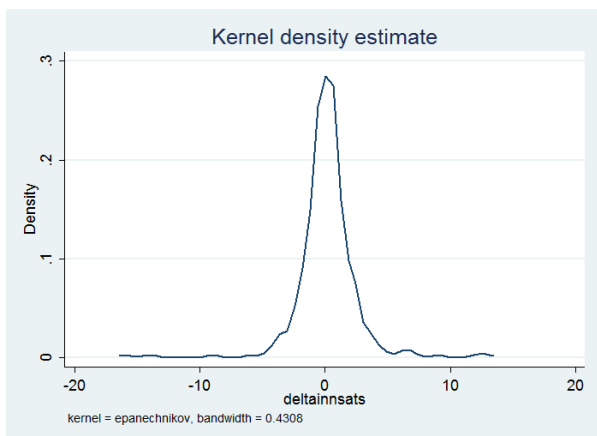


Figur 8.3
– Variasjon i arbeidsinnsats før avsluttende eksamen for hele utvalget

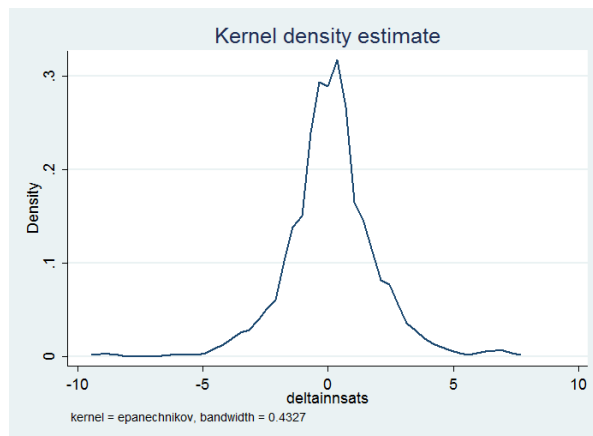


Figur 8.4
– Variasjon i arbeidsinnsats før avsluttende eksamen for utvalget uten ekstremobservasjoner

8.1.3 Δ Innsats

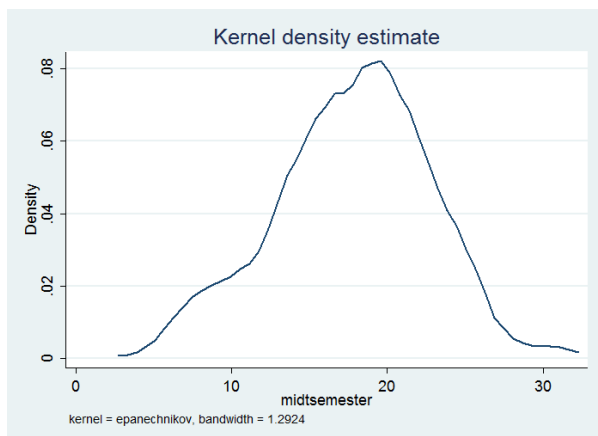


Figur 8.5
– Variasjon i endring i arbeidsinnsats for hele utvalget

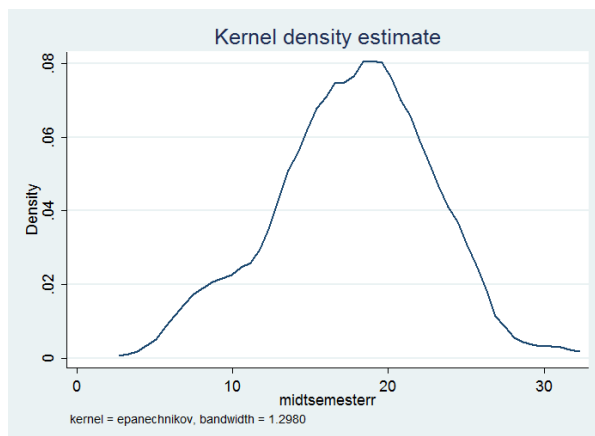


Figur 8.6
– Variasjon i endring i arbeidsinnsats for utvalget uten ekstremobservasjoner

8.1.4 Resultat midtsemestereksamen

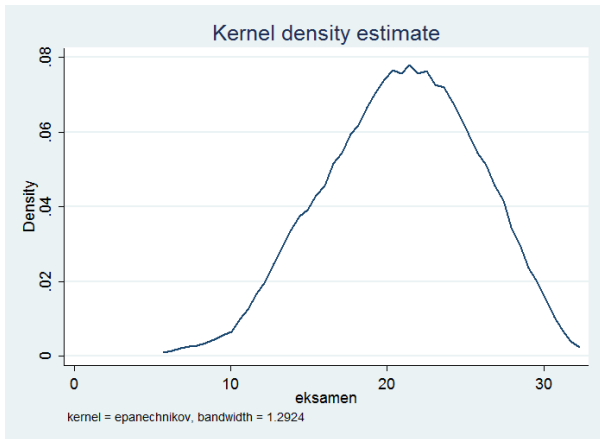


Figur 8.7
– Variasjon i resultat på midtsemestereksamen for hele utvalget

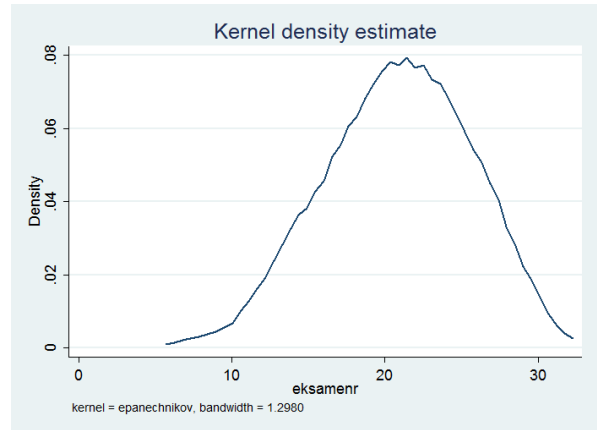


Figur 8.8
– Variasjon i resultat på midtsemstereksamen for utvalget uten ekstremobservasjoner

8.1.5 Resultat avsluttende eksamen



Figur 8.9
– Variasjon eksamensresultat for hele utvalget



Figur 8.10
– Variasjon eksamensresultat for utvalget uten ekstremobservasjoner

8.2 Tabeller

Tabell 8.1 – Resultater minste kvadraters metode – Robusthetstest

	Arbeidsinnsats			Resultater	
	Oktober	Semesterslutt		Midtsemester	Eksamen
Kjønn	-0,437* (0,217)	-0,580* (0,250)	-0,510 (0,264)	0,322 (0,529)	1,183* (0,521)
Opptakspoeng	0,002 (0,030)	0,031 (0,035)	0,024 (0,037)	0,252*** (0,073)	0,350*** (0,072)
Konstant	2,953 (1,624)	1,896 (1,871)	2,187 (2,022)	0,682 (3,963)	-0,531 (3,910)
2008	1,025** (0,312)	0,337 (0,360)	0,621 (0,384)	3,495*** (0,761)	3,465*** (0,750)
2009	0,467 (0,278)	-0,390 (0,321)	-0,341 (0,347)	4,259*** (0,679)	1,825** (0,670)
2010	1,136*** (0,325)	0,405 (0,374)	0,662 (0,347)	6,450*** (0,793)	3,781*** (0,782)
Forventet-Faktisk			0,047 (0,031)		
N	279	279	279	279	279

* p<0.05; ** p<0.01; *** p<0.001

Tabell 8.2 MKM resultater mikroøkonomi

	Resultater
	Mikroøkonomi
Kjønn	1,704
Opptakspoeng	0,087*
Konstant	-1,688
2008	0,408
2009	-0,266
2010	1,123*
N	279

Tabell 8.3 – Difference in difference – Robusthetstest

	ΔResultater			Kjønn		Opptakspoeng	
				Gutter	Jenter	Høy	Lav
ΔInnsats	0,538** (0,173)	0,459** (0,173)	0,460** (0,174)	0,876** (0,274)	0,202 (0,225)	0,042 (0,278)	0,657** (0,227)
2008		0,367 (0,892)	0,287 (0,894)	0,484 (1,338)	0,333 (1,195)	-0,295 (1,360)	0,827 (1,201)
2009		-1,850* (0,792)	-2,040* (0,805)	-0,828 (1,173)	-2,979** (1,099)	-2,691* (1,162)	-1,727 (1,164)
2010		-2,407** (0,930)	-2,334* (0,932)	-1,103 (1,425)	-3,098* (1,240)	-4,860** (1,723)	-1,535 (1,137)
Kjønn			0,914 (0,616)			0,197 (0,956)	-1,642* (0,813)
Opptakspoeng			0,098 (0,085)	-0,057 (0,139)	0,181 (0,109)	0,042 (0,154)	-0,100 (0,248)
N	279	279	279	125	154	111	168

* p<0.05; ** p<0.01; *** p<0.001

Tabell 8.4 – Deskriptiv statistikk fordelt på kjønn og opptakspoeng

	Totalt	Kjønn		Opptakspoeng	
		Gutt	Jente	Høy	Lav
Opptakspoeng	53,92	53,54	54,22		
Arbeidsinnsats oktober	3,32	2,99	3,58	3,52	3,18
Arbeidsinnsats semesterslutt	3,48	3,23	3,68	3,51	3,46
Resultat midtsemester	17,74	17,95	17,57	18,67	17,13
Resultat eksamen	20,87	21,54	20,31	21,95	20,15
Forventet resultat midtsemester	15,32	16,07	14,64	16,31	14,64
Δ Innsats	0,16	0,23	0,09	-0,01	0,27
Δ Resultat	3,12	3,59	2,75	3,28	3,02

9 Referanser

Betts, J. R. & Grogger, J. (2003): “The impact of grading standards on student achievement, educational attainment, and entry-level earnings.” *Economics of Education Review*, 22(4): 343–352.

Bonesrønning, H. (1999): “The variation in teachers’ grading practices: causes and consequences.” *Economics of Education Review*, 18: 89–105.

Bonesrønning, H. & Opstad, L.: “How much is Students’ College Performance Affected by Quantity of Study?” *International Review of Economics Education* (forthcoming)

Didia, D. & Hasnat, B. (1998): “The determinants of performance in the university introductory finance course.” *Financial Practice and Education*, 8: 102–07.

Douglas, S. & Sulock, J. (1995): “Estimating educational production functions with correction fordrops.” *Journal of Economic Education*, 26: 101–12.

Figlio, D. N. & Lucas, M. E. (2004): “Do high grading standards affect student performance?” *Journal of Public Economics*, 88(9–10): 1815–1834.

Krohn, Gregory A. & O’Connor, Catherine M. (2005): “Student Effort and Performance over the Semester.” *The Journal of Economic Education*, 36:1, 3-28

Raaum, O. (1999) “Inntektseffekter av utdanning i Norge - en litteraturoversikt”, *Arbeidsnotat Frischsenteret*, 2/1999.

Schmidt, R. M. (1983): “Who maximizes what? A study in student time allocation.” *American Economic Review*, 73: 23–28.

Stinebrickner, T. & R. Stinebrickner. (2008): “The causal effect of studying on academic performance.” *The B.E. Journal of Economic Analysis & Policy*, 8(1), Article 14

Woolridge, Jeffrey M. (2009): "Introductory econometrics a modern approach, fourth edition." *Thomson South-Western*