

## Forord

Denne oppgaven markerer slutten på studietiden min ved NTNU. Det har vært spennende, lærerike og utfordrende år som har gitt meg mange gode minner.

Jeg vil rette en stor takk til min veileder Hans Bonesrønning for glimrende og grundig veiledning, nyttige kommentarer og god tilgjengelighet. Tips og tilbakemeldinger har vært verdifullt gjennom hele prosessen, og gitt meg inspirasjon, nye ideer og en bedre faglig forståelse.

En stor takk til min kjære ektemann, Sverre Georg, for oppmuntring, gode innspill og korrekturlesing. Takk også til mamma og pappa, Erlend og Andreas for at dere heier på meg og alltid er der. Til slutt vil jeg også takke venner og bekjente som har kommet med råd og oppmuntring underveis.

Trondheim 28.11.2012

Marie Husøy Sæther



# Innholdsfortegnelse

<b>1 Innledning</b>	<b>1</b>
1.1 Problemstilling	3
1.2 Disposisjon	3
<b>2 Et utvalg tidligere empirisk forskning på kausaliteten mellom familiestruktur og elevers prestasjoner i skolen</b>	<b>5</b>
<b>3 Datamateriale</b>	<b>9</b>
3.1 Deskriptiv statistikk	9
3.1.1 Familiestruktur	10
3.1.2 Forklaringsvariabler, utvalg fra 5. klassetrinn	12
3.1.3 Forklaringsvariabler, hele utvalget	15
3.1.4 Testscore ved nasjonale prøver i lesing og matematikk	16
<b>4 Estimeringsstrategi</b>	<b>17</b>
4.1 Minste Kvadraters Metode	18
4.2 Difference in differences	20
<b>5 Resultater og diskusjon</b>	<b>25</b>
5.1 Forholdet mellom familiestruktur og familiespesifikke karakteristika	25
5.1.1 Diskusjon	27
5.2 Forholdet mellom familiestruktur og elevresultater	30
5.2.1 Testscore som funksjon av familiestruktur	30
5.2.2 Testscore som funksjon av familiestruktur og individspesifikk variabel, kjønn	31
5.2.3 Testscore som funksjon av familiestruktur, individ- og familiespesifikke variabler	31
5.2.4 Diskusjon	37
5.3 Estimeringsresultater difference in differences	39
5.3.1 Utvikling i prestasjoner som funksjon av familiestruktur	39
5.4 Effekten av familiestrukturendring i forkant av familieoppløsningen	42

5.5 Forskjeller blant familietyper i kategorien intakt familie	46
5.6 Heterogene effekter	48
5.6.1 Forskjeller i effekt av familieoppløsning mellom kjønn	48
5.6.2 Effekten av familieoppløsning blant elever i høyt og lavt utdannede familier	49
<b>6 Konklusjoner</b>	<b>53</b>
<b>7 Referanser</b>	<b>57</b>

# 1 Innledning

Flere forskere hevder at barn som opplever en familieoppløsning eller vokser opp i en ikke-intakt familie har mindre gode forutsetninger for å oppnå sterke resultater i skolen, enn barn fra intakte kjernefamilier. Mindre tid til oppfølging og færre økonomiske ressurser tilgjengelig for barn av aleneforeldre kan være avgjørende for hvordan disse elevene presterer på skolen. Oppvekst med kun en forelder gir dessuten barna færre rollemodeller å strekke seg etter i utviklingen. Opplevelsen av en familieoppløsning påvirker også i seg selv barna negativt i det en trygg hverdag brytes opp.

Utdanningsøkonomisk forskning har forsøkt å undersøke hvilke faktorer som spiller inn i elevenes resultater i skolen. En viktig variabel å inkludere har vært familiestruktur, sammensetningen av barn og voksne i nærmeste familie. Det er funnet mye støtte for en negativ sammenheng mellom elevresultat og oppvekst i ikke-intakte familier, men det har vært vanskelig å fastslå om det dreier seg om et rent kausalt forhold, der familiestruktur er bestemmende for elevenes forutsetninger i skolen.

I den politiske debatten innen utdanning og skole diskuteres det ofte hvordan elevenes læringshverdag kan effektiviseres og hvordan elevprestasjonene kan forbedres. For å kunne utforme gode politiske strategier for å styrke elevenes læring og utdanning er det nødvendig å avdekke hvilke forhold som virker inn på barnas resultat i skolen. USA har innført politiske tiltak for å øke andelen ekteskap og intakte familier som et ledd i en forbedring av barns oppvekstvilkår, (1996 Welfare reforms og Healthy Marriage Initiative, 2006 TANF reauthorization, fra Finlay og Neumark (2009)). En slik kampanje har blant annet bakgrunn i funn som underbygger en sammenheng mellom svakere elevprestasjoner og oppvekst i en blandet eller ikke-intakt familie. Dersom denne sammenhengen er overtolket, vil den politiske strategien være rettet mot gal faktor. For å kunne designe tiltak rettet mot faktorene som er av størst betydning er det relevant å forsøke å avdekke de virkelige årsakssammenhengene.

Til tross for at flere studier legger frem empiriske argument for at nivå på elevprestasjoner og familiestruktur korrelerer, er det risikabelt å tolke dette som en direkte årsakssammenheng uten utdypende analyser. Flere har funnet indikasjoner på at det ikke kun er familiebrudd som er årsaken til spredning i elevprestasjoner, men at familienes underliggende egenskaper

påvirker elevenes forutsetninger og er avgjørende for prestasjonene i skolen (for eksempel Ginther og Pollak (2004).)

Sannsynligheten for at en selvseleksjon inn i familieoppløsning foreligger er stor. Årsakene til at familier går i oppløsning kan være mange. Noen av faktorene som gjør det vanskeligere å holde familien intakt er problematisk å observere for forskere. Eksempelvis kan familiekarakteristika som mismatch mellom mor og far, svak familieorientering hos foreldre, barn som er urolige og krevende å oppdra, eller andre vanskelige forhold innad i familien, øke sannsynligheten for at familien ikke fungerer og splittes opp. Dette er familiespesifikke egenskaper det er vanskelig å få øye på og kontrollere for i studier. Videre kan man argumentere for at disse faktorene også er viktige determinanter for hvordan barna mestrer skolehverdagen og hvilke resultater de oppnår, i det konsentrasjonsevne, oppfølging fra foreldre og miljøet hjemme, er vesentlig for barns utvikling. Uobserverbare faktorer kan altså parallelt redusere forutsetningene for å holde familien intakt og svekke barnas forutsetninger i skolen. Selvseleksjon vil da konkret bety at barn fra hjem med visse karakteristika assosiert med svakere skolerresultater, oftere havner i gruppen ikke-intakte familier. Slik vil utvalget i denne gruppen i gjennomsnitt inneha andre egenskaper, enn utvalget i intakte familier. Dette kan bety at barn som vokser opp under disse forutsetningene ville prestert dårligere enn sine klassekamerater, også dersom familien deres ikke gjennomgikk en oppløsning.

I tillegg til at familiespesifikke karakteristika kan være avgjørende for elevresultater, er det naturlig å tenke seg at opplevelsen av selve familieoppløsningen ikke er uten påvirkning. Et familiebrudd fører med seg mange endringer i barnas hverdag som også kan ha betydning for elevenes mestring i skolen. Disse er begge viktige faktorer i en studie av sammenhengen mellom familiestruktur og elevers resultater.

Effekten av selve familieoppløsningen kan være sammensatt og problematisk å analysere i detalj. I tiden rundt familieoppløsningen vil sorgreaksjoner, opplevelsen av en ustabil hverdag hvor et trygt hjemmemiljø brytes opp, fiendtlig klima mellom foreldre, og endringer i hverdagen med bytte av nærmiljø, vennekrets, skole og lignende ha mye å si for barnets trivsel, konsentrasjon og skoleprestasjoner. Dette er spesielt aktuelt på kort sikt, men det kan også tenkes at noen av forholdene er av betydning over et lengre tidsrom. Familier med kun en forsørger har dessuten gjerne lavere inntekt, mindre tid til å følge opp barna, færre rollemodeller og et mindre sosialt nettverk. Dette er konsekvenser av familiestrukturen som

virker inn på barnas hverdag på en mer permanent basis, og som kan ha en effekt på elevenes prestasjoner. Vi ser altså at mange av faktorene som påvirker elevprestasjoner endres simultant ved en familieoppløsning. Det kan imidlertid være problematisk å isolere de enkelte variabelenes betydning.

## 1.1 Problemstilling

Flere internasjonale forskere observerer at elever fra ikke-intakte familier oppnår svakere resultater i skolen, enn sine medelever fra intakte familier. I denne oppgaven undersøkes sammenhengen mellom elevprestasjoner og familiestruktur for et norsk utvalg for å forsøke å komme nærmere en forståelse av om det finnes et kausalt forhold mellom variablene. Som mål på elevprestasjonene brukes resultater fra nasjonale prøver i lesing og matematikk for 5., 8. og 9. trinn i grunnskolen fra årene 2007-11. En enkel regresjon mellom testresultater og familiestruktur etableres først. Disse resultatene justeres så for kjente observerbare familiekarakteristika som man tenker seg samvarierer med familiestrukturer for å få et bedre bilde av familieoppløsningens effekt og en indikasjon på om selvseleksjon kan eksistere. Gjennom disse resultatene kan man få en indikasjon på hvor sannsynlig det er at korrelasjonen mellom elevresultater og familiestruktur gjenspeiler et kausalt forhold.

Den aktuelle problemstillingen da være:

**Kan det eksistere en kausal sammenheng mellom familiestruktur og elevers prestasjoner i den norske skolen?**

## 1.2 Disposisjon

I kapittel 2 presenteres et utvalg tidligere forskning på sammenhengen mellom familiestruktur og elevers resultat i skolen. Kapittel 3 beskriver datamaterialet og inkluderer oppgavens deskriptive statistikk. Kapittel 4 skildrer analysen med dens estimeringsstrategier og økonometriske utfordringer. Resultatene diskuteres i kapittel 5. Her presenteres også alternative analysemetoder som et supplement til punkter hvor hovedanalysen kommer til kort. Kapittel 6 presenterer oppgavens konklusjon og oppsummerer funnene.





## 2 Et utvalg tidligere empirisk forskning på kausaliteten mellom familiestruktur og elevers prestasjoner i skolen.

Det er gjort få empiriske studier i Norge som søker å finne en sammenheng mellom familiestruktur og elevprestasjoner, men en del internasjonal forskning finnes. Det er generelt bred enighet om at det eksisterer en korrelasjon mellom familiestruktur og skoleprestasjoner. Björklund, Ginther og Sundström (2006) finner en negativ sammenheng mellom oppvekst i oppløst familie og prestasjoner i skolen for et utvalg fra Sverige og USA. Francesconi, Jenkins og Siedler (2010) trekker samme konklusjon med begrunnelse i data fra Tyskland. Ginther og Pollak (2004) viser, med grunnlag i data fra USA, at prestasjoner i lesing og matematikk hos elever fra blandede familier (familier med stesøsken og halvsøsken), er merkbart lavere enn hos elever fra den tradisjonelle kjernefamilie.

Det foreligger imidlertid ikke en samlet oppfattelse av en kausal sammenheng mellom variablene. Cherlin et al. (1991) finner at elever i Storbritannia som opplevde en familieoppløsning i løpet av barneskoleårene, oppnådde dårligere resultater enn barn fra intakte familier også i forkant av familieoppløsningen. Dette indikerer at andre forhold, uavhengig av familiestruktur, er avgjørende for barnas prestasjoner i skolen. Deres funn argumenterer for en selvseleksjon i utvalget, og det kunne derav ikke konkluderes med et kausalt forhold mellom variablene. Painter og Levine (2000) går derimot ikke bort fra teorien om en årsakssammenheng mellom familiestruktur og elevprestasjoner. Som andre studier fant de at ikke-tradisjonelle familiestrukturer påvirket elevenes evner negativt. Resultatene viste imidlertid ingen klar sammenheng mellom skilsmisse eller nye ekteskap hos foreldrene og familiespesifikke karakteristika eller egenskaper hos barna. Dette tyder på at selvseleksjon ikke var et problem i utvalget i studien. Deres estimat viste heller ingen forskjell i prestasjoner mellom barn fra intakte familier og barn fra senere oppløste familier i tiden før endringen i familiestruktur inntraff. Studier for Øst-Tyskland gjort av Francesconi, Jenkins og Siedler (2010) kunne heller ikke forkaste hypotesen om kausalitet mellom oppvekst i en ikke-intakt familie og ufordelaktige forutsetninger i skolen.

Eksisterende litteratur innen emnet redegjør for mulige løsninger på endogenitetsproblemer som selvseleksjon. Ginther og Pollak (2004) inkluderer familiespesifikke karakteristika som samvarierer med familiestruktur i regresjonen for å finne et så rent uttrykk for

familiestrukturen som mulig. De viser at når det kontrolleres for andre variabler (for eksempel familieinntekt), svekkes den negative sammenhengen mellom familiestruktur og elevresultater, og blir i noen tilfeller også ikke-signifikant. Francesconi, Jenkins og Siedler (2010) finner for et utvalg fra Vest-Tyskland at den negative sammenhengen mellom oppvekst i ikke-intakte familier og elevprestasjoner halveres når det kontrolleres for foreldrenes inntekt og arbeidsstatus. Effekten er fremdeles signifikant, men reduseres sterkt. Disse studiene argumenterer for at flere faktorer relevant for elevresultatene også er årsaken til deler av familiestrukturens effekt på elevenes prestasjoner, og videre at en kausal sammenheng er mindre sannsynlig. Dette belyser viktigheten av å avdekke de ulike faktorenes påvirkning for bedre å forstå sammenhengen som undersøkes.

Også alternative metoder presenteres for å komme nærmere en konklusjon på kausalitetshypotesen. Ginther og Pollak (2004) har undersøkt forskjellene i elevprestasjoner hos halv- og stesøsken innad i den enkelte blandede familie. Det kan argumenteres for at en hverdag med begge biologiske foreldrene tilstede gir et fortrinn sammenliknet med å leve med en biologisk og en steforelder. Ginther og Pollak (2004) finner ikke en signifikant fordel ved å bo sammen med to biologiske foreldre i en blandet familie. Alle typer barn i en blandet familie ser ut til å ha dårligere forutsetninger for å oppnå gode resultat i skolen enn barn i tradisjonelle kjernefamilier. Dette kan tolkes som en støtte for at det ikke er opplevelsen av et familiebrudd eller en hverdag uten to biologiske foreldre til stede som er relevant for elevers prestasjoner, men at familiens spesifikke egenskaper og miljø avgjør barnas forutsetninger for læring og prestasjon.

Det synes vanskelig å enes om en konklusjon om en eventuell årsakssammenheng mellom familiestruktur og elevprestasjoner. Forskning av Painter og Levine (2000) kan ikke utelukke et kausalt forhold mellom familieoppløsning og elevresultater. Funn fra Ginther og Pollak (2004) og Cherlin et al. (1991) argumenterer derimot for at korrelasjonen reduseres, og i noen tilfeller blir ikke-signifikant, når familiespesifikke variabler innføres. Dette antyder en sammenheng mellom familieegenskaper og elevresultat, og familieegenskaper og familiestruktur, men i mindre grad elevresultat og familiestruktur. Disse støtter altså ikke kausalitetshypotesen, men konkluderer med at visse familier selekterer seg inn i ikke-intakte familier, og at egenskapene bak seleksjonen har noe å si for barnas skoleprestasjoner.

Biblarz og Raftery (1999) har i sin studie satt søkelyset på hvorfor forskere som undersøker denne sammenhengen lander på så ulike konklusjoner. De mener forklaringen ligger i valg av kontrollvariabler som inkluderes i regresjonene i studiene. Mange inkluderer ulike familiespesifikke variabler for å teste sammenhengen mellom familiestruktur og elevenes resultater, men hvilke som tas med varierer. Deres studie undersøkte hvorvidt familiestrukturens virkning endret seg når ulike forklaringsvariabler ble inkludert. Med utgangspunkt i data fra USA fra en 30-årsperiode, fant de grunnlag for å konkludere med at dette var tilfellet. De fant for eksempel at den negative effekten på elevenes prestasjoner av å vokse opp med kun en forelder, spesifikt alenemor, forsvant når mors sosioøkonomiske status og arbeidssituasjon ble inkludert som forklaringsvariabel.

Familiens inntekt og sosioøkonomiske status synes å være en viktig familiespesifikk variabel å inkludere. Ginther og Pollak (2004) og Biblarz og Raftery (1999) fant begge at denne i stor grad avgjorde familiestrukturens estimerte virkning på elevprestasjonene. Det å ha ressurssterke foreldre med høy utdanning og høy inntekt svekker effekten av en endring i familiestruktur på resultat i skolen. Stimuli og oppmuntring til læring, eller humankapital og økonomiske midler som investeres i barnet kan holde et relativt høyt nivå, selv med kun en forelder. Dette indikerer at ressurser og egenskaper i familiemiljøet eller hos barnet selv er viktige faktorer som bestemmer hvordan det takler endringer i hverdagen i hjemmet og i hvilken grad hendelser påvirker prestasjonene i skolen.

Et interessant spørsmål er hvorvidt jenter og gutter responderer likt på oppvekstvilkår og familiestruktur. Bertrand og Pan (2011) har undersøkt dette med fokus på atferds- og sosio-emosjonelle problemer. De fant at spesielt gutters utvikling synes å være meget var for oppvekstmiljø og stimuli hjemmefra. En hverdag med kun en forelder tilstede vil i mange tilfeller bety færre økonomiske midler og mindre tid investert i barnet, noe som ifølge forskningen var avgjørende især for gutters utvikling. Sett i sammenheng med skolehverdagen kan atferdsproblemer gi konsentrasjonsvansker, vanskeligere innpass i klassemiljø eller konflikter med lærere, og dermed redusert stimuli til læring og tregere utvikling av evner. Dersom gutters utvikling i større grad enn jenters påvirkes av trygghet, stabilitet og oppfølging hjemme, vil hendelser som familieoppløsninger eller visse egenskaper i familien mulig innvirke i større grad på gutteelevenenes prestasjoner.

Fremgangsmåten i denne oppgaven vil likne metoden brukt i Painter og Levine (2000). Analysen vil forsøke å estimere effekten av familiestruktur på elevers resultater i den norske skolen. En viktig del blir å undersøke i hvilken grad familiestrukturens eventuelle samvariasjon med elevprestasjonene reflekterer en årsakssammenheng, eller hvorvidt underliggende familiespesifikke variabler er de virkelige determinantene og en selvseleksjon finnes. Dette gjøres ved å inkludere ulike forklaringsvariabler i regresjonen, og se hvorvidt estimatoren for familiestruktur endres ved dette. Estimeringsstrategien har også mange likheter med Ginther og Pollak (2004). Funn i Cherlin et al. (1991) og Bertrand og Pan (2011) er dessuten motivasjon for alternative analyser i to av underkapitlene i kapittel 5, *Resultater og diskusjon*. Denne oppgavens bidrag til emnet vil være å komme nærmere en forståelse av hvordan elevresultatene i den norske skolen samvarierer med familiestruktur, og om det kan eksistere en selvseleksjon inn i familiestrukturer.

### 3 Datamateriale

Denne oppgaven benytter data fra nasjonale prøver i lesing og matematikk for 5., 8. og 9. trinn i den norske grunnskolen fra årene 2007-11. Til sammen omfatter dette datamaterialet over 700 000 observasjoner av elevresultater. Antallet observasjoner er noe mindre for enkelte av forklaringsvariablene analysen inkluderer, ettersom komplett informasjon ikke er tilgjengelig for alle individ. Analysen vil ta utgangspunkt i 677 119 observasjoner, andelen av utvalget som har tilknyttet informasjon om alle ønskede variabler. Den viktigste variabelen å undersøke vil være familiestruktur. Videre omfatter oppgavens analyse informasjon om elevenes kjønn, foreldres inntekt, utdanning og alder, samt om familien har innvandrerbakgrunn. Disse inkluderes som forklaringsvariabler i regresjonene av familiestruktur og testscore på nasjonale prøver.

Analysen tar utgangspunkt i en sammenlikning av testresultatene fra høstsemesteret på 5. og 8. klassetrinn. Mange elever bytter skole, klasse og lærere i overgangen mellom barne- og ungdomsskole, og får gjennom disse endringene nye forutsetninger for læring og prestering. Fordi prøvene fra 8. klasse avholdes så snart etter skolebyttet, vil elevenes faglige grunnlag fra barneskolen trolig være mer avgjørende for resultatet, enn det nye læringsmiljøet på ungdomsskolen. Slik preges sammenlikningen av elevresultat mellom disse årene i mindre grad av andre endringer, og gjenspeiler i større grad egenskaper hos elevene og familiene deres. Ved en sammenlikning mellom 5. og 9. trinn har utvalget trolig en påvirkning fra endrede omgivelser i overgangen til nytt skolemiljø som kan ha noe å si for elevenes prestasjoner og deres faglige utvikling.

#### 3.1 Deskriptiv statistikk

Tabellene under oppsummerer deskriptiv statistikk for variablene i analysen i denne oppgaven, sammenhengen mellom elevresultater og familiestruktur. Tallene omfatter gjennomsnitt, standardavvik, antall observasjoner, og maksimums- og minimumsverdier der det er relevant.

### 3.1.1 Familiestruktur

Oppgaven fokuserer på familiestrukturers rolle i barns utvikling og prestasjoner i skolen, med hovedvekt på en sammenlikning av intakte og ikke-intakte familier. For enkelt å skille mellom de to primærstrukturene, kategoriseres familiene i utvalget ved hjelp av en binær variabel, *Familiestruktur*. Denne kan anta to verdier, 1 for familier hvor foreldrene bor sammen, i ekteskap eller som samboere, og 0 i familier hvor foreldrene ikke bor sammen.

Tabell 3.1 oppsummerer utvalgets fordeling over intakte og ikke-intakte familier, gitt ved gjennomsnitt og standardavvik, for hele utvalget og spesifikt 5. og 8. trinn. Datasettet inkluderer resultat fra nasjonale prøver i 5., 8. og 9. klasse. Når oppgaven henviser til *hele utvalget* er det disse tre klassetrinnene det siktes til. Tallene viser at andelen elever som lever i en intakt familie ligger på rundt 71% når alle observasjonene inkluderes. For 5. klassetrinn ligger tallet litt høyere, 74%, mens andelen elever på 8. klassetrinn som lever i en intakt familie ligger noe under nivået for andelen i hele utvalget.

Tabell 3.1: Deskriptiv statistikk dummyvariabel familiestruktur

Variabel	Observasjoner	Gjennomsnitt	Standardavvik
Hele utvalget	677 119	0.7102	0.4537
5.klassetrinn	275 353	0.7423	0.4374
8.klassetrinn	286 170	0.6931	0.4612

Datasettet oppgaven tar utgangspunkt i angir informasjon om elevenes familiestruktur gjennom kategorier av familietyper fra 1 til 6. Eksempelvis er en familie hvor begge foreldrene lever sammen i ekteskap familietyptype 1, en familie hvor foreldrene er samboere familietype 2. Fordelingen av familiestrukturene over verdiene 1-6 er gjort som vist under. I en familie hvor begge foreldrene lever sammen, altså familier av type 1 eller 2, vil den binære *Familiestruktur*-variabelen ta verdien 1. For resten av utvalget, hvor foreldrene ikke bor sammen, familietyperne 3-6, tar *Familiestruktur* verdien 0.

Familietyper:

1. Begge foreldre og gifte
2. Begge foreldre og samboere
3. Bare mor
4. Mor og stefar
5. Bare far
6. Far og stemor

Tabell 3.2 presenterer fordelingen av elever over de seks familietypekategoriene. Den største prosentandelen elever finnes under kategori 1, gifte foreldre. Familietyper som inkluderer nest flest elever er alenemødre. En annen stor gruppe er samboende foreldre. Familier med kun far eller far og stemor utgjør de minste gruppene i utvalget. Familier med mor og stefar favner i underkant av en tidel av utvalget. Når antall familier i de ulike familietyper undersøkes nærmere, ser man at andelen intakte familier reduseres i løpet av tiden mellom 5. og 8. klasse. I særlig grad gjelder oppløsningen familier hvor foreldrene er samboere. Andelen familier fra familietype 1, gifte foreldre, holdes mer stabil. Alle typer ikke-intakte familier, dvs. kategori 3-6, øker i prosentandel fra 5. til 8. klasses trinn. Familietype 3 og 4, hvor barnet bor hos sin biologiske mor, men ikke biologiske far, har den største økningen.

Tabell 3.2: Deskriptiv statistikk familietyper.

Familie- type	Hele utvalget		5. klasse- trinn		8. klasse- trinn		Endring 5-8 kl. %poeng
	Observasjoner	%	Observasjoner	%	Observasjoner	%	
1	399 542	59.0	164 981	59.9	168 053	58.7	Ned 1.2
2	81 366	12.0	39 403	14.3	30 292	10.6	Ned 3.7
3	104 116	15.4	40 164	14.6	45 102	15.8	Opp 1.2
4	58 547	8.7	19 848	7.2	27 233	9.5	Opp 2.3
5	23 322	3.4	7 924	2.9	10 591	3.7	Opp 0.8
6	10 226	1.5	3 033	1.1	4 899	1.7	Opp 0.6
Sum	677 119	100	275 353	100	286 170	100	-

### 3.1.2 Forklaringsvariabler, utvalg fra 5.klasstrinn

Tabellene 3.3-3.4 viser deskriptiv statistikk for forklaringsvariablene. Disse sammenfatter gjennomsnitt, standardavvik, antall observasjoner og minimum og maksimum for forklaringsvariabler i analysen for et utvalg avgrenset til 5. klasstrinn. Elevens kjønn, foreldres inntekt, utdanning og alder, samt eventuell innvandrerbakgrunn er variabler som inkluderes i regresjonen av elevresultater. Variablene inkluderes også i en regresjon av familiestruktur for å avdekke om selvseleksjon inn i familietyper kan eksistere.

Variabelen kjønn kan anta to verdier, 1 dersom eleven er gutt, 2 for jente. Tallene for foreldrenes inntekt er avrundet til nærmeste krone. Observasjoner hvor inntekten er lavere enn null er utelatt, mens utvalgets høyeste inntekt er inkludert som en referanse til spredningen i utvalget. Variablene for foreldrenes utdanning kan anta verdiene 0-8. Hvordan de ulike nivåene er fordelt over numrene oppsummeres under. Andelen i utvalget hvor utdanningsnivå er ukjent er for øvrig eliminert fra datamaterialet som brukes i analysen. Utvalget er også begrenset til å favne observasjoner hvor foreldrene er maksimum 80 år gamle, dette for å luke ut feilrapporteringer. Laveste og høyeste alder for hvert utvalg er oppgitt i parentes. Første- og andregenerasjonsinnvandrere er dummyvariabler som tar verdien 1 for henholdsvis første- og andregenerasjons innvandrere, 0 for elever som har minst en etnisk norsk forelder.

Utdanningsnivå:

0. Ingen utdanning etter førskole

1. Barneskoleutdanning

2. Ungdomsskoleutdanning

3. Videregående, grunnutdanning (GK/VG1, VK1/VG2, realskole og folkehøgskole)

4. Videregående, avsluttende utdanning (VK2/VG3 og gymnasium)

5. Påbygg til videregående utdanning, forkurs

6. Universitets- og høyskoleutdanning, lavere nivå

7. Universitets- og høyskoleutdanning, høyere nivå

8. Forskerutdanning



Tabell 3.3: Forklaringsvariabler, 5.klasstrinn, 2007-09

	2007 n = 55 759		2008 n = 55 404		2009 n = 55 926	
	Gjennom- snitt	Standard- avvik	Gjennom- snitt	Standard- avvik	Gjennom- snitt	Standard- avvik
Kjønn	1.4902	0.4999	1.4918	0.4999	1.4894	0.4999
Fars inntekt	511 836 maks: 4.23e+7	510 916	511 972 maks: 2.89e+7	443 153	511 973 maks: 2.36e+7	425 533
Mors inntekt	297 309 maks: 7 030 265	206 399	292 492 maks: 4 291 073	198 633	291 368 maks: 8 623 921	207 165
Fars utdanning	4.2355	1.7378	4.2881	1.7329	4.33	1.716
Mors utdanning	4.4037	1.6809	4.4706	1.6707	4.5288	1.6529
Fars alder	42.1096 (24-78)	5.9164	42.2693 (26-80)	5.8787	42.4553 (25-80)	5.8956
Mors alder	39.2629 (23-61)	5.0132	39.4403 (24-62)	4.9841	39.5907 (24-60)	4.9763
Innvandrer førstegen	0.0135	0.1152	0.0131	0.1138	0.0109	0.104
Innvandrer andregen	0.0377	0.1904	0.0401	0.1962	0.0438	0.2045

Tabellene 3.3-3.4 angir gjennomsnittsverdiene av variablene for ulike årsklasser av utvalget fra 5. trinn, samt spredningen i observasjonene. Observasjonene er i all hovedsak jevnt fordelt over kjønnene, men utvalget har en liten overvekt av gutter. Fars inntekt ligger i gjennomsnitt drøye 200 000 over mors inntekt. Når det gjelder utdanningsnivå, ligger imidlertid tallene for mor noe høyere enn de tilsvarende for far. Forskjellen er omtrent 0.2 over hele utvalget. Dette er en relativt liten differanse, men siden den viser seg som systematisk over hele utvalget er den verdt å kommentere. Fars alder ligger i gjennomsnitt jevnt knappe 3 år over mors alder, henholdsvis på 42- og 39-tallet.

Tabell 3.4: Forklaringsvariabler, 5.klasstrinn, 2010-11.

	2010 n = 55 448		2011 n = 52 816		Alle år n = 275 352	
	Gjennom- snitt	Standard- avvik	Gjennom- snitt	Standard- avvik	Gjennom- snitt	Standard- avvik
Kjønn	1.4884	0.4999	1.492	0.4999	1.4903	0.4999
Fars inntekt	510 711 maks: 4.08e+7	467 752	517 136 maks: 3.08e+7	425 421	512 681 maks: 4.23e+7	455 990
Mors inntekt	286 889 maks: 5 713 632	200 756	302 287 maks: 4 573 901	203 628	293 990 maks: 8 623 921	203 416
Fars utdanning	4.3715	1.6972	4.4047	1.7185	4.3251	1.7216
Mors utdanning	4.5678	1.6627	4.6374	1.6584	4.5205	1.6671
Fars alder (25-79)	42.5742	5.8666	42.7349 (25-75)	5.8457	42.4255 (24-80)	5.8851
Mors alder (25-58)	39.7497	5.0045	39.9163 (25-61)	5.0104	39.585 (23-62)	5.0027
Innvandrer førstegen	0.0108	0.1034	0.0103	0.1011	0.01174	0.1077
Innvandrer andregen	0.0465	0.2106	0.0453	0.2081	0.0427	0.2021

Andelen andregenerasjonsinnvandrere øker med tiden, mens prosenten førstegenerasjonsinnvandrere reduseres marginalt over årene. Utvalget fra 2007 viser en andel første- og andregenerasjonsinnvandrere på 1.4% og 3.8%. Dette endres til 1% og 4.5% i 2011. I gjennomsnitt ligger andelene i utvalget fra hele tidsperioden på 1.2% og 4.3%. Kort oppsummert er det få uventede funn i denne deskriptive statistikken. Utvalget har omtrent lik andel fra begge kjønn, far har i gjennomsnitt høyere inntekt enn mor og er noen år eldre, utdanningsnivå synes å være relativt likt på tvers av kjønn, og andelen innvandrere er relativt liten.

### 3.1.3 Forklaringsvariabler, hele utvalget

Neste tabell inneholder gjennomsnitt, standardavvik og antall observasjoner for forklaringsvariablene for hele utvalget, altså 5., 8. og 9. klasstrinn for alle årene 2007-11.

*Tabell 3.5: Forklaringsvariabler, hele utvalget*

Variabel	Observasjoner	Gjennomsnitt	Standardavvik
Kjønn	677 119	1.4888	0.4999
Fars inntekt	677 119	513 136	470 536
Mors inntekt	677 119	299 225	206 985
Fars utdanning	677 119	4.2617	1.7358
Mors utdanning	677 119	4.3424	1.6781
Fars alder	677 119	44.1009	6.0659
Mors alder	677 119	41.256	5.2108
Innvandrer, førstegen	677 119	0.0144	0.119
Innvandrer, andregen	677 119	0.0396	0.1951

Tallene i denne tabellen er i stor grad forenlig med resultatene diskutert over. Det finnes en marginal overvekt av gutter i utvalget, fars inntekt ligger omtrent 200 000 over mors, og mors utdanning ligger i gjennomsnitt litt høyere enn fars. Det er fremdeles tre års forskjell i gjennomsnittlig alder for far og mor, men aldersgjennomsnittet er noe høyere når alle klasstrinn inkluderes. Andelene av utvalget første- og andre generasjons innvandrere er henholdsvis 1.4% og 4%.

### 3.1.4 Testscore ved nasjonale prøver i lesing og matematikk

Tabell 3.6: Deskriptiv statistikk resultat på nasjonale prøver. Antall observasjoner i parentes.

År	Variabel	5. trinn			8. trinn		
		Gjennomsnitt	Standardavvik	Poengrange	Gjennomsnitt	Standardavvik	Poengrange
2007	LES	17.0926 (54 172)	6.4627	0-29	27.436 (54 525)	7.9385	0-43
	MAT	31.5846 (52 244)	10.5611	0-51	40.9909 (54 408)	15.4865	0-76
2008	LES	18.5927 (52 135)	5.4539	0-29	27.1907 (53 389)	9.3211	0-48
	MAT	26.3203 (52 442)	9.4276	0-47	33.1335 (54 492)	12.0101	0-58
2009	LES	20.0442 (52 054)	6.6635	0-32	28.6562 (53 858)	8.344	0-46
	MAT	26.1003 (53 111)	9.2287	0-48	29.7538 (55 225)	11.9839	0-59
2010	LES	19.0641 (50 604)	6.4443	0-33	23.1041 (53 233)	7.4093	0-42
	MAT	25.9183 (53 186)	8.8965	0-44	31.3304 (55 114)	11.6111	0-58
2011	LES	21.5318 (50 072)	5.8101	0-32	27.811 (53 655)	8.0112	0-45
	MAT	26.2205 (50 155)	8.4562	0-45	32.7333 (53 484)	11.5795	0-58

Tabell 3.6 angir deskriptiv statistikk for testscorene ved nasjonale prøver i lesing og matematikk for 5. og 8. klassetrinn. Poeng-rangen for de ulike fagene over analyseårene varierer, slik at en direkte sammenlikning av resultatene er vanskelig. I oppgavens analyse standardiseres poengscorene til gjennomsnitt 0 og standardavvik 1.

## 4 Estimeringsstrategi

Det finnes ulike årsaker til at familier går i oppløsning. Ofte kan kilden til vanskelighetene med å holde familien intakt ligge i enhetens egne egenskaper og hvordan samspillet fungerer medlemmene imellom. Noen av disse egenskapene er lett å observere og måle, mens andre er mer problematisk å sette fingeren på. Eksempler på slike karakteristika er presentert innledningsvis i oppgaven. Dersom man tenker seg at egenskapene som øker sannsynligheten for en familieoppløsning parallelt også svekker elevenes forutsetninger i skolen, står vi overfor et problem med selvseleksjon i analysen. Familier med visse egenskaper går oftere i oppløsning enn andre familier. Dersom elever fra familier med slike egenskaper også i gjennomsnitt presterer dårligere på skolen enn sine klassekamerater, vil elever fra denne gruppen fremstå som svakere i skolen enn elever fra intakte familier. Dette betyr ikke nødvendigvis at det foreligger en kausal sammenheng mellom å leve i en oppløst familie og å prestere dårligere enn elever fra intakte hjem, men snarere at andre egenskaper som avgjør at elevprestasjonene i gjennomsnitt er relativt svake, simultant setter elevene i gruppen ikke-intakt familie.

Noen av disse familiekarakteristikaene er observerbare og mulig å kontrollere for i studier. Det kan tenkes at for eksempel inntekts- og utdanningsnivå hos foreldre innvirker på elevenes læreevner og resultat, og i tillegg har noe å si for forutsetningene for at familien forblir intakt. Informasjon om slike variabler er lett å inkludere i en analyse, slik at sammenhengen rundt faktorene kan undersøkes. Vanskeligere observerbare variabler som mismatch mellom foreldre, urolig hjemmemiljø eller mangel på familieorientering er det imidlertid mer problematisk å hente inn informasjon om. Det at familier som selekterer seg til oppløsning kan ha uobserverbare karakteristika, gjør det problematisk å finne betydningen av selve familieoppløsningen.

Mange av familiens egenskaper endres også i tiden rundt et familiebrudd. Familiens inntekt reduseres i mange tilfeller, foreldre har mindre tid til oppfølging av barna, og hverdagsmønsteret endres i det mange barn pendler mellom flere boliger eller bytter bosted og skolekrets. Dette er eksempler på omveltninger i barnas oppvekstmiljø grunnet et familiebrudd. I en studie av familiestrukturens innvirkning på skoleprestasjoner er det relevant å forsøke å finne disse enkeltendringenes betydning for elevenes testresultater. Er

bortfall av økonomiske midler spesielt vesentlig, eller har tap av rollemodeller større konsekvenser for barna? Å isolere og måle disse enkelteffektene kan være problematisk.

De overnevnte aspektene er denne studiens to største økonometriske utfordringer.

#### 4.1 Minste Kvadraters Metode

For å undersøke problemstillingen om hvordan elevenes prestasjoner i skolen og familiestruktur henger sammen benyttes først regresjoner basert på *Minste Kvadraters Metode (MKM)*. Elevenes resultater fra nasjonale prøver i lesing og matematikk for 5. og 8. trinn for 2007-11 modelleres som en lineær funksjon av familiestruktur, som etter hvert bygges ut med individ- og familiespesifikke forklaringsvariabler. Variablene  $D_{i,t}$  er tids-dummyer for årene  $t=2008-11$ . 2007 er valgt som basisår. Dummyene inkluderes for å fange opp generelle effekter som påvirker prøveresultatene for hele utvalget et gitt år. Den avhengige variabelen  $LES_{i,k,t}$  er estimert testresultat i lesing for elev  $i$ , på klassetrinn  $k$  for året  $t$ .  $REG_{i,k,t}$  er tilsvarende testscoreestimat i matematikk.  $\varepsilon_i$  og  $v_i$  er restledd i modellen. Produktfunksjonene for elevresultat i lesing og matematikk tar denne formen, og bygges suksessivt ut med individ- og familiespesifikke variabler:

$$LES_{i,k,t} = \alpha_0 + \alpha_1 \text{familiestruktur} + \varepsilon_i$$

$$LES_{i,k,t} = \alpha_0 + \alpha_1 \text{familiestruktur} + \alpha_2 \text{kjønn} + \varepsilon_i$$

$$\begin{aligned} LES_{i,k,t} = \alpha_0 + \alpha_1 \text{familiestruktur} + \alpha_2 \text{kjønn} + \alpha_3 \text{farsinntekt} + \alpha_4 \text{morsinntekt} \\ + \alpha_5 \text{farsutdanning} + \alpha_6 \text{morsutdanning} + \alpha_7 \text{farsalder} \\ + \alpha_8 \text{morsalder} + \alpha_9 \text{førstegeninnvandrere} + \alpha_{10} \text{andregeninnvandrere} \\ + \alpha_{11} D_{i,2008} + \alpha_{12} D_{i,2009} + \alpha_{13} D_{i,2010} + \alpha_{14} D_{i,2011} + \varepsilon_i \end{aligned}$$

$$REG_{i,k,t} = \beta_0 + \beta_1 \text{familiestruktur} + v_i$$

$$REG_{i,k,t} = \beta_0 + \beta_1 \text{familiestruktur} + \beta_2 \text{kjønn} + v_i$$

$$\begin{aligned} REG_{i,k,t} = \beta_0 + \beta_1 \text{familiestruktur} + \beta_2 \text{kjønn} + \beta_3 \text{farsinntekt} + \beta_4 \text{morsinntekt} \\ + \beta_5 \text{farsutdanning} + \beta_6 \text{morsutdanning} + \beta_7 \text{farsalder} \\ + \beta_8 \text{morsalder} + \beta_9 \text{førstegeninnvandrere} + \beta_{10} \text{andregeninnvandrere} \\ + \beta_{11} D_{i,2008} + \beta_{12} D_{i,2009} + \beta_{13} D_{i,2010} + \beta_{14} D_{i,2011} + v_i \end{aligned}$$

$$k = 5, 8 \quad t = [2007, 2011]$$

Innledningsvis i kapitlet ble to aspekter som kan komplisere studien av effekten av familiestruktur på elevresultater presentert. Disse løses ikke ved å bruke MKM på tverrsnittsdataene. For det første kan det eksistere en selvseleksjon inn i familiestrukturer basert på uobserverbare kjennetegn ved familiene eller elevene. Dette tilsier at MKM-estimatene kan være skjeve fordi variabelen familiestruktur korrelerer med restleddet. Ved å legge til observerbare familiekarakteristika som kontrollvariabler elimineres noe av faren for skjevhet i estimatoren for *Familiestruktur*, men dersom uobserverbare egenskaper er tilstede, vil det fremdeles eksistere en skjevhet. En alternativ analysemetode som løsning på dette problemet er *Difference in differences*, en analyse basert på endringsvariabler. Ved hjelp av denne metoden elimineres konstante familiekarakteristika, dermed også uobserverbare faktorer, slik at effekten av selve endringen i familiestruktur står alene tilbake. For det andre kan det være vanskelig å isolere effektene av enkeltfaktorer som virker inn på elevresultatene, og som endres simultant med en familieoppløsning. En mulig løsning er å inkludere slike variabler man tror endres i en *difference in differences*-regresjon. Dersom disse varierer over tidsrommet vil de ikke differensieres bort, og effektene vil kunne avdekkes. Endringsvariabler lages for familieegenskapene, og estimatorene for disse sier noe om faktorenes innvirkning på elevenes progresjon i skolen.

Analysen møter i tillegg noen utfordringer med tanke på avgrensningen av utvalget. Som nevnt i kapittel 3 er noen observasjoner trukket ut av datamaterialet grunnet mangelfull informasjon. Eksempler er ukjent nivå på utdanning, feilrapporteringer om foreldres alder eller oppgitt negativ inntekt. Det kan tenkes at dette i særlig grad berører visse deler av befolkningen, eksempelvis grupper med lavere sosioøkonomisk status, og dermed at utvalget som danner grunnlaget for denne analysen ligger kunstig høyt for variabler som inntekt og utdanning. Dette berører også i stor grad elever i innvandrergruppene.

## 4.2 Difference in differences

Denne analysen benyttes for å kartlegge elevenes progresjon i skoleprestasjoner. Regresjoner av endringsvariabler gir en indikasjon på bevegelser i faktorer over tid. Metoden *Difference in differences* estimerer forskjeller i endring mellom individ fra ulike andeler av utvalget, en kontroll- og en behandlingsgruppe. I dette tilfellet vil analysen springe ut fra gruppene intakte og ikke-intakte familier, hvor sistnevnte gruppe ”behandles” med familieoppløsning. Strategien er å studere om det finnes en forskjell i utvikling mellom 5. og 8. klassetrinn for elever fra de ulike familiekategoriene; i hvilken grad elevenes progresjon mellom prøvene påvirkes av en endring i familiestruktur. Denne metoden er et godt supplement til MKM-regresjonene, i det analysen kan være en løsning på de økonometriske utfordringene med selvseleksjon og begrenset mulighet for å isolere effekten av endringer i enkeltfaktorer.

Datamaterialet som benyttes til denne delen av analysen er observasjoner av elever på 5. klassetrinn fra 2007, og 8. klassetrinn i 2010. Observasjonene følger da de samme individene over disse årene. Elevene som representerer ikke-intakte familier begrenses til elever som opplever et familiebrudd i tiden mellom 5. og 8. klasse. Alle inkluderte individ lever altså i perioden før 5. klassetrinn i intakte familier. Utvalget deler seg så inn i gruppene intakte og ikke-intakte familier innen elevene når 8. klasse. Ved kun å inkludere i utgangspunktet intakte familier kan man bedre undersøke effekten av en familiestrukturendring. Alle elever registrert med ikke-intakt familie i 8. klasse da vil ha gjennomgått den familieendringen som er fokus i undersøkelsen i løpet av perioden, og elever som gjennom hele perioden tilhører en ikke-intakt familie vil ikke forstyrre den undersøkte effekten. Dette utvalget av datamateriale sørger også for at en registrert endring i familiestruktur over årene kun kan bety at en familieoppløsning har funnet sted. Familier som går fra ikke-intakt til intakt i tidsrommet undersøkt, og dermed også har en endring i *Familiestruktur*, inkluderes ikke i analysen.

Endringen finnes ved å ta differansen mellom produktfunksjonene for poengscore i nasjonale prøver for 5. og 8. trinn, gitt de samme variablene inkludert i begge likningene. Konstante individ- og familiespesifikke variabler, som elevens kjønn, innvandrerbakgrunn og foreldres utdanning (tilnærmet) vil da differensieres bort. Dette gjelder også uobserverbare konstante familiekarakteristika. *Difference in differences*-uttrykket står nå igjen med endringen i familiestruktur mellom klassetrinnene som variabel, sammen med faktorer som også endres over årene. Den estimerte effekten av familiestruktur er et uttrykk for betydningen av en familieoppløsning, uten konstante familiekarakteristika som kan påvirke elevresultatene



uavhengig av familiestruktur, og skaper endogenitetsproblemet med selvseleksjon. Dersom denne metoden lykkes i å differensiere bort familiespesifikke egenskaper, kan man få et mer spesifikt mål på familiestrukturens virkning. Ved deretter suksessivt å legge til variabler man antar endres simultant med en endring i familiestruktur, kan enkelteffekter isoleres.

En familiespesifikk variabel som i mange tilfeller ikke faller bort ved denne metoden er inntekt. Det er for det første naturlig å tenke seg at familiens inntekt varierer over år, selv uten andre endringer i familien. For det andre er inntekt en faktor som vil kunne absorbere deler av effekten av en familieoppløsning. Mange forskere argumenterer for at effekten av en familieoppløsning går via blant annet økonomiske ressurser, i det disse i mange tilfeller reduseres på familienivå når familien splittes opp. Ved å inkludere denne faktoren som endringsvariabel, kan denne påstanden analyseres. Dette er et ledd i å isolere effektene av de mange endringene i familieegenskaper en familie står overfor ved en familieoppløsning, og viser til en mulig løsning på en av hovedutfordringene ved MKM.

Difference in differences-funksjonene formuleres som under. Det forutsettes at restledd og inkluderte variabler ikke korrelerer. Etter differensieringen står likningen igjen med variablene for familiestruktur og noen andre familieegenskaper som også endres over tid. Dersom familien holder sammen gjennom hele perioden vil familiestruktur-dummyen ha verdien 1 for begge klassetrinnene. Variabelen vil i dette tilfellet differensieres bort, og  $\Delta\text{familiestruktur}$  vil bli 0. En familie hvor en oppløsning finner sted i løpet av tidsrommet vil få  $\Delta\text{familiestruktur} = -1$ , i det dummyvariabelen har verdien 1 i første periode, 0 i andre. Endringsvariabelen for familiestruktur som benyttes i selve regresjonen under avsnitt 5.3 antar som over verdien 0 for familier som forblir intakte over perioden, og for enkelthets skyld verdien 1 dersom en familieoppløsning finner sted.

$$\begin{aligned}
 NPLES_{i,8} - NPLES_{i,5} &= \alpha_0 + \alpha_1(\text{familiestruktur}_{i,8} - \text{familiestruktur}_{i,5}) \\
 &+ \alpha_2(\text{farsinntekt}_{i,8} - \text{farsinntekt}_{i,5}) \\
 &+ \alpha_3(\text{morsinntekt}_{i,8} - \text{morsinntekt}_{i,5}) \\
 &+ \alpha_4(\text{farsalder}_{i,8} - \text{farsalder}_{i,5}) \\
 &+ \alpha_5(\text{morsalder}_{i,8} - \text{morsalder}_{i,5}) \\
 &+ \varepsilon_i
 \end{aligned}$$

$$\begin{aligned}
NPREG_{i,8} - NPREG_{i,5} &= \beta_0 + \beta_1(familiestruktur_{i,8} - familiestruktur_{i,5}) \\
&+ \beta_2(farsinntekt_{i,8} - farsinntekt_{i,5}) \\
&+ \beta_3(morsinntekt_{i,8} - morsinntekt_{i,5}) \\
&+ \beta_4(farsalder_{i,8} - farsalder_{i,5}) \\
&+ \beta_5(morsalder_{i,8} - morsalder_{i,5}) \\
&+ v_i
\end{aligned}$$

$$\begin{aligned}
\Delta NPLES_i &= \alpha_0 + \alpha_1 \Delta familiestruktur_i + \alpha_2 \Delta farsinntekt_i + \alpha_3 \Delta morsinntekt_i \\
&+ \alpha_4 \Delta farsalder_i + \alpha_5 \Delta morsalder_i + \varepsilon_i
\end{aligned}$$

$$\begin{aligned}
\Delta NPREG_i &= \beta_0 + \beta_1 \Delta familiestruktur_i + \beta_2 \Delta farsinntekt_i + \beta_3 \Delta morsinntekt_i \\
&+ \beta_4 \Delta farsalder_i + \beta_5 \Delta morsalder_i + v_i
\end{aligned}$$

En difference in differences-analyse kan bidra til å finne løsninger på de økonometriske utfordringene diskutert under estimeringsstrategien MKM. Kilden til skjevheter i MKM-estimatene grunnet utelatte variabler som korrelerer med restleddet elimineres når konstante variabler differensieres bort. Videre kan enkelteffekter av faktorer som endres simultant med en familieoppløsning isoleres ved å inkludere endringen i andre variabler til regresjonen. Dette er de store styrkene ved å bruke denne metoden i analysen. Det siste poenget er imidlertid ikke ukomplisert for dette utvalget.

Et problem med gjennomførelsen av analysen som skissert over er at datamaterialet for hver observasjon inkluderer informasjon om foreldrenes separate inntekter, og ikke et samlet tall for familiens økonomiske midler. Det kan slik bli vanskelig å finne den korrekte familieinntekten etter familieoppløsningen. I tillegg finnes et mer fundamentalt problem i datasettet. Foreldres inntekt er angitt på bakgrunn av tall fra 2010, og ikke basert på informasjon fra de enkelte årene. Det betyr at inntekten oppgitt for hvert individ er stabil og lik inntekten fra 2010 gjennom perioden, og at denne variabelen dermed vil differensieres bort med difference in differences-metoden. Det er slik ikke mulig å inkludere en variabel som fanger opp endringen i inntekt simultant med familiestrukturendringen.

Opgaven benytter seg av informasjon om familietype etter familieoppløsningen for å lage et alternativt mål på familieinntekt etter familiestruktureringen i en utvidelse av difference in differences-analyse. Hvem eleven bor hos blir et anslag for økonomiske midler tilgjengelig. Denne løsningen har imidlertid også noen svakheter. Det er for det første problematisk å stadfeste sikkert hvem av foreldrene elevene mottar økonomiske ressurser fra. I noen tilfeller bidrar begge de biologiske foreldrene, i andre faller det økonomiske ansvaret i større grad på forelderen som har barnet boende hos seg. For det andre er ikke informasjonen om bosituasjon en refleksjon av inntekt alene. Et estimat på grunnlag av dette sier noe om den samlede effekten av å bo hos mor eller far, ikke direkte betydningen av deres inntekt.

En annen svakhet ved å bruke difference in differences i denne analysen er at vi har mangelfull informasjon om gruppene i utvalget i tiden før familieoppløsningen. Metoden krever at gruppene som undersøkes er like i forkant av behandlingstidspunktet. I dette utvalget lever alle individene i en intakt familie på 5. klassetrinn, og familiestrukturen er slik lik. Man kan imidlertid ikke være sikker på at de samlede forutsetningene er de samme på tvers av gruppene. For eksempel er det grunn til å tro at elevgruppene ikke er like i forkant av familiestruktureringen dersom en selvseleksjon finnes. Kapittel 5.4 presenterer en sammenlikning av gruppene på 5. klassetrinn som en strategi for å kartlegge utgangsforskjellene. Dette gir en indikasjon på om det finnes en forskjell mellom elevene, men viser imidlertid kun situasjonen på et gitt tidspunkt. Difference in differences estimerer påvirkningen av endringen i variabler, og i den forbindelse ville kunnskap om forskjeller i *progresjon* på tvers av gruppene før familiestruktureringen være av større nytte. Dersom difference in differences-estimatet indikerer en forskjell i progresjon mellom elevgruppene, er det vanskelig å vite om hele eller deler av denne forskjellen også var tilstede i utvalget før oppløsningen inntraff.



## 5 Resultater og diskusjon

I dette kapitlet presenteres estimeringsresultatene. Første underkapittel tar for seg utfordringene med selvseleksjon i utvalget. Andre underkapittel viser resultatene fra analysen av sammenhengen mellom elevprestasjoner og familiestruktur. MKM benyttes i regresjonene i begge delkapitlene. Tredje underkapittel viser resultat fra difference in differences-estimatene av familiestrukturens betydning. Videre presenteres alternative analysemetoder som søker å løse noen av utfordringene som står igjen etter de første delkapitlene. Heterogene effekter inkluderes til slutt.

### 5.1 Forholdet mellom familiestruktur og familiespesifikke karakteristika

Dette avsnittet vil adressere problematikken rundt selvseleksjon inn i familiestrukturer og undersøke hvorvidt dette ser ut til å være en utfordring i analysen. En regresjon av familiestruktur på familiespesifikke variabler stadfester forholdet mellom faktorene, og indikerer i hvilken grad det eksisterer en samvariasjon mellom familiens egenskaper og sannsynligheten for at den holdes intakt.

Tabell 5.1 viser resultatene av en regresjon av den binære variabelen *Familiestruktur* på et utvalg familiespesifikke variabler og tidsdummyer for årene 2008-11. 2007 er valgt som basisår. Siden denne variabelen kun kan anta verdiene 0 og 1, blir det problematisk å snakke om marginale effekter av endringer i forklaringsvariablene, slik man gjør for variabler som kan ta en rekke verdier. Effektene tolkes som endringer i sannsynlighet for at familiestrukturdummyen tar verdien 1, at foreldrene lever sammen i ekteskap eller som samboere, ved en enhets endring i regresjonsvariablene. Regresjonsresultatene gir en indikasjon på hvordan elevenes kjønn, foreldres utdanning, inntekt og alder, og etnisk bakgrunn i familien virker inn på sannsynligheten for at familien holdes intakt.

Kort oppsummert finnes en positiv effekt av variablene foreldres inntekt og utdanning, mors alder og innvandrerbakgrunn i familien på sannsynligheten for at familien er intakt. Disse resultatene indikerer at familier hvor foreldrene er gifte eller samboere i gjennomsnitt har høyere inntekt og utdanning hos foreldre, høyere alder for mor og innvandrerbakgrunn.

Tabell 5.1: MKM-regresjon, Familiestruktur-dummy

	Hele utvalget n = 677 119	5. klasstrinn n = 275 353	8. klasstrinn n = 286 170
Konstant	0.0171*** (0.0049)	0.0024 (0.0073)	-0.1150** (0.0079)
Kjønn	- 0.0027* (0.0011)	- 0.0014 (0.0016)	-0.0040* (0.0017)
Fars inntekt	8.14e-08*** (1.19e-09)	7.95e-08*** (1.85e-09)	7.92e-08*** (1.78e-09)
Mors inntekt	4.98e-09 (2.84e-09)	1.48e-08** (4.36e-09)	6.04e-09 (4.41e-09)
Fars utdanning	0.0426*** (0.0003)	0.0408*** (0.0005)	0.0418*** (0.0005)
Mors utdanning	0.0135*** (0.0004)	0.0138*** (0.0006)	0.0103*** (0.0006)
Alder far	- 0.0004** (0.0001)	-0.0004* (0.0002)	0.0003 (0.0002)
Alder mor	0.0107*** (0.0002)	0.0119*** (0.0002)	0.0128*** (0.0002)
Innvandrer, første-gen	0.2088*** (0.0045)	0.2008*** (0.0075)	0.2165*** (0.0068)
Innvandrer, andregen	0.1587*** (0.0028)	0.1333*** (0.004)	0.1667*** (0.0044)
År 2008	- 0.0068*** (0.0018)	- 0.0033 (0.0025)	- 0.0097*** (0.0026)
År 2009	- 0.0114*** (0.0018)	- 0.0039 (0.0025)	- 0.0184*** (0.0026)
År 2010	- 0.0318*** (0.0017)	- 0.0037 (0.0025)	- 0.0202*** (0.0026)
År 2011	- 0.0367*** (0.0017)	- 0.0109*** (0.0025)	- 0.0231*** (0.0026)

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

Det kan være nyttig å endre koeffisienttallene noe for å få et mer lettfattelig bilde av variablenes effekter på sannsynligheten for at familien holdes intakt. Hver variabels estimerte virkning, gitt ved koeffisientene fra regresjonen, multipliseres med deres respektive standardavvik. Dette produktet, rapportert i kolonne 3 i tabell 5.2, gir et bedre mål på effektene. Videre kan produktet divideres med standardavviket til den binære variabelen, *Familiestruktur*, som er lik 0.45. Man oppnår da et uttrykk for endringen i sannsynlighet for at *Familiestruktur*=1 angitt per standardavvik, ved et standardavviks endring i en forklaringsvariabel. Tabellen under oppsummerer disse effektene. Disse produktene er bakgrunnen for diskusjonen og sammenlikning av variablenes virkninger.

Tabell 5.2 rapporterer de standardavvikjusterte estimatene fra regresjonen av familiestruktur på ulike forklaringsvariabler. Tallene er hentet fra regresjonen for hele utvalget. Resultatene diskuteres under.

Tabell 5.2: Effekter justert med standardavvik

Variabel	Koeffisient	Standardavvik	Effekt	Effekt/Standardavvik familiestruktur
Kjønn	- 0.0027	0.4999	-0.0013	- 0.003
Fars inntekt	8.14e-08	470 536	0.0383	0.085
Mors inntekt	4.98e-09	206 985	0.001	0.0023
Fars utdanning	0.0426	1.7358	0.0739	0.1643
Mors utdanning	0.0135	1.6781	0.0227	0.0503
Fars alder	- 0.0004	6.0659	- 0.0024	- 0.0054
Mors alder	0.0107	5.2108	0.0558	0.1239
Innvandrer, førstegen	0.2088	0.119	0.0248	0.0552
Innvandrer, andregen	0.1587	0.1951	0.031	0.0688

### 5.1.1 Diskusjon

Kjønnsvariabelen er en binær variabel som tar verdiene 1 for gutt, 2 for jente. Den har en negativ, men liten koeffisient. Resultatet indikerer at familier med jenter har en mindre sannsynlighet for å holdes intakt enn gutter, effekten er imidlertid svært beskjeden sammenliknet med virkningen av de andre variablene. Dette estimatet har også den svakeste signifikansen og bør tolkes med forsiktighet.

Effektene av foreldres utdanning og inntekt går begge i positiv retning. Virkningen av høyere utdanning og inntekt hos far er dessuten større enn den tilsvarende effekten av mors nivå på disse variablene. Dersom fars utdanning øker med ett standardavvik følger det at det gir en større sannsynlighet for å være i en intakt familie lik 0.1643 standardavvik, tilsvarende effekt av mors utdanning er 0.0503 standardavvik. Foreldres inntekt er en variabel som ofte inkluderes i studier av familiestrukturens rolle i elevers skoleprestasjoner og som det ofte konkluderes med at er av signifikant betydning. Flere forskere har funnet at effekten av familiestruktur på elevresultat reduseres sterkt når inntektsvariabelen inkluderes. Dersom familiens inntekt har en sentral rolle i å bestemme forutsetningene for at familien fungerer over tid og holdes intakt, vil dette komme til syne ved at variabelen har en positiv effekt på familiestruktur. Det aktuelle utvalget kjennetegnes ved denne positive koeffisienten, men effekten er merkbart mindre enn virkningen av utdanningsnivå. Dette argumenterer for at det i dette utvalget er vel så viktig å inkludere foreldres utdanning som inntekt i en analyse av forholdet mellom familiestruktur og elevresultat for å få en så forventningsrett estimator som mulig.

Mors alder har en stor positiv effekt, mens fars alder har en liten negativ virkning. Dersom mors alder øker med et standardavvik, 5.2108 år, vil sannsynligheten for å tilhøre en intakt familie øke med 0.1239 standardavvik. Denne korrelasjonen kan reflektere individene i utvalget med svært unge mødre. Mødre som får barn i tenårene har ikke utenkelig en lavere sannsynlighet for fremdeles å bo under samme tak som far, når barnet når storskoletrinnet, enn kvinner som er mer etablert når de blir mødre. Dersom dette er korrekt kan det forklare hvorfor relativt høy alder hos mor henger sammen med en intakt familiestruktur. Dette er en hypotese som kan etterprøves i en eventuell senere analyse av familieoppløsning. Virkningen av en økning i fars alder er, i likhet med kjønnsvariabelen, av svært liten betydning, og sannsynligheten for at familien er intakt endres med kun  $-0.0054$  standardavvik per standardavviks endring i fars alder.

Som nevnt over har elever med innvandrerbakgrunn en større sannsynlighet for å leve i en intakt familie enn etnisk norske elever. Man kan argumentere for at dette reflekterer kulturforskjellene mellom de etniske gruppene, ettersom en stor del av Norges innvandrere kommer fra muslimske samfunn hvor familie og ekteskap er viktige verdier. Man kan i tillegg tenke seg at denne gruppen elever jevnt over har dårligere forutsetninger for å prestere i skolen enn etnisk norske elever, i det for eksempel språkrelaterte utfordringer kan virke inn på



hvor raskt elevene lærer. Dersom denne antakelsen er korrekt, vil denne faktoren redusere effekten av familiestruktur på elevresultat, ettersom gjennomsnittsprestasjonene for denne gruppen elever reduseres når elever med innvandrerbakgrunn inkluderes. Effekten av familiestruktur på elevprestasjoner vil trolig ha en negativ forventningsskjevhet dersom innvandrerbakgrunn ikke formuleres som egen variabel.

Innledningsvis presenterte oppgaven et mulig endogenitetsproblem med selvseleksjon inn i visse familietyper. Lettere og vanskeligere observerbare egenskaper i familien kan parallelt svekke barns prestasjoner i skolen og forutsetningene for at familien holdes intakt over tid. Dersom dette er tilfellet, vil elever med ufordelaktige forutsetninger i skolen oftere havne i gruppen ikke-intakt familie enn sine klassekamerater. Videre vil individene i denne familiestrukturkategorien i gjennomsnitt inneha visse egenskaper. Dersom dette er tilfellet vil det kunne komme til syne i regresjonen av familiestruktur på familiekarakteristika.

Resultatene over indikerer at det foreligger en selvseleksjon inn i de ulike familiestrukturene. Kort oppsummert, kan man fra funnene over konkludere med positiv virkning av foreldres alder, utdanning og inntekt på sannsynligheten for å være i en intakt familie. Dette betyr at familier av typene 1 og 2, hvor foreldrene er gift eller samboere, i gjennomsnittsnitt har noe høyere alder (for mor), utdanning og inntekt hos foreldre, sammenliknet med familier hvor foreldrene ikke bor sammen. Dersom dette har betydning for barnas prestasjoner i skolen og disse egenskapene ikke inkluderes som egne variabler, vil estimatet av virkningen av familiestruktur være forventningsskjev i positiv retning. Fars utdanning og mors alder er variablene som synes å være av størst betydning for sannsynligheten for at familien er intakt.

## 5.2 Forholdet mellom familiestruktur og elevresultater

Denne seksjonen presenterer analysen av hvordan elev- og familieegenskaper innvirker på prestasjoner i skolen. Resultatene viser elevenes score på nasjonale prøver som funksjon av familiestruktur og noen individ- og familiespesifikke karakteristika. De avhengige variablene er estimert poengsum for prøvene i lesing og matematikk, standardisert til gjennomsnitt 0 og standardavvik 1. Først estimeres prøveresultatene i en enkel regresjon på bakgrunn av én variabel, *Familiestruktur*. Videre bygges modellen ut med den individspesifikke variabelen *Kjønn*, så gradvis med familiespesifikke karakteristika som foreldres inntekt, utdanning og alder, og familiens bakgrunn. Hensikten med trinnvis å bygge ut regresjonen er å ta hensyn til at det kan eksistere en selvseleksjon inn i familietyperne. Dette vil komme til syne i en endring i estimert effekt av familiestruktur under utvidelsen. Dette indikerer en forventningsskjevhet i det første estimatet grunnet utelatte variabler som korrelerer med testscore og *Familiestruktur*.

### 5.2.1 Testscore som funksjon av familiestruktur

Første runde regresjoner er poengscore på nasjonale prøver i lesing og matematikk som funksjon av familiestruktur. Resultatene presenteres i tabell 5.3. Alle funnene argumenterer for at elever fra intakte familier, *Familiestruktur*=1, oppnår høyere poengsum enn elever fra ikke-intakte familier. Denne konklusjonen samsvarer med flere tidligere studier, for eksempel Björklund, Ginther og Sundström (2006), Francesconi, Jenkins og Siedler (2010) og Ginther og Pollak (2004) som også finner en negativ sammenheng mellom oppvekst i oppløst familie og elevprestasjoner i skolen. Fortrinnet av å leve i intakte familier er dessuten større for faget matematikk enn lesing, og større for 8. enn 5. trinn.

Tabell 5.3: Resultat nasjonale prøver som funksjon av familiestruktur

	Lesing		Matematikk	
	5. klassetrinn n = 259 037	8. klassetrinn n = 286 660	5. klassetrinn n = 263 138	8. klassetrinn n = 272 723
Konstant	- 0.1254*** (0.0039)	- 0.1476*** (0.0035)	- 0.1653*** (0.0069)	- 0.1855*** (0.0035)
Familiestruktur	0.1681*** (0.0045)	0.2118*** (0.0042)	0.2217*** (0.0045)	0.2661*** (0.0041)

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

## 5.2.2 Testscore som funksjon av familiestruktur og individspesifikk variabel, kjønn

Tabell 5.4: Resultat nasjonale prøver, inkl. variabel kjønn

	Lesing		Matematikk	
	5. klassetrinn n = 259 037	8. klassetrinn n = 268 660	5. klassetrinn n = 263 138	8. klassetrinn n = 272 723
Konstant	- 0.3897*** (0.0075)	- 0.5305*** (0.0067)	0.0611*** (0.0069)	0.0334*** (0.0066)
Familiestruktur	0.1691*** (0.0045)	0.2134*** (0.0041)	0.221*** (0.0044)	0.2652*** (0.0041)
Kjønn	0.1762*** (0.0039)	0.2559*** (0.0038)	- 0.1512*** (0.0039)	- 0.1465*** (0.0038)

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

Tabell 5.4 inkluderer, i tillegg til familiestruktur, elevenes kjønn i regresjonen. Sammenliknet med resultatene fra tabell 5.3, har dette lite å si for estimert effekt av familiestruktur, estimatoren ligger stabil for begge klassetrinn og fag. Å være i en intakt familie gir fremdeles en høyere estimert testscore enn alternative familiestrukturer. Kjønn har en positiv estimert virkning på resultat i lesing. Dette indikerer at jenter, ceteris paribus, i gjennomsnitt presterer bedre enn gutter. I matematikk har vi motsatt mønster, gutter presterer bedre enn jenter. Dette mønsteret i fagprestasjoner på tvers av kjønn består gjennom hele analysen.

## 5.2.3 Testscore som funksjon av familiestruktur, individ- og familiespesifikke variabler

De neste regresjonene, tabellene 5.5-5.8, inkluderer, i tillegg til familiestruktur og kjønn, suksessivt familiespesifikke variabler i analysen.

Tabell 5.5 presenterer resultatene etter inntekt inkluderes som variabel. Estimert effekt av familiestruktur svekkes merkbart. Dette resultatet er ikke overraskende sett i lys av sammenhengen mellom familiestruktur og foreldreinntekt funnet i avsnittet om selvseleksjon. Fordi inntekt korrelerer med både testscore og familiestruktur vil en utelatelse av denne variabelen fra regresjonen gi en skjevhet i estimatoren for familiestruktur, og estimert virkning av å bo i en intakt familie blir for høy. Denne konklusjonen stemmer godt med

resultatene fra tidligere forskning, for eksempel Ginther og Pollak (2004) og Biblarz og Raftery (1999), som også finner at effekten av familiestruktur svekkes når en variabel for familieinntekt legges til analysen. Trendene fra tabell 5.3 og 5.4 er fremdeles tilstede, effekten av familiestruktur er større for 8. enn for 5. trinn, og større for matematikk enn lesing. Testscore for lesing på 5. klassetrinn påvirkes, i motsetning til funnene over, nå negativt av å være i en intakt familie. Alt annet likt, vil testscoren for en elev i intakt familie ligge omtrent 0.1 standardavvik under en elev fra en ikke-intakt familie. Økning i familiens inntekt estimeres til å gi en økning i elevenes prestasjoner, mors inntekt har noe større virkning enn fars.

Tabell 5.5: Resultat nasjonale prøver, inkl. variabler for foreldres inntekt.

	Lesing		Matematikk	
	5. klassetrinn n = 259 037	8. klassetrinn n = 268 660	5. klassetrinn n = 263 138	8. klassetrinn n = 272 723
Konstant	- 0.6534*** (0.0075)	- 0.8438*** (0.007)	- 0.2351*** (0.0074)	- 0.2022*** (0.007)
Familiestruktur	- 0.1118** (0.0045)	0.1558*** (0.0041)	0.1564*** (0.0044)	0.2059*** (0.004)
Kjønn	0.1783*** (0.0038)	0.2562*** (0.0037)	- 0.1492*** (0.0038)	- 0.1465*** (0.0037)
Fars inntekt	1.79e-07*** (4.24e-09)	2.02e-07*** (3.89e-09)	2.10e-07*** (4.21e-09)	2.09e-07*** (3.82e-09)
Mors inntekt	7.08e-07*** (9.53e-09)	8.17e-07*** (9.04e-09)	7.84e-07*** (9.44e-09)	8.52e-07*** (8.99e-09)

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

Tabell 5.6: Resultat nasjonale prøver, inkl. variabler for foreldres utdanning.

	Lesing		Matematikk	
	5. klassetrinn n = 259 037	8. klassetrinn n = 268 660	5. klassetrinn n = 263 138	8. klassetrinn n = 272 723
Konstant	- 1.2707*** (0.0085)	- 1.4812*** (0.008)	- 0.7802*** (0.0085)	- 0.882*** (0.008)
Familiestruktur	0.0125** (0.0044)	0.0633*** (0.004)	0.0672*** (0.0044)	0.1181*** (0.004)
Kjønn	0.1809*** (0.0037)	0.2589*** (0.0036)	- 0.1475*** (0.0037)	- 0.1441*** (0.0036)
Fars inntekt	5.33e-08*** (4.25e-09)	7.39e-08*** (3.88e-09)	9.53e-08*** (4.25e-09)	8.98e-08*** (3.83e-09)
Mors inntekt	2.14e-07*** (9.99e-09)	2.99e-07*** (9.44e-09)	3.50e-07*** (9.97e-09)	3.78e-07*** (9.45e-09)
Fars utdanning	0.0852*** (0.0012)	0.0917*** (0.0012)	0.08*** (0.0012)	0.0904*** (0.0012)
Mors utdanning	0.1163*** (0.0013)	0.1219*** (0.0012)	0.0988*** (0.0013)	0.1074*** (0.0012)

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

Når foreldres utdanning inkluderes som egen variabel forsvinner store deler av det estimerte fortrinnet med å leve i en intakt familie. Effekten på estimert resultat i de andre nasjonale prøvene mer enn halveres sammenliknet med funn fra forrige regresjon, uten utdanning inkludert som forklaringsvariabel. Som argumentert for under regresjonen hvor inntektsnivå hentes inn, er dette ingen overraskelse gitt kunnskap om selvseleksjon i utvalget basert på blant annet disse variablene. Grunnet korrelasjonen mellom utdanning, familiestruktur og testscore vil en ekskludering av utdanning fra regresjonen gi en skjevhet i estimatoren for familiestrukturen. Dette indikerer at estimert effekt av *Familiestruktur* var for høy i de første regresjonene.

Elever med høyt utdannede foreldre oppnår i gjennomsnitt høyere testresultater enn elever med lavere utdannede foreldre. Utdanningsnivå hos mor har dessuten større innvirkning enn

fars. For øvrig er effekten av foreldres utdanning på elevresultat sterkere enn effekten av familiestruktur.

Betydningen av foreldres inntekt svekkes dessuten når utdanningsnivå inkluderes. Dette peker mot en positiv sammenheng mellom inntekt og utdanning. Ved å inkludere variabler for utdanning, får vi nå et mer forventningsrett uttrykk for inntektseffekten. Mors inntekt har fremdeles en større effekt enn fars.

Tabell 5.7: Resultat nasjonale prøver, inkl. variabler for foreldres alder.

	Lesing		Matematikk	
	5. klassetrinn n = 259 037	8. klassetrinn n = 268 660	5. klassetrinn n = 263 138	8. klassetrinn n = 272 723
Konstant	- 1.6183*** (0.0168)	- 1.9387*** (0.0168)	- 0.8486*** (0.0167)	- 1.0940*** (0.0169)
Familiestruktur	- 0.0021 (0.0044)	0.0451*** (0.004)	0.064*** (0.0044)	0.1100*** (0.004)
Kjønn	0.1806*** (0.0037)	0.2582*** (0.0036)	- 0.1475*** (0.0037)	- 0.1444*** (0.0036)
Fars inntekt	4.91e-08*** (4.25e-09)	7.15e-08*** (3.88e-09)	9.42e-08*** (4.26e-09)	8.93e-08*** (3.83e-09)
Mors inntekt	1.68e-07*** (1.00e-08)	2.73e-07*** (9.46e-09)	3.44e-07*** (1.0e-08)	3.67e-07*** (9.48e-09)
Fars utdanning	0.0828*** (0.0013)	0.0886*** (0.0012)	0.0794*** (0.0012)	0.0893*** (0.0012)
Mors utdanning	0.1149*** (0.0013)	0.12*** (0.0012)	0.0985*** (0.0013)	0.1065*** (0.0012)
Alder far	0.0004 (0.0004)	0.0006 (0.0004)	- 0.0006 (0.0004)	0.0015** (0.0004)
Alder mor	0.0093*** (0.0005)	0.0113*** (0.0005)	0.0026*** (0.0005)	0.0039*** (0.0005)

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

Tabell 5.7 inkluderer foreldres alder som forklaringsvariabel. Resultatene viser at fars alder har svært liten effekt på testscorene i lesing og matematikk, og estimatoren er i noen tilfeller heller ikke signifikant ved et 5%-nivå. Mors alder har en positiv, men relativt beskjeden effekt. Et standardavviks økning i alder gir en økning i testscore på kun 0.03-0.1 standardavvik. Betydningen av familiestruktur reduseres noe når aldersvariablene innføres, i størst grad for lesing. Endringen er imidlertid ikke på langt nær så tydelig som effekten av å inkludere utdanning i regresjonen. Utelatelse av alder har ikke skapt like store skjevheter i estimatoren for *Familiestruktur*. Betydningen av foreldres inntekt reduseres noe når alder inkluderes. Virkningen er også her begrenset, men resultatet kan være en indikasjon på en sammenheng mellom alder og nivå på inntekt. Dette er ikke en overraskende observasjon, ettersom man kan tenke seg at lønnsnivå og ansiennitet ofte henger sammen.

Tabell 5.8 henter inn informasjon om innvandrerbakgrunn hos elever. Første- og andre-generasjonsinnvandreres estimerte testscore ligger mellom 0.2-0.5 og 0.03-0.24 standardavvik under etnisk norske elevers. Forskjellene mellom norske elever og elever med innvandrerbakgrunn er større i lesing enn matematikk, og jevnt over større på 8. enn på 5. klassetrinn.

Effekten av familiestruktur øker når variablene for innvandrerbakgrunn inkluderes. Sett i lys av innvandrerbakgrunnens negative effekt på elevprestasjoner oppsummert over og den positive korrelasjonen mellom sannsynligheten for å leve i en intakt familie og innvandrervariablene, er dette resultatet som forventet. Før familiens bakgrunn ble inkludert som egen variabel, lå de ufordelaktige forutsetningene for skoleprestasjoner hos innvandrerelevne inn under familiestrukturens estimerte effekt, og estimatoren fikk en negativ forventningsskjevhet. Ved å legge til innvandrervariablene reduseres de i gjennomsnitt dårligere forutsetningene for læring og prestasjoner for gruppen intakt familie, og effekten av å tilhøre denne familiestrukturen øker og blir mindre skjev.

Tabell 5.8: Resultat nasjonale prøver, inkl. variabler for innvandrerbakgrunn.

	Lesing		Matematikk	
	5. klasstrinn n = 259 037	8. klasstrinn n = 268 660	5. klasstrinn n = 263 138	8. klasstrinn n = 272 723
Konstant	- 1.5926*** (0.0168)	- 1.8912*** (0.0168)	- 0.8321*** (0.0168)	- 1.071*** (0.0169)
Familiestruktur	0.0066 (0.0044)	0.0591*** (0.004)	0.0686*** (0.0044)	0.1153*** (0.004)
Kjønn	0.1805*** (0.0037)	0.2581*** (0.0036)	- 0.1475*** (0.0037)	- 0.1444*** (0.0036)
Fars inntekt	4.10e-08*** (4.25e-09)	5.94e-08*** (3.88e-09)	8.97e-08*** (4.27e-09)	8.46e-08*** (3.84e-09)
Mors inntekt	1.68e-07*** (1.01e-08)	2.46e-07*** (9.45e-09)	3.34e-07*** (1.01e-08)	3.57e-07*** (9.50e-09)
Fars utdanning	0.0828*** (0.0013)	0.0888*** (0.0012)	0.0795*** (0.0012)	0.0896*** (0.0012)
Mors utdanning	0.1111*** (0.0013)	0.1153*** (0.0012)	0.0965*** (0.0013)	0.1049*** (0.0012)
Alder far	0.0016*** (0.0004)	0.002*** (0.0004)	- 0.0001 (0.0004)	0.0019*** (0.0004)
Alder mor	0.0082*** (0.0005)	0.0096*** (0.0005)	0.002*** (0.0005)	0.0032*** (0.0005)
Innvandrer, førstegegn	- 0.2996*** (0.018)	- 0.4702*** (0.0147)	- 0.2434*** (0.0178)	- 0.298*** (0.0147)
Innvandrer, andregne	- 0.1938*** (0.0095)	- 0.2445*** (0.0095)	- 0.0668*** (0.0094)	- 0.0256** (0.0096)

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



Betydningen av de andre forklaringsvariablene er relativt uendret etter at modellen bygges ut med variablene for innvandrerbakgrunn. Dette argumenterer for lav korrelasjon mellom etnisk bakgrunn og karakteristika som utdanning, inntekt og alder. Det er gode grunner til å sette spørsmålstegn ved dette, ettersom man kan tenke seg at foreldregenerasjonen hos første- og andregenerasjonsinnvandrere i gjennomsnitt har lavere inntekt og utdanning enn gjennomsnittsforelderen til etnisk norske elever. Mange innvandrere har levd deler av livet på flukt eller i land hvor gjennomsnittlig utdanningsnivå ligger under Norges. Datamaterialet denne oppgaven tar utgangspunkt i har gjennomgått en seleksjon, i det eksempelvis individ med foreldre med ukjent utdanningsnivå og med negativ inntekt er eliminert fra utvalget. Denne tilpasningen berører i stor grad observasjoner for innvandrere. Antallet første- og andregenerasjonsinnvandrere i datamaterialet inkludert i analysen er henholdsvis 9 732 og 26 826. Utvalget favner imidlertid hele 22 830 og 32 742 observasjoner av elever med innvandrerbakgrunn når observasjoner med ukjent utdanningsnivå ikke elimineres. Dette peker i retning av at særlig gruppen førstegenerasjonsinnvandrere kan stå i fare for å ikke ha et representativt utvalg med i studien. Dette kan være en forklaring på hvorfor sammenhengen mellom innvandrerbakgrunn og eksempelvis utdanning og inntekt er svakere enn forventet.

Regresjonen er også utført med års-dummyer inkludert. Effekten av familiestruktur endret seg imidlertid svært lite med dette. Disse resultatene vil derfor ikke bli diskutert nærmere.

#### 5.2.4 Diskusjon

Analysene av hvordan elevenes prestasjoner påvirkes av ulike faktorer har gitt resultater som samsvarer med tidligere forskning. Å leve i en ikke-intakt familie ser i utgangspunktet ut til å ha en negativ virkning på elevenes resultater i skolen. Når studien bygges ut med utvalgte individ- og familiespesifikke variabler, endres imidlertid bildet noe. Effekten av familiestruktur blir svakere, spesielt når foreldres inntekt og utdanning inkluderes. Dette peker i retning av en selvseleksjon inn i familiestrukturene. Observerbare og uobserverbare familieegenskaper korrelerer med både familiestruktur og elevenes skoleprestasjoner, og estimatet fra den enkle regresjonen får en forventningsskjevhet ettersom disse variablene er utelatt. At virkningen av familiestruktur reduseres når denne informasjonen legges til bekrefter nettopp dette. Oppvekst i en ikke-intakt familie gir ikke så ufordelaktige forutsetninger som den første regresjonen argumenterer for, en del av de estimerte fortrinnene i intakte familier kan knyttes til at disse i gjennomsnitt har foreldre med høyere utdanning og høyere inntekt. Selv om effekten av familiestruktur svekkes med utvidelsen av modellen

eksisterer fremdeles en sammenheng mellom variabelen og elevenes resultater. Denne er imidlertid relativt liten, mellom 0.007 og 0.12 standardavvik høyere estimert score for elever i intakte familier. Spørsmålet er om relasjonen kan kalles et kausalt forhold eller om korrelasjonen gjenspeiler virkningen av andre faktorer som det ikke kontrolleres for. At sammenhengen synes å være svak er dessuten et argument som peker i retning av at en kausalitet er mindre sannsynlig.

Etter disse analysene er det fremdeles uklarhet rundt noen av de økonometriske utfordringene presentert tidligere i oppgaven. Ett problem er at vi kan ha at familier som selekterer seg inn i gruppen ikke-intakte familier har uobserverbare karakteristika som skiller dem fra intakte familier. Det kan justeres for noen typer familieegenskaper i en studie, slik som over hvor familiespesifikke variabler legges til regresjonen. Uobserverbare karakteristika er imidlertid vanskeligere å legge til regresjonen slik at de kan inkluderes som egne variabler. Virkningen av disse absorberes dermed under *Familiestruktur*, og estimatoren har fremdeles en skjevhet. Dette problemet diskuteres videre i avsnitt 5.3 *Estimeringsresultater difference in differences*. I tillegg problematiseres analysen av at flere variabler endres simultant med en endring i familiestruktur. Et relevant eksempel er familiens inntekt. For å kartlegge hvilke av endringene i familiens hverdag som er av størst betydning må enkelteffektene isoleres. Resten av kapitlet undersøker relasjonen mellom elevprestasjoner og familiestruktur ytterligere og presenterer forslag til løsninger på utfordringene nevnt over.

### 5.3 Estimeringsresultater difference in differences

Dette avsnittet presenterer estimeringsresultatene fra difference in differences-analysen. Metoden undersøker hvorvidt det finnes en forskjell i utvikling mellom 5. og 8. klassetrinn for elever fra intakte familier og familier som opplever en oppløsning i løpet av årene mellom prøvene.  $\Delta$ familiestruktur er en endringsvariabel som antar verdien 1 dersom eleven gjennomgår en familieoppløsning mellom 5. og 8. klassetrinn, 0 dersom elevens familie forblir intakt i denne perioden.

#### 5.3.1 Utvikling i prestasjoner som funksjon av familiestruktur

Tabell 5.9: Utvikling i prestasjoner som funksjon av familiestruktur

	Lesing	Matematikk
	Utvikling 5.-8.trinn n = 37 570	Utvikling 5.-8.trinn n = 38 854
Konstant	0.0283*** (0.004)	0.0279*** (0.0035)
$\Delta$ familiestruktur	- 0.0557*** (0.0153)	- 0.0851*** (0.0134)

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

Resultatene fra tabell 5.9 viser en negativ effekt på endringen i testscore på nasjonale prøver for elever som opplever en familieoppløsning i tidsperioden mellom 5. og 8. klasse. De negative koeffisientene for familiestrukturendringen indikerer at elever som gjennomgår en familieoppløsning i gjennomsnitt har en noe svakere utvikling enn sine klassekamerater i tiden mellom 5. og 8. klasse. Konstantleddet er positivt, som betyr at elever som ikke gjennomgår en endring i familiestruktur i løpet av perioden som undersøkes har en progresjon for testresultatene som ligger over gjennomsnittet. Denne enkle regresjonen argumenterer for at et familiebrudd har en negativ virkning på elevenes prestasjoner. Det er imidlertid vanskelig å stadfeste hvor stor selve familieoppløsningens effekt er på utviklingen. Som argumentert for under kapitlet *Estimeringsstrategi* har ikke analysen nok informasjon til å si noe om hvorvidt det eksisterte forskjeller i progresjon på tvers av gruppene før familiebruddet. Analysen kan bekrefte en forskjell i etterkant, men om denne var tilstede også før 5. klassetrinn er usikkert.

Som i analysen basert på MKM kan familiespesifikke variabler legges til regresjonen ettersom analysen utvides. Karakteristika som er konstante over tidsperioden, som kjønn, innvandrerbakgrunn og foreldres utdanning, differensieres bort under regresjonen som måler endringen i progresjonen i elevenes resultater. Enkeltfaktorer som endres simultant med en familieoppløsning, for eksempel inntekt, beholdes, og kan isoleres fra andre forhold og gjøres målbare med denne metoden. Som diskutert i kapitlet om estimeringsstrategi er det problematisk å lykkes med dette for det aktuelle utvalget, i det datamaterialet har begrenset med informasjon for inntektsvariablene.

På bakgrunn av den deskriptive statistikken i tabell 3.2 kan det argumenteres for hvordan en familieoppløsning kan prege hverdagen gjennom endringer i familieinntekt. Tabell 3.2 viser hvordan fordelingen av elever over ulike familietyper endres over årene mellom de aktuelle klassetrinnene. Andelen familier med gifte og samboende familier reduseres med 1.2% og 3.7%. Andelen elever i ikke-intakte familier øker tilsvarende, men det er en tydelig forskjell i hvor mye av økningen familietyperne absorberer. Familier med alenemor eller mor og stefar øker med 1.2% og 2.3%, mens tilsvarende familiestrukturer med far øker med kun 0.8% og 0.6%. Dette indikerer at de fleste av barna som opplever at familien splittes opp bor med mor etter oppløsningen. Gitt kunnskap om at mors inntekt i dette utvalget i gjennomsnitt ligger 200 000 under fars, vil en stor del av disse elevene antakelig oppleve en reduksjon i familieinntekt etter familiebruddet. Regresjonene av resultater ved nasjonale prøver på ulike forklaringsvariabler viser et positivt forhold mellom foreldreinntekt og poengscore. Av dette kan det argumenteres for at fall i inntekt som følge av familieoppløsning, kan virke i negativ retning på elevenes prestasjoner.

Tabell 5.10 presenterer en utvidelse av difference in differences-analysen over, hvor kunnskap om ny familiestruktur etter et familiebrudd utnyttes. Hos hvem av foreldrene barna bor etter familieoppløsningen kan gi en indikasjon på ressurser tilgjengelig, og videre i elevenes progresjon i prestasjoner i årene etter familieoppløsningen. Det er imidlertid viktig å påpeke at dette ikke er en fullverdig erstatning for resultatene man kunne fått ved å inkludere en variabel for familieinntekt i difference in differences-analysen.

Regresjonen tar nå utgangspunkt i to endringsvariabler for familiestruktur.  $\Delta$ familiestruktur, mor fanger opp effekten av et familiebrudd på elevenes utvikling i tilfellene hvor eleven etter oppløsningen bor hos mor. Tilsvarende er  $\Delta$ familiestruktur, far et mål på virkningen av

familieoppløsningen for gruppen av elever som i etterkant bor hos far. Konstantleddet påvirkes ikke av denne endringen, ettersom utvidelsen skjer i formuleringen av variablene for elevene som opplever oppløsning. Estimert utvikling for elevene som gjennom hele perioden lever i en intakt familie påvirkes ikke.

Oppløsningen har som over en negativ virkning på progresjonen i skoleresultater. Effekten av en endring i familiestruktur er størst hos elevene som bor hos mor i etterkant av oppløsningen. Denne effekten er dessuten større for denne elevgruppen, enn for gjennomsnittlig effekt for *alle* elever som gjennomgår et familiebrudd. Effekten for elevene som blir boende med far er følgelig mindre negativ enn både for den samlede gruppen og blant elever som bor sammen med mor. Dette kan gjenspeile den overnevnte gjennomsnittlige forskjellen i inntekt hos mødre og fedre. Analysene i kapittel 5.2 viser en positiv sammenheng mellom foreldres inntekt og elevers prestasjoner. Det kan tenkes at denne relasjonen nå inkorporeres i variablene for oppløsning i det informasjonen om hvem av foreldrene barnet bor hos indirekte sier noe om økonomiske ressurser. Som nevnt under *Estimeringsstrategi* er det risikabelt å være for selvsikker i å tolke dette som en refleksjon av inntektseffekten, siden familietyperen også kan fange opp andre effekter.

Tabell 5.10: Utvikling i prestasjoner som funksjon av familiestruktur

	Lesing	Matematikk
	Utvikling 5.-8.trinn n = 37 570	Utvikling 5.-8.trinn n = 38 854
Konstant	0.0283*** (0.004)	0.0279*** (0.0035)
Δfamilie- struktur, mor	- 0.0734*** (0.0182)	- 0.0875*** (0.0159)
Δfamilie- struktur, far	- 0.0158 (0.027)	- 0.0789** (0.0236)

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

## 5.4 Effekten av familiestrukturændring i forkant av familieoppløsningen

Avsnittene om selvseleksjon argumenterer for at fordelingen av utvalget over intakte og ikke-intakte familier ikke skjer tilfeldig. Familier med visse egenskaper opplever oftere en oppløsning enn familier med andre karakteristika, noe som indikerer at sannsynligheten for at familien splittes opp avhenger av faktorer innad i familien. Barn som opplever en familieoppløsning mellom 5. og 8. klassetrinn vil kunne påvirkes av disse ufordelaktige familieforholdene allerede før selve familieoppløsningen, siden familieegenskapene former oppvekstvilkårene også mens familien er intakt. En hypotese som følger av dette resonnementet er at elever fra familier som splittes opp i tidsrommet mellom 5. og 8. trinn, i gjennomsnitt vil ha dårligere resultat i skolen også i forkant av familieoppløsningen.

Difference in differences-metoden benyttet over krever at individgruppene i analysen står overfor det samme utgangspunktet før familieoppløsningen inntreffer. Dette kravet er vanskelig forenelig med teorien om at barn fra senere oppløste familier påvirkes av uobserverbare familieegenskaper også før familiebruddet. Ved å sammenlikne prestasjonene for elever fra ikke-intakte familier før og etter familieoppløsningen, med tilsvarende resultater for elever som gjennom hele skoleløpet lever i en intakt familie, kan man se om elevene som havner i gruppen ikke-intakt familie i utgangspunktet har ufordelaktige forutsetninger for å prestere i skolen. Dette kan gi en indikasjon på om forutsetningene i difference in differences er tilfredsstillende, men som diskutert under *Estimeringsstrategi* gir det imidlertid kun et bilde av forskjellene på et gitt tidspunkt, og ikke en forskjell i *progresjon*.

Tidligere forskning har undersøkt sammenhengen mellom familiestruktur og barnas prestasjoner i skolen ved å sammenlikne resultat hos elever fra intakte og ikke-intakte familier i tiden før familieoppløsningen fant sted. Cherlin et al. (1991) fant at barn som opplevde et familiebrudd i løpet av barneskoleårene oppnådde dårligere resultater enn barn fra intakte familier, også i forkant av familiebruddet. Regresjonen under tar utgangspunkt i denne ideen, og presenterer en analyse for dette utvalget.

Tabell 5.11 rapporterer gjennomsnittlig testscore for elever som senere opplever av familien går i oppløsning og elever fra familier som holdes intakte. Poengscorene er standardisert med gjennomsnitt 0 og standardavvik 1 for hele utvalget. Som før tar analysen utgangspunkt i 5. klasseelevene fra 2007, hvor familien er intakt.

Tabell 5.11: Gjennomsnittsscore nasjonale prøver på 5. trinn 2007

Utvalg 5. trinn	n	Lesing		n	Matematikk	
		Gjennomsnitt	Standardavvik		Gjennomsnitt	Standardavvik
Elever fra intakte familier	36 868	0.0534	0.9909	36 910	0.071	0.9848
Elever som senere opplever familieoppløsning	2 784	- 0.0352	1.0176	2 782	- 0.624	1.0006

Gjennomsnittsscorene for andelene av utvalget viser en forskjell i prestasjoner allerede før endringen i familiestruktur har funnet sted mellom elever som vokser opp i en intakt familie og elever som gjennomgår en familieoppløsning. Elevene som i løpet av tiden mellom 5. og 8. klassetrinn opplever at familien splittes opp presterer 0.6 standardavvik under gjennomsnittet i matematikk og 0.04 standardavvik under gjennomsnittet i lesing. Elevene fra familier som forblir intakte i det aktuelle tidsrommet har en poengscore marginalt over gjennomsnittet. Dette indikerer at det initialt finnes en forskjell i familieforholdene som er av betydning for elevenes prestasjoner, og med det at familiens egenskaper har noe å si for barnas skolerresultater uavhengig av familiestruktur. Ved å legge til en variabel til en MKM-regresjon som angir om eleven kommer til å oppleve en familieoppløsning de neste årene, kan størrelsen på denne effekten sammenliknes med effekten av andre inkluderte kontrollvariabler. Det blir også mulig å kontrollere for observerbare familieegenskaper som kan ha en innvirkning. Uobserverbare familiekarakteristika er vanskelig å inkludere i en slik empirisk analyse, og kan følgelig ikke kontrolleres for.

Regresjonen baseres på minste kvadraters metode. Utvalget begrenses til elever med utgangsstruktur intakt familie på 5. klassetrinn i 2007 som også har ønsket informasjon inneliggende for 8. klassetrinn i 2010. En binær variabel, *Oppløsning*, tar verdien 1 dersom eleven opplever oppløsning i tiden mellom de aktuelle klassetrinnene, 0 for elever som lever i en intakt familie hele perioden. Estimatoren viser effekten av en fremtidig familieoppløsning på resultater ved nasjonale prøver i lesing og matematikk på 5. trinn. Ved å sammenlikne

denne med de andre estimatorene, kan man se størrelsen av familiestrukturendringens betydning relativt til andre familiespesifikke variabler.

Tabell 5.12: Resultat nasjonale prøver 5. trinn, inkl. dummy for senere familieoppløsning

5. klasstrinn 2007		
Initial familiestruktur: Intakt familie		
	Lesing n = 39 652	Matematikk n = 39 692
Konstant	- 1.5864*** (0.0445)	- 0.9517*** (0.0444)
Oppløsning	- 0.0025 (0.0184)	- 0.0495* (0.0183)
Kjønn	0.2072*** (0.0093)	- 0.0672*** (0.0093)
Fars inntekt	3.45e-08*** (8.99e-09)	7.21e-08*** (9.00e-09)
Mors inntekt	1.90e-07*** (2.50e-08)	3.07e-07*** (2.50e-08)
Fars utdanning	0.0994*** (0.0033)	0.096*** (0.0033)
Mors utdanning	0.1084*** (0.0034)	0.1034*** (0.0034)
Alder far	0.0012 (0.0012)	0.0004 (0.0012)
Alder mor	0.0068*** (0.0014)	0.0019 (0.0014)
Innvandrer, førstegegn	- 0.3682*** (0.0396)	- 0.4097*** (0.0392)
Innvandrer, andregen	- 0.1901*** (0.024)	- 0.0664* (0.0239)

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



Tabell 5.12 oppsummerer resultatene av analysen. En oppløsning i løpet av årene mellom 5. og 8. klassetrinn har en relativt liten effekt, estimert testscore reduseres med 0.003-0.05 standardavvik ved en fremtidig oppløsning. Oppløsningsvariabelen virker negativt inn på elevenes resultater, i størst grad for fagområdet matematikk. Betydningen av en fremtidig familieoppløsning er beskjeden sammenliknet med de andre familiespesifikke variablene, sett bort fra foreldrenes alder. For lesing er estimatoren heller ikke signifikant ved et 5%-signifikansnivå. At betydningen av en fremtidig oppløsning viser seg å være liten når observerbare familiespesifikke variabler inkluderes, kan være et argument for at forskjellen i prestasjoner mellom elevgruppene som først ble funnet skyldes familiens karakteristika og dermed reflekterer selvseleksjonen inn i familietyper. Når det kontrolleres for blant annet foreldres inntekt, utdanning og alder forsvinner mye av effekten av en fremtidig oppløsning. En viss effekt eksisterer likevel fremdeles. Denne kan gjenspeile uobserverbare faktorer det ikke kontrolleres for.

At utvalget i denne delen av analysen er snevret inn kan være problematisk. Potensielt viktig informasjon undertrykkes, i det informasjon om hvilke individ som har gjennomgått en familieoppløsning hentes inn på kun ett tidspunkt, i 8. klasse mot slutten av perioden. Når eleven opplever at familien splittes opp er relevant for tolkningen av oppløsningens effekt. Er familieoppløsningen nært forestående når observasjonen hentes inn, vil hendelsen sannsynligvis ha større påvirkning på elevens prestasjonsevne på observasjonstidspunktet enn en familieoppløsning som forekommer 2-3 år etter. Prestasjonene registrert i 5. klasse vil sannsynligvis i større grad påvirkes av et familiebrudd i løpet av 5. trinn enn tilsvarende familiestrukturering på 8. trinn. Det kan godt tenkes at familien preges av andre egenskaper og har en annen atmosfære enn normalt i tiden like før en oppløsning. Senere og mer omfattende studier av dette emnet bør utnytte informasjon om oppløsningstidspunkt for å oppnå en bedre forståelse av sammenhengen.

## 5.5 Forskjeller blant familietyper i kategorien intakt familie

Analysen har så langt dreid seg om forskjeller i elevresultater mellom to hovedkategorier innen familiestrukturer, intakte og ikke-intakte familier. Dette er en svært grov gruppering som overser nyanser i utvalget. Man kan tenke seg at det for eksempel finnes en forskjell i effekten av en familieoppløsning for elever fra hjem med en utgangsstruktur med gifte foreldre, sammenliknet med hjem der foreldrene var samboere. Ved å gjennomføre separate difference in differences-analyser for de to overnevnte familietyper finnes effektene av et familiebrudd på elevenes progresjon for de to gruppene. Disse sammenliknes så for å undersøke om det eksisterer en forskjell i påvirkning. Regresjonene følger samme mønster som over. Utvalget omfatter 5. klassetrinn hvor familiestrukturen er intakt. Først estimeres effekten av en endring i familiestruktur på progresjonen i elevresultater,  $\Delta_{\text{familiestruktur}} = 1$ , for den delen av utvalget som i utgangssituasjonen hadde familietype 1, gifte foreldre. Deretter følger en tilsvarende regresjon for elever fra familietype 2 på 5. klassetrinn, hvor foreldrene er samboere. Koeffisientene for  $\Delta_{\text{familiestruktur}}$  reflekterer nå effekten av en familieoppløsning på utviklingen i elevprestasjoner for de to utgangsstrukturene. Poengscoren er fremdeles standardisert til gjennomsnitt 0 og standardavvik 1.

Tabell 5.13 rapporterer funnene fra denne analysen. For begge andelene av utvalget har en familieoppløsning negativ effekt på utviklingen for de to fagområdene, lesing og matematikk. Elevene fra familie med utgangsstruktur med samboende foreldre har en utvikling under gjennomsnittet, uavhengig av en eventuell familieoppløsning. Tilsvarende ligger elever fra hjem med gifte foreldre på 5. klassetrinn like over gjennomsnittet. Den negative effekten av et familiebrudd er dobbelt så sterk hos elever fra familietyper med samboende foreldre, sammenliknet med hjem hvor foreldrene i utgangspunktet er gift. Sammen med det negative estimatet for konstantleddet, viser dette at elever fra hjem med familietype 2 i gjennomsnitt ligger under utviklingsnivået til elever fra familier av typen 1. Elever med i utgangspunktet gifte foreldre trekkes også under gjennomsnittlig utvikling ved en familieoppløsning. I lesing er imidlertid effekten svært liten. Det er i praksis problematisk å konkludere med en svakere progresjon enn gjennomsnittet. I matematikk er effekten noe større. De største effektene av en familieoppløsning er å finne hos elever fra familier med samboende foreldre.

Den negative effekten av familieoppløsningen på utvikling mellom 5. og 8. klassetrinn er større for resultatene i matematikk enn i lesing. Denne trenden samsvarer med funn fra

tidligere regresjoner i analysen, sammenhengen mellom resultater og familiestruktur, utvikling i prestasjoner og familiestrukturendringer, samt initiale forutsetninger i forkant av familieoppløsningen.

*Tabell 5.13: Sammenlikning av utvikling i prestasjoner, familietype 1 og 2*

	Elever fra familietype 1, gifte foreldre		Elever fra familietype 2, samboende foreldre	
	Lesing	Matematikk	Lesing	Matematikk
	Utvikling	Utvikling	Utvikling	Utvikling
	5.-8.trinn	5.-8.trinn	5.-8.trinn	5.-8.trinn
	n = 30 840	n = 31 856	n = 6 730	n = 6 998
Konstant	0.0309*** (0.0044)	0.0367*** (0.0039)	- 0.0673*** (0.0098)	- 0.1130*** (0.0086)
$\Delta$ familie- struktur	- 0.0381* (0.0177)	- 0.0665*** (0.0155)	- 0.1033** (0.0308)	- 0.1303*** (0.0271)
*p<0.05; **p<0.01; ***p<0.001				

Resultatene i dette underkapittelet avdekker at det finnes nyanser i utvalget hovedanalysen ikke fanger opp. Tabellen over indikerer betydelige forskjeller mellom utviklingen hos elever som i utgangspunktet er klassifisert i samme familiestrukturkategori.

## 5.6 Heterogene effekter

### 5.6.1 Effekten av familieoppløsning blant gutter og jenter

Bertrand og Pan (2011) argumenterer for at gutters utvikling i større grad enn jenters avhenger av oppvekstmiljø og familiestruktur, samt endringer i disse variablene. Med utgangspunkt i difference in differences-metoden kan det undersøkes om dette er tilfellet også i det norske utvalget. Regresjoner gjøres for to separate deler av utvalget, gutter og jenter. Variabelen  $\Delta$ familiestruktur angir som før effekten av en familieoppløsning på elevenes progresjon i prestasjoner mellom 5. og 8. klassetrinn. Tabell 5.14 oppsummerer resultatene. Estimaten i denne regresjonen bør tolkes i lys av varierende signifikansnivå for koeffisientene.

Tabell 5.14: Sammenlikning av utvikling i prestasjoner, kjønn

	Gutter		Jenter	
	Lesing	Matematikk	Lesing	Matematikk
	Utvikling	Utvikling	Utvikling	Utvikling
	5.-8.trinn	5.-8.trinn	5.-8.trinn	5.-8.trinn
	n = 19 068	n = 19 831	n = 18 502	n = 19 023
Konstant	0.0106 (0.0058)	0.0589*** (0.0049)	0.0464*** (0.0056)	- 0.0042 (0.005)
$\Delta$ familie- struktur	- 0.0381* (0.0177)	- 0.0797*** (0.0186)	- 0.067** (0.0214)	- 0.0911*** (0.0192)

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

Resultatene argumenterer for at en familieoppløsning har negativ innvirkning på utviklingen innen begge fagområder for både gutter og jenter. Estimert effekt på progresjon i matematikk er mer negativ, enn tilsvarende tall i lesing for begge kjønn. Effekten av en familieoppløsning er noe større for jenter enn for gutter i begge fag, men forskjellen er imidlertid relativt beskjeden. Disse resultatene samsvarer ikke med konklusjonen til Bertrand og Pan (2011).

Et naturlig steg videre i denne analysen er å undersøke om effektene for gutter og jenter er signifikant forskjellige. En mulig analysemetode er å inkludere et interaksjonsledd mellom kjønn og endring i familiestruktur til regresjonen. Tabell 5.15 viser resultatene fra denne analysen. Endringsvariabelen  $\Delta$ familiestruktur og interaksjonsleddet  $Kj\ddot{o}nn \times \Delta$ familiestruktur er ikke signifikante. Dette gir bakgrunn for å konkludere med at det ikke eksisterer noen signifikant forskjell i effekten av en familieoppløsning på tvers av kjønnene. Denne oppgavens analyse kan altså ikke bekrefte at funnene fra Bertrand og Pan (2011) eksisterer for et utvalg fra Norge.

Tabell 5.15: Utvikling i prestasjoner som funksjon av familiestruktur og kjønn

	Lesing	Matematikk
	Utvikling 5.-8.trinn	Utvikling 5.-8.trinn
	n = 37 570	n = 38 854
Konstant	- 0.0252* (0.0127)	0.1216*** (0.0111)
Kjønn	0.0358*** (0.0081)	- 0.0629*** (0.007)
$\Delta$ familiestruktur	- 0.0224 (0.0482)	- 0.0684 (0.042)
$Kj\ddot{o}nn \times \Delta$ familiestruktur	- 0.0223 (0.0306)	- 0.0113 (0.0268)

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

### 5.6.2 Effekten av familieoppløsning blant elever i høyt og lavt utdannede familier

En annen problemstilling det kan være interessant å undersøke er om det eksisterer en forskjell i effekten av en familieoppløsning på skoleprestasjoner for elever med høyt og lavt utdannede foreldre. Utdanningsnivå er en variabel som virker å ha stor betydning for elevers skoleprestasjoner, og det er ikke utenkelig at effekten av en familieoppløsning er ulik på tvers av familier med ulike utdanningsnivå. Med utgangspunkt i kategoriseringen av utdanningsvariabelen i kapittel 3 defineres høy utdanning for denne analysen som nivå  $\geq 6$ ,

fullført lavere universitetsutdanning eller høyere. Denne regresjonen vil undersøke forskjellen i progresjon i elevresultater mellom 5. og 8. klasstrinn for individ med høyt og lavt utdannede foreldre. Som i analysen over benyttes difference in differences-metoden. Familier med utdanningsnivå  $\geq 6$  hos begge foreldrene kategoriseres som ”høyt utdannede”, familier med utdanningsnivå  $\leq 5$  for begge foreldrene som ”lavt utdannede”. Denne inndelingen fører til at familier hvor foreldrene har grad av utdanning på hver side av grensen mellom høy og lav utdanning elimineres fra utvalget. Omtrent 10 000 observasjoner trekkes fra utgangsutvalget for hvert fagområde ved denne strategien.

Resultatene rapporteres i tabell 5.16. Utviklingen i prestasjoner i både lesing og matematikk påvirkes negativt av en familieoppløsning for begge utdanningsnivå. Effekten er som rapportert tidligere større i faget matematikk enn i lesing. Den negative virkningen av en familieoppløsning er dobbel så stor for elever i lavt utdannede familier som i høyt utdannede. Dette resultatet må sies å være i tråd med forventningene, ettersom et høyt utdanningsnivå i familien er assosiert med sterke prestasjoner i begge fag for både 5. og 8. klasstrinn. Det er naturlig å tenke seg at familier med høyt utdannede foreldre er relativt ressurssterke selv om en forelder og/eller en inntektskilde kan falle bort ved en familieoppløsning.

Tabell 5.16: Sammenlikning av utvikling i prestasjoner, foreldres utdanningsnivå

Foreldres utdanningsnivå	Lavt $\leq 5$		Høyt $\geq 6$	
	Lesing	Matematikk	Lesing	Matematikk
	Utvikling 5.-8.trinn n = 17 358	Utvikling 5.-8.trinn n = 17 953	Utvikling 5.-8.trinn n = 9 414	Utvikling 5.-8.trinn n = 9 717
Konstant	0.0011 (0.0061)	- 0.0103 (0.0054)	0.0633*** (0.0077)	0.0857*** (0.0066)
$\Delta$ familie- struktur	- 0.0431* (0.0215)	- 0.0718*** (0.0189)	- 0.0248 (0.0339)	- 0.0447 (0.0291)

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

Som i analysen av effekten for gutter og jenter testes også disse funnenes signifikans. En binær variabel *Høy utdanning*, =1 for utdanningsnivå  $\geq 6$  og =0 ellers, legges til regresjonen sammen med et interaksjonsledd mellom denne og  $\Delta$ *familie-struktur*.

Tabell 5.17: Utvikling i prestasjoner som funksjon av familiestruktur og utdanning

	Lesing	Matematikk
	Utvikling 5.-8.trinn	Utvikling 5.-8.trinn
	n = 26 772	n = 27 670
Konstant	0.0011 (0.006)	- 0.0103 (0.0052)
Høy utdanning	0.0622*** (0.01)	0.0959*** (0.0087)
$\Delta$ familie- struktur	- 0.0431* (0.0211)	- 0.0718*** (0.0184)
Høy utdanning $\times$ $\Delta$ familie- struktur	0.0183 (0.041)	0.0271 (0.0359)

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

Resultatene fra tabell 5.17 sår tvil om det finnes en forskjell i effekten av en familieoppløsning blant utdanningsnivågruppene. Interaksjonsleddene er ikke signifikante ved et 5% nivå. Konstantleddene fra regresjonen er heller ikke signifikante.





## 6 Konklusjoner

I denne oppgaven analyseres forholdet mellom familiestruktur og elevers prestasjoner i skolen, gitt ved resultater på nasjonale prøver i lesing og matematikk for 5. og 8. klassetrinn over årene 2007-11. Problemstillingen spør om det kan eksistere en kausal relasjon mellom disse variablene. Resultatene fra den innledende regresjonen av testscore på variabelen *Familiestruktur* argumenterer for en signifikant negativ korrelasjon mellom testresultat og det å vokse opp i en ikke-intakt familie. Dette samsvarer med tidligere forskning referert til innledningsvis i oppgaven, som Björklund, Ginther og Sundström (2006), Francesconi, Jenkins og Siedler (2010) og Ginther og Pollak (2004). Korrelasjonen svekkes imidlertid i det familiespesifikke variabler som foreldres inntekt, utdanning og alder, samt innvandrerbakgrunn inkluderes. Det er også et resultat som kan finnes igjen i tidligere studier, for eksempel Ginther og Pollak (2004) og Biblarz og Raftery (1999). Det er dermed usikkert hvor sterk korrelasjonen mellom familiestruktur og elevprestasjoner egentlig er, og om denne representerer et kausalt forhold.

Modereringen av konklusjonen etter den første enkle regresjonen kan knyttes opp mot et problem med selvseleksjon i utvalget, som gir estimatene en forventningsskjevhet grunnet utelatte variabler. Mange av de estimerte fortrinnene med å leve i en intakt familie, kan spores til at disse hjemmene i gjennomsnitt har relativt ressurssterke foreldre. Den negative sammenhengen funnet mellom familiestrukturen ikke-intakt familie og testscore gjenspeiler dermed at familiene til elever i denne familiekategorien i stor grad har egenskaper som gir elevene ufordelaktige forutsetninger i skolen. Det er slik ikke selve familiestrukturen som har størst innvirkning på elevenes prestasjoner, men visse familieegenskaper som forekommer oftere innen visse grupper i utvalget. Nivå på foreldres utdanning viser seg å være en spesielt avgjørende variabel for barnas skoleprestasjoner.

Inntekt er en variabel mange studier trekker frem som vesentlig i analysen av hvordan familiestruktur påvirker barnas hverdag. I internasjonal forskning har familieinntekt vist seg som en viktig determinant for elevers prestasjoner. I dette utvalget synes denne effekten å være tilstede, den er imidlertid beskjeden sammenliknet med effekten av utdanningsnivå hos foreldre. Dette kan reflektere systematiske ulikheter på tvers av landegrensene, som at det i Norge er en svakere link mellom utdanning og inntekt enn i andre økonomier. Med bakgrunn

i politiske trender i Norge kan det argumenteres for at inntektsnivået i befolkningen her er jevnere fordelt enn i land med mindre grad av velferdsordninger. I et land med færre utjevneende ordninger vil trolig nivå på inntekt spres mer med grad av utdanning, og familieinntekt bli et bedre mål på hvor ressurssterk familien er. I Norge er kanskje ikke lønnsforskjellene store nok til å vise dette, og utdanning sier følgelig mer enn inntekt om familiens karakteristika på dette området.

Etter de første analysene gjenstår fremdeles noe usikkerhet knyttet til uobserverbare egenskaper ved familier og hvordan disse påvirker estimatet, samt effekten av enkeltendringer i forbindelse med en familieoppløsning.

Opgaven favner videre undersøkelser av elevenes progresjon i prestasjoner mellom 5. og 8. klasse. I en analyse av forskjellen mellom utviklingen blant individ fra ulike familiestrukturer, finnes det bevis for at barn fra familier som splittes opp har en svakere utvikling sammenliknet med barn fra intakte familier. Dette predikerer at gapet mellom disse elevgruppene også vil øke over år. I estimeringsmetoden brukt her er konstante familiekarakteristika differensiert bort, slik at familieforholdenes påvirkninger minimeres og estimatoren blir mindre forventningsskjev. Analysen gir imidlertid ikke optimal uttelling i denne oppgaven. Mangelfull informasjon i datasettet gjør det vanskelig å ta høyde for endringer i familiekarakteristika man kan tenke seg skjer simultant med en familieoppløsning. Familiens økonomiske ressurser er en variabel som har vært viktig i analysen. Denne vil i mange tilfeller endres ved et familiebrudd. En undersøkelse av dens effekt har vært vanskelig i denne oppgaven.

I en sammenlikning av resultater fra 5. klassetrinn mellom elever som senere vil oppleve en familieoppløsning og elever fra familier som holdes intakt over hele undersøkelsesperioden, finner oppgaven at elevene som de neste årene opplever et familiebrudd, allerede på dette stadiet har svakere resultater enn andre elever. Dette kan linkes opp mot konklusjonen fra de første analysene, hvor det avdekkes at familiespesifikke egenskaper har mye å si både for skoleprestasjoner og familiens evne til å holde sammen. Disse resultatene stemmer dessuten overens med konklusjonen i Cherlin et al. (1991). Virkningen av en fremtidig oppløsning er for øvrig liten sammenliknet med effekten av familiespesifikke karakteristika.

For å kunne tolke resultatene i progresjonsanalysene nevnt over på best mulig måte er det relevant å få oversikt over forskjeller i nivå og utvikling blant elevene i tiden før en familieoppløsning finner sted. Sammenlikningen av resultat på 5. trinn gir et bilde av situasjonen på dette tidspunktet og dekker slik behovet for informasjon om nivå. Analysene gir imidlertid ikke en forståelse av om det eksisterer en forskjell i progresjon blant elevene i forkant av familiestrukturendringen. Resultatene viser at elever fra ikke-intakte familier har en svakere utvikling i prestasjoner i skolen etter familiebruddet. Om denne trenden eksisterte også i forkant av bruddet er usikkert. Forskjellen i prestasjonsnivå mellom elevgruppene på 5. trinn kan være et resultat av ulike utviklingsmønstre allerede før familiestrukturendringen.

Opgaven presenterer videre en mer nyansert analyse av gruppen elever kategorisert i intakte familier. Resultatene viser at et familiebrudd har en større negativ innvirkning på elever fra familier med samboende foreldre enn elever med gifte foreldre. Dessuten avdekker analysen en utgangsforskjell i prestasjoner mellom elever fra disse familietyper, hvor elever med gifte foreldre presterer noe bedre enn elever med samboende. Dette kan reflektere at også forskjellen i virkningen av en familieoppløsning er et resultat av ulikheter i gjennomsnittlige familieegenskaper over familiekategoriene.

Analysen av heterogene effekter kan ikke konkludere med at det eksisterer en signifikant forskjell på tvers av kjønnene i hvordan barna påvirkes av familiestruktur, slik det er funnet i Bertrand og Pan (2011). Dette gjelder også for effektene av familiestruktur for elevgrupper med høyt og lavt utdannede foreldre. Resultatene viser ingen signifikant forskjell i betydning av en familieoppløsning for progresjonen i lesing og matematikk på tvers av elevgruppene.

Spørsmålet om det kan eksistere en kausalitet mellom familiestruktur og elevers prestasjoner i skolen gjenstår. Det er problematisk å trekke en konklusjon om hvorvidt et kausalt forhold foreligger mellom familiestruktur og elevers prestasjoner. MKM-analysene taler for at familieegenskaper er de avgjørende faktorene for prestasjonsmønsteret observert på tvers av elevgruppene. Difference in differences-analysene viser en forskjell i progresjon blant elevene, selv etter at påvirkningen av konstante familiekarakteristika er eliminert. Om denne trenden eksisterte også i forkant av familiebruddet er imidlertid usikkert.

Etter denne studien finnes fremdeles noen løse tråder som overlates til senere forskningsarbeid å nøste opp i. Betydningen av endring i familieinntekt som følge av en

familieoppløsning er fremdeles ukjent. Ved å hente inn informasjon om familieinntekt i etterkant av en familieoppløsning kan en enkel difference in differences-regresjon indikere hvordan en familieoppløsning påvirker progresjon i skolerestater via inntektsendringer. I tillegg er bakgrunnsinformasjonen for difference in differences-analysen begrenset. En mulig strategi for å undersøke hvorvidt det er forskjeller i utvikling i skoleprestasjoner mellom elever fra intakte og senere oppløste familier allerede i forkant av familiestruktureringen, er å benytte seg av observasjoner fra 9. klassetrinn. Observasjoner av individ gjennom 5., 8. og 9. klassetrinn gir materiale for å estimere forskjeller i prestasjonsutvikling frem til en familieoppløsning og videre den umiddelbare effekten av hendelsen på elevresultatene. En slik analyse vil da ta utgangspunkt i en sammenlikning av elever som opplever at familien splittes opp mellom 8. og 9. trinn og elever fra intakte familier, og deres utvikling før dette, mellom 5. og 8. klassetrinn.

Analysen kan dessuten også forbedres ved i tillegg å inkludere informasjon om testscore ved kartleggingsprøver på 11. klassetrinn, VG1. Med data for fire observeringstidspunkt kan elevenes progresjon og dens påvirkning av en familieoppløsning undersøkes mer inngående. Et forslag til utvidelse av modellen kan være å sammenlikne elevens utvikling mellom 5. og 8. klassetrinn og utviklingen mellom 9. og 11. klassetrinn for elever som opplever en familieoppløsning mellom de to periodene (mellom 8. og 9. klassetrinn) og elever fra intakte familier i en difference in differences-analyse. Dette gir et mer solid mål på virkningen av en familieoppløsning, også for elevenes utvikling over årene i etterkant av familiestruktureringen. I tillegg kan man få kunnskap om en eventuell forskjell i utviklingen også i forkant av familieoppløsningen.

## 7 Referanser

Bertrand, M. & Pan, J. (2011): "The Trouble with Boys: Social Influences and Gender Gap in Disruptive Behaviour." NBER Working Paper No. 17541

Biblarz, T. & Raftery, A. E. (1999): "Family Structure, Educational Attainment, and Socioeconomic Success: Rethinking the "Pathology of Matriarchy." American Journal of Sociology, Vol. 105, No. 2

Björklund, A., Ginther, D. K. & Sundström, M. (2006): "Family structure and child outcomes in the USA and Sweden." Journal of Population Economics, Vol. 20, pp. 183-202.

Cherlin, A. J., Furstenberg, F. F, Lindsay, C-L, Kiernan, K. E., Robins, P. K., Morrison, D. R. & Teitler, J. O. (1991): "Lognitudal studies of effects of divorce on children in Great Britain and the United States." Science Vol. 252, No 5011, pp. 1386-1389

Francesconi, M., Jenkins, S. P. & Siedler, T. (2010): "Childhood family structure and schooling outcomes: evidence for Germany." Journal of Population Economic, Vol. 23, pp. 1201-1231

Ginther, D. K. & Pollak, R. A. (2004): " Family structure and children's educational outcomes: blended families, stylized facts and descriptive regressions." Demography, Vol. 41, No 4, pp.671-696

Painter, G. & Levine, D. I. (2000): "Family structure and youths' outcomes. Which correlations are causal?" The Journal of Human Resources, Vol 35, No. 3, pp. 524-549

Wooldridge, J. M. (2009): "Introductory econometrics. A modern approach. Fourth edition." South-Western.