

**En undersøkelse av sammenhengen mellom
ulikhet og økonomisk vekst i OECD-landene:
Et nytt bidrag til sentrale samfunnsspørsmål**

Julian Paulsen Blytt

Masteroppgave

Masteroppgaven er levert for å fullføre graden

Master i samfunnsøkonomi

Universitetet i Bergen, Institutt for økonomi

Juni 2019



UNIVERSITETET I BERGEN

Forord

Alf Erling Risa og Kjell Vaage er veiledere for denne oppgaven.

Sammendrag

Denne oppgaven tar for seg problemstillingen: «Er det en sammenheng mellom ulikhet og økonomisk vekst i OECD-landene?». Spørsmålet undersøkes i en panelanalyse med fasteffekt- og førstedifferansemodeller. Stata 15.0. benyttes til den statistiske analysen. Undersøkelsen bygger på tidligere analyser av temaet. Hovedbidraget mitt til forskningsfeltet er anvendelsen av den nyeste versjonen av Standardized World Income Inequality Database fra 2019, med noen andre modellspesifikasjoner enn hva som er anvendt i tidligere undersøkelser. Et sentralt formål med oppgaven er å sammenligne mine, og tidligere analyseresultater, i lys av nye data og metodiske forskjeller.

Jeg finner tilsynelatende en positiv sammenheng mellom høyere ulikhet og større økonomisk vekst i en grunnleggende modell. Når jeg kontrollerer for østeuropeiske land, forsvinner den generelle sammenhengen. Det observeres fortsatt en positiv, signifikant sammenheng mellom høyere ulikhet og større vekst i Øst-Europa. Jeg finner ingen lignende resultater for andre grupper av land i analysen. Min vurdering er at dette resultatet ikke kan generaliseres, ettersom Øst-Europa har gjennomgått en historisk utvikling som ikke kan sammenlignes med de øvrige OECD-landene. Resultatene mine antyder derfor at det ikke er en generell sammenheng mellom økonomisk ulikhet og vekst i OECD som helhet. Ingen andre modellspesifikasjoner eller nye kontroller som benyttes i analysen har ytterligere implikasjoner for det grunnleggende funnet.

Noen resultater i analysen varierer med ulike modellspesifikasjoner, og skiller seg fra flere tidligere funn i den øvrige litteraturen om temaet. Resultatene i andre analyser har også stor variasjon. Jeg diskuterer at dette kan skyldes ulike metodiske valg og datakilder i forskjellige undersøkelser. Dette sår tvil om den eksterne gyldigheten av forskningen om økonomisk ulikhet, noe som impliserer at spørsmålet fortsatt preges av stor usikkerhet.

Innholdsfortegnelse

1. Innledning	1
2. Historisk utvikling av ulikhet	3
3.1 Tidligere forskning om sammenhengen mellom ulikhet og vekst	5
3.2 Kuznetskurven.....	5
3.3 Statistiske undersøkelser av økonomisk ulikhet og økonomisk vekst	7
4.1 Hvordan måle ulikhet.....	9
4.2 Gini-koeffisienten.....	10
4.3 Gjennomgang av ulikhetsdata	12
4.4 Standardized World Income Inequality Database	13
5.1 Diskusjon av metode.....	15
5.2 Endogenitet.....	17
5.2a Målefeil	17
5.2b Samtidighet, revers kausalitet og utelatt variabel-skeivhet	18
5.2c Instrumentvariabler	19
5.3 Fasteffekt og førstedifferanse.....	20
6.1 Presentasjon av datasett	22
6.2 Valg av land til analysen.....	23
6.3 Valg av tidsperiode.....	23
6.4 Presentasjon av variabler.....	25
6.5 Begrunnelse av kontrollvariabler	25
6.6 Data til variabler.....	27
6.7 Håndtering av endogenitetsproblemet	29
6.8 Dynamisk modell?	30
6.9 Presentasjon av modell.....	31
7.1 Deskriptiv statistikk.....	32
7.2 Grafisk oversikt	36
8. Diskusjon av teorier for sammenhengen mellom ulikhet og vekst	38
9.1 Presentasjon av resultater	41
9.2 Modell uten Øst-Europa	43
9.3 Modell med forskjellige tidsperioder.....	45
9.4 Sammenligning med tidligere forskning	47
9.5 Diskusjon av resultatene	48
9.6 Veien videre	50
10. Konklusjon.....	52

Figurer

Figur 1: Kuznetskurven.....	6
Figur 2: Gini-koeffisienten.....	11

Grafer

Graf 1: Data fra World Inequality Database, pre-skatt nasjonalinntekt https://wid.world/	3
Graf 2: Gjennomsnittlig BNP-vekstrate i OECD 1950-2015	4
Graf 3: Sør-Korea BNP-vekstrate og BNP per capita 1980-2015	26
Graf 4: Gjennomsnittlig vekstrate og ulikhet i OECD 1980-2015	36
Graf 5: Vest-Europa.....	36
Graf 6: Øst-Europa	36
Graf 7: USA.....	36
Graf 8: Norge.....	36
Graf 9: Gjennomsnittlig BNP per Capita i OECD 1980-2015	49

Tabeller

Tabell 1: Gini-data fra SWIID for imputasjon 1 og 100	32
Tabell 2: Korrelasjonsmatrise mellom Gini-data fra SWIID, OECD, World Bank, LIS	33
Tabell 3: Deskriptiv statistikk for andre variabler i analysen	34
Tabell 4: Korrelasjonsmatrise for variabler i analysen.....	35
Tabell 5: Førstedifferanse- og fasteffektmodeller for hele OECD	41
Tabell 6: Førstedifferanse- og fasteffektmodeller uten Øst-Europa	45
Tabell 7: Førstedifferanse- og fasteffektmodeller, 1980-2000 og 1990-2015.....	46

1. Innledning

Økonomisk ulikhet i samfunnet er et sentralt økonomisk og politisk spørsmål. De normative og praktiske implikasjonene av ulikhet debatteres flittig, uten et entydig svar. I økonomifaget er et særlig betydningsfullt spørsmål hvilke implikasjoner fordelingen av velstanden har for et land sine økonomiske fremtidsutsikter. Dersom utjevning av ulikhet har negative konsekvenser for langsiktig vekst står man ovenfor en vanskelig avveining mellom to ønskede, men motstridende politiske mål. Skal samfunnskaken vokse, eller fordeles jevnere? Dette spørsmålet har nødvendigvis store teoretiske og praktiske følger, og er derfor blitt viet stor oppmerksomhet i økonomisk teori og empirisk forskning.

De siste 25 årene har mange statistiske undersøkelser blitt utført vedrørende sammenhengen mellom økonomisk ulikhet og økonomisk vekst. I denne litteraturen har resultatene hatt stor variasjon, noe som bidrar til at det fremdeles er uenighet vedrørende temaet. I 1994 ble to av feltets mest innflytelsesrike undersøkelser publisert. Persson og Tabellini (AER) og Alesina og Rodrik (QJE) fant begge en signifikant sammenheng mellom lavere ulikhet og høyere vekst. I 2000 kom Forbes (AER) til motsatt konklusjon, og antydte at det finnes en trade-off mellom ulikhet og vekst. Andre undersøkelser både underbygger og undergraver begge resultatene. Det er lite som tyder på at dette spørsmålet blir avklart i nær fremtid, men stadig nye undersøkelser kan bidra til å øke forståelsen av temaet. Et sentralt spørsmål jeg stiller i denne oppgaven er om tidligere resultater står seg i lys av nye data.

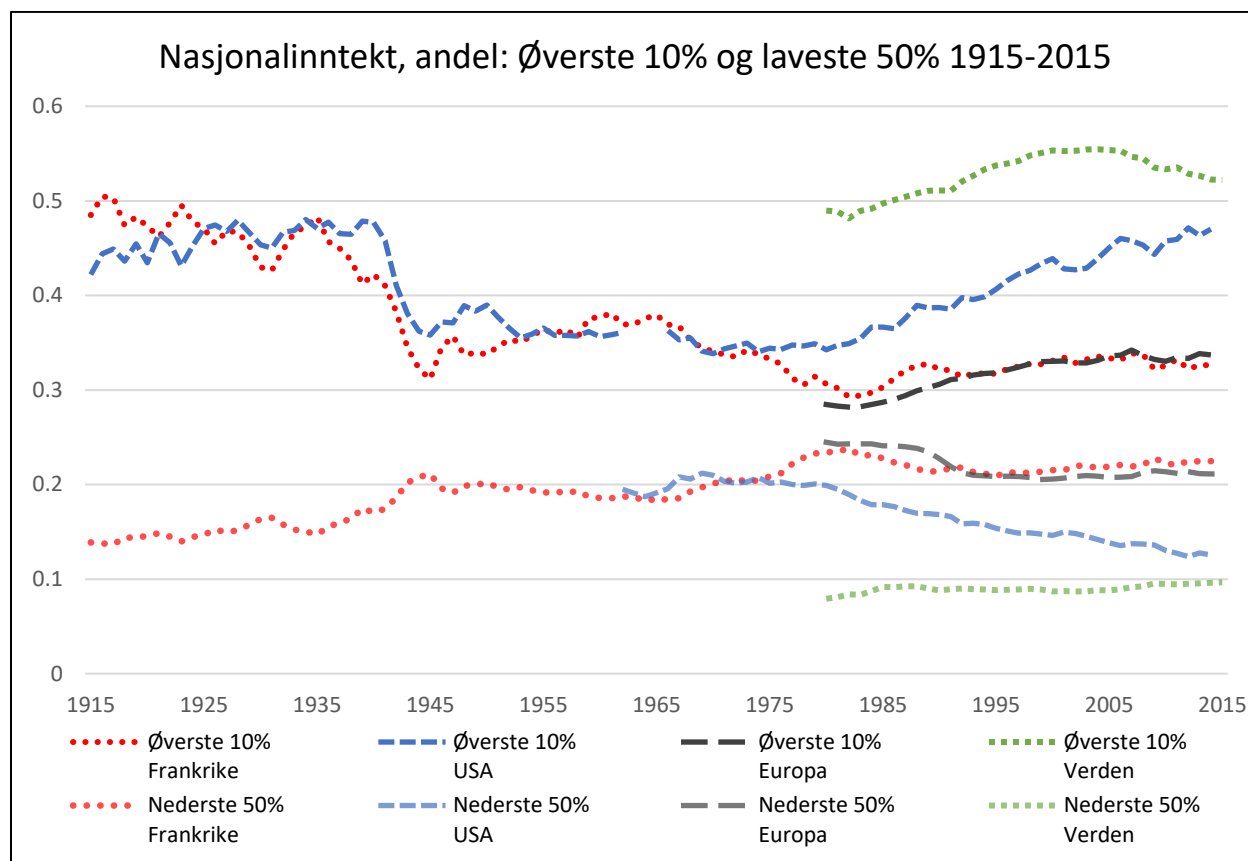
I denne oppgaven utfører jeg en regresjonsanalyse med paneldata. Økonomisk vekst er avhengig- og ulikhet er uavhengig variabel. Jeg utfører analysen med fasteffekt- og førstedifferansemodeller i tidsperioden fra 1980 til 2015, men med noen observasjoner lengre tilbake i tid. Panelet dekker alle landene som i dag er medlemmer av OECD. Til datapanelet benyttes den nyeste versjonen av «Standardized World Income Inequality Database» (SWIID) som datakilde for økonomisk ulikhet. Problemstillingen for analysen er «Er det en sammenheng mellom ulikhet og økonomisk vekst i OECD-landene?».

Hovedbidraget mitt til den øvrige litteraturen er anvendelsen av SWIID-datasettet som ble publisert i 2019. Videre så dekker min analyse et sett av land med større intern sammenlignbarhet enn mange tidligere opplegg. Dette underbygger resultatene sin pålitelighet. Jeg utfører også noen modellspesifikasjoner som skiller seg fra tidligere forskning om temaet. Undersøkelsen er særlig relevant i dag ettersom ulikheten er økende i vestlige land. Sammenhengen mellom ulikhet og vekst er også spesielt relevant for Norge, ettersom utjevning av økonomisk ulikhet har vært en viktig politisk prioritering i landet.

I analysen min finner jeg at det ikke er en generell statistisk sammenheng, hverken positiv eller negativ, mellom økonomisk ulikhet og økonomisk vekst generelt i panelet. Jeg finner derimot en særegen, positiv og signifikant sammenheng mellom økt ulikhet og økonomisk vekst i Øst-Europa, men mener at dette resultatet har begrenset ekstern gyldighet på grunn særtrekk ved disse landene. Ingen andre modellspesifikasjoner i analysen min ser ut til å ha ytterligere implikasjoner for resultatene. Jeg utelukker ikke at det kan være underliggende endogenitetsproblemer i analysen, som kan så tvil om resultatene.

Videre i oppgaven gjør jeg kort rede for den historiske utviklingen av ulikhet i vestlige land, og gjennomgår tidligere forskning som denne oppgaven bygger på. Deretter gjør jeg rede for datainnsamlingen til panelet. Da begrunner jeg også hvorfor jeg har valgt SWIID-datasettet, samt hvilke fordeler og utfordringer dette medfører. Før analysen diskuterer jeg også det metodiske grunnlaget for opplegget og gjør rede for endogenitetsproblemer. Videre legger jeg frem deskriptiv statistikk fra panelet. Til slutt presenterer og diskuterer jeg resultatene fra panelanalysen min, og trekker konklusjoner basert på funne mine.

2. Historisk utvikling av ulikhet



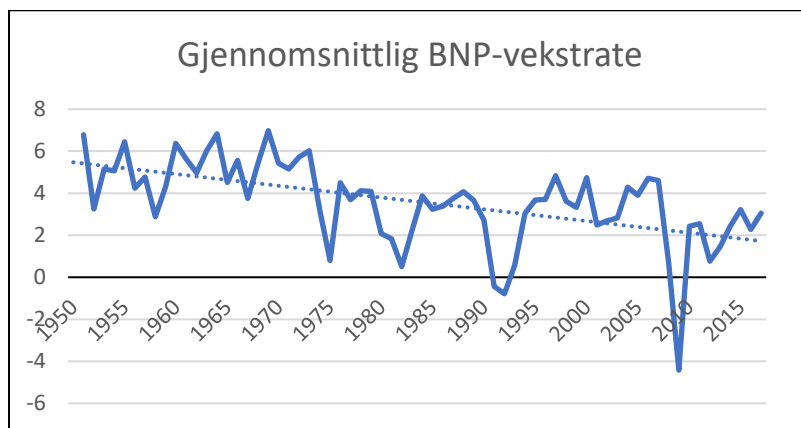
Graf 1: Data fra World Inequality Database, pre-skatt nasjonalinntekt <https://wid.world/>

I begynnelsen av den industrielle æraen så man en skarp økning i den økonomiske ulikheten i vestlige samfunn, i regi av den økende økonomiske innflytelsen til industriere vis a vis den fremvoksende arbeiderklassen. Videre utover 1900-tallet økte kravene om økonomisk omfordeling, og teknologiske innovasjoner kom flere til gagn. Likevel økte inntektskonsentrasjonen i det økonomiske toppsjiktet i flere tiår inn i 1900-tallet, og nådde sin topp på 1920 og 1930-tallet. Denne trenden er observert for de fleste vestlige samfunn, som dokumentert i arbeidet til eksempelvis Atkinson, Piketty og Saez, som har utført omfattende datainnsamlinger for den historiske utviklingen av ulikhet. Eksempelvis i Atkinson m. fl. (2011) gir en dyptgående gjennomgang av utviklingen av ulikhet i USA, og Piketty (2013), setter søkelys på implikasjonene av den historiske utviklingen av ulikhet i verden. Deres, og flere andre data er samlet i World Inequality Database (2018), og fremstilles i graf 1.

Den store depresjonen og andre verdenskrig ledet til et dramatisk paradigmeskifte i politisk tankegang og krav fra folket, noe som banet vei for sosialdemokratiets gullalder i etterkrigstiden. I denne perioden opplevde man en stadig vekst i gjennomsnittlige inntekter, og relativ reduksjon i gapet mellom rike og fattige, samtidig som den økonomiske veksten var stor.

Den fallende trenden i økonomisk ulikhet var dog ikke varig, og etter stagflasjonen på 70-tallet så man igjen en endring i økonomisk tankegods i favør av nyklassisk økonomisk teori om åpne markeder og mindre inngripen for å opprettholde den langsiktige veksten. Økende globalisering i de neste tiårene skapte også nye utfordringer for velferdsstatens rigiditeter i et stadig mer kompetitivt internasjonalt klima. Demografiske endringer som økende andel av eldre i vestlige samfunn, og større innvandring, har også medført økt press på veletablerte velferdsordninger, som har ledet til politiske reformer. Til i dag har ulikheten igjen vokst, særlig i USA, men også i Europa. Avviklingen av kommunisme i Øst-Europa forsterket også den pågående trenden, som følge av den dramatiske overgangen fra planøkonomi med sterkt fokus på økonomisk utjevning, til kapitalistiske, liberale demokratier med begrensede sikkerhetsnett. Det er denne perioden av økonomisk historie som står i fokus i denne oppgaven.

Samtidig med den økende ulikhetstrenden i moderne tid, har graden av økonomisk vekst falt for vestlige land siden etterkrigstiden. Dette illustreres i graf 2. Det finnes ulike teorier for hva dette kan skyldes, men for denne oppgaven er det sentrale spørsmålet i hvilken grad den fallende veksttrenden skyldes økonomisk ulikhet.



Graf 2: Gjennomsnittlig BNP-vekstrate i OECD 1950-2015

3.1 Tidligere forskning om sammenhengen mellom ulikhet og vekst

Diskusjonen av økonomisk ulikhet i samfunnet er et tidløst tema som omfatter en rekke forskjellige fagfelter. Hva enn de normative merittene av ulikhetsutjevning angår, har denne oppgaven interesse for temaet fra et økonomisk vekstperspektiv. På hvilke måter, og i hvilken grad har inntektssammensetningen i et land implikasjoner for et land sin økonomiske vekst? Klassisk økonomisk teori advarte mot offentlig inngripen, selv der hvor formålene var altruistiske. Følgelig er det da slik at normative ønsker om å fordele velstanden jevnere vil virke mot sin hensikt dersom dette forårsaker lavere økonomisk vekst på sikt. Den store depresjonen kastet i tvil økonomiske konvensjoner. Nye bidrag til fagfeltet, som Keynes (1936) representerte et paradigmeskifte i mange økonomers syn på offentlig inngripen. Uten å gjennomgå hele utviklingen av økonomisk tankegods siden den gang, er det klart at økonomisk teori i dag har stor variasjon. Det er mange motstridende ideer om betydningen av statlig inngripen for et samfunns fremtidige velstand. Hva enn de teoretiske rammeverkene sier om saken, mener jeg at man må gå empirisk til verks for å virkelig forstå sammenhengen mellom økonomisk ulikhet og vekst.

3.2 Kuznetskurven

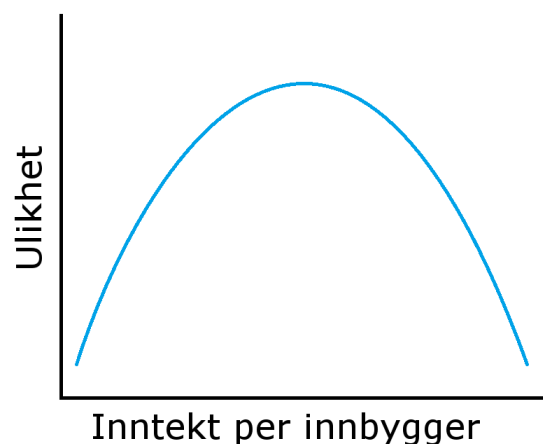
Et av de første og mest innflytelsesrike bidragene til denne litteraturen er Simon Kuznets sitt arbeid på 50- og 60-tallet. Basert på de historiske trendene i økonomisk vekst og ulikhet, slik han så dem, presenterte Kuznets en hypotese om at ulikheten først øker, og deretter faller, ved økt økonomisk velstand (1955).

Kuznets mente at den tidlige industrialiseringen av et samfunn ville lede til store økonomiske omveltninger med negative konsekvenser for det lavere samfunnssjiktet, i form av økt urbanisering med negativt lønnspress og dårligere vilkår for den fremvoksende arbeiderklassen. Samtidig opplever høyere samfunnslag økende profitt og forhandlingsmakt. I denne perioden vokser ulikheten. Etter hvert som økonomier modnes øker kravene om demokratisering og velferdsordninger. Samtidig blir humankapital stadig mer viktig for videre effektivitetsøkning. I regi av disse faktorene ser man en reduksjon i ulikheten i mer avanserte

økonomier. Kuznets legger derfor frem en hypotese om at endringen i ulikhet ved økt inntekt per capita i et samfunn følger en omvendt U-kurve, etter hvert som landet utvikler seg. Denne kurven bærer i dag Kuznets sitt navn og er illustrert i figur 1.

Denne sammenhengen virket veldig klar tatt i betraktning dataene som var tilgjengelige på Kuznets sin tid, hvor man så en stadig reduksjon i ulikheten. På tross av dette innrømmet Kuznets selv at sammenhengen var spekulativ, med svakt empirisk grunnlag (1955). Kuznetskurven har i senere år blitt mer omstridt, ettersom det igjen observeres økende ulikhet. Eksempelvis Piketty (2013, s. 24) skriver at Kuznets feiltolket den observerte nedgangen i ulikhet fra 1930-1950-tallet som sluttpunktet i utviklingen av forholdet mellom økonomisk vekst og ulikhet. I senere år viser det seg at trenden har snudd, og ulikheten er økende. Kuznets sin teori har likevel hatt stor innflytelse på økonomifaget.

Det er verd å nevne at Kuznets så økonomiske vekst som den ledende faktoren i sammenhengen mellom de to parameterne, mens ulikheten endrer seg i kjølvannet av denne. Slik berører Kuznets sin teori en sentral diskusjon i studiet av temaet, nemlig spørsmålet om kausalretning. Spørsmålet er om det er den økonomiske veksten som leder ulikheten, eller ulikheten som leder den økonomiske veksten. Denne problemstillingen, og følgene det har for min statistiske analyse, vil diskuteres ytterligere senere i denne oppgaven.



Figur 1: Kuznetskurven

3.3 Statistiske undersøkelser av økonomisk ulikhet og økonomisk vekst

De første store statistiske analysene av vekst og ulikhet av tradisjonen som denne oppgaven inngår i, kom på 1990-tallet. Blant de første og mest innflytelsesrike bidragene var Persson og Tabellini (1994) og Alesina og Rodrik (1994), som ble nevnt innledningsvis. Begge forskningsoppleggene kom til samme konklusjon om at større ulikhet reduserer den økonomiske veksten. Undersøkelsene skapte stor oppmerksomhet rundt temaet, og senere bidrag som Perotti (1996) kom til samme konklusjon. I etterkant har det blitt reist kritikk av disse tidlige undersøkelsene. Analysene benyttet relativt svake datakilder, med lav sammenlignbarhet på tvers av land, og mindre sofistikerte statistiske metoder. Tverrsnittsanalyser ble utført der hvor panelanalyser nå gjør seg mest gjeldende på feltet. I del 5 diskuterer jeg hvorfor mitt synspunkt også er at panelanalyser er bedre egnet til statistiske undersøkelser av ulikhet og økonomisk vekst.

Et gjennombrudd i feltet var publikasjonen av Deininger og Squire (1996) sitt datasett om ulikhet, som av noen ble ansett for å være en klar forbedring fra datamaterialet i tidligere opplegg. Dette åpnet for anvendelsen av sterkere metodikk i senere statistiske analyser. Dette såkalte DS1996-datasettet ble senere anvendt i undersøkelser av blant andre Kristin Forbes (2000) og Robert Barro (2000). Forbes utførte en panelanalyse med samme spesifikasjoner som Perotti (1996), hvor hensikten var å etterprøve tidligere resultater, men med metodiske forbedringer og rikere datakilder. I strid med tidligere resultater, finner Forbes tvert imot at landene med høyere ulikhet har opplevd større vekst. Barro (2000) fant en statistisk sammenheng mellom høyere ulikhet og lavere vekst i utviklingsland, og høyere ulikhet og høyere vekst i avanserte økonomier. Selv om Atkinson og Brandolini (2001) kritiserte DS1996-datasettet og Roodman (2009) fant svakheter ved Forbes sitt forskningsopplegg, har undersøkelsene tidlig på 2000-tallet vist seg å være et veikart for senere forskning på feltet. Det store nye bidraget i litteraturen på denne tiden var bruk av panelmodeller for å bedre redegjøre for endogenitet. Dette har blitt standarden i senere forskning om ulikhet og vekst.

Dollar og Kraay (2001) utfører et lignende opplegg, med et utvidet datasett, flere kontroller og andre innfallsvinkler til problematikken. De finner at vekst er gunstig for de fattige, mens forsøk på økonomisk utjevning ikke har den tilsiktede effekten. Castelló-Climent

(2010) utførte en analyse med særlig fokus på land- og regionspesifikk variasjon, med søkelys på avanserte økonomier. Hun finner at ulikhet i inntekt og humankapital virker negativt på veksten i utviklingsland, men at sammenhengen forsvinner, eller peker i motsatt retning i avanserte økonomier. Halter m. fl. (2014) utførte en analyse med særlig vektlegging av forskjellige tidsperspektiver. De finner varierte resultater vedrørende sammenhengen mellom ulikhet og vekst på kortere og lengre sikt. Scholl og Klasen (2018) utfører en analyse basert på undersøkelsen til Forbes (2000), men finner at mye av Forbes sine resultater skyldes den store veksten og dramatiske økningen i ulikhet i Øst-Europa på 1990-tallet. Kontrollert for dette, ble det observert lite sammenheng mellom variablene i opplegget. Et annet nytt bidrag i litteraturen er Brueckner og Lederman (2018), som setter belyser viktigheten av et lands initiale inntekt for hvordan ulikheten påvirker veksten i senere perioder.

Det er flere grunnleggende innsikter å hente fra undersøkelsene som har blitt utført om sammenhengen mellom ulikhet og vekst. Forskningsoppleggene har gått fra teoriens verden, til å anvende stadig mer velutviklede statistiske modeller. Dette har kun latt seg gjøre ettersom datagrunnlaget for statistiske analyser har blitt bedre. Dette har redusert risikoen for problemer som endogenitet. Observasjonene har blitt mer utfyllende og sammenlignbare, med mer informasjon både på tvers av land, og langs tidsaksen.

En annen viktig lærdom fra forskningen som har blitt gjort så langt er hvor stor variasjon det har vært i resultatene fra ulike forskningsopplegg. Et eksempel er ulike resultater i overgangen fra tversnittdata på 90-tallet, til paneler i senere tid. Likeledes finner man også klar variasjon i resultatene avhengig av hvilke land som analyseres og hvilke kontroller som benyttes. Barro (2000), Castelló-Climent (2010) og Brueckner og Lederman (2018) finner ulike sammenhenger i lavinntektsland og høyinntektsland. Scholl og Klasen (2018) finner variasjon i resultatene når man gjør rede for den særegne situasjonen i Øst-Europa. Halter m. fl. (2014) finner ulike resultater på kortere og lengre sikt, og med forskjellige metodiske innfallsvinkler.

På bakgrunn av dette er det viktig å holde tungen bent i munnen vedrørende hvor generaliserbare funnene fra ulike opplegg kan sies å være. Valg av land og valg av metode kan ha en innvirkning på hvilke resultater som fremkommer i analysen. Statistiske undersøkelser av

ulikhet og vekst er også et relativt nytt felt, hvor det stadig introduseres nye innovasjoner i de metodiske oppleggene som utføres. Alle nye publikasjoner på feltet må forstås som et bidrag til en gradvis økning av forståelsen i fagfeltet, heller enn et endelig resultat som avgjør debatten en gang for alle. Forskningsfeltet har fortsatt en lang vei å gå.

4.1 Hvordan måle ulikhet

Det kanskje mest sentrale spørsmålet forut for enhver analyse av økonomisk ulikhet er hvordan ulikhet skal defineres og måles. Jeg understreker at i min definisjon av begrepet må ikke ulikhet forveksles med fattigdom. Ulikhet forstås i denne oppgaven som den relative fordelingen av velstand mellom forskjellige deler av populasjonen i et land. Graden av økonomisk ulikhet sier ikke nødvendigvis noe om velferdsnivået til lavinntektsindivider i, eller på tvers av samfunn.

Det sentrale spørsmålet må være hvordan man best definerer en parameter som tilfredsstillende representerer den økonomiske fordelingen i et samfunn på en meningsfylt måte. Mer praktisk er det også et spørsmål om å finne en måleenhet hvor datainnsamling lar seg gjøre med tilgjengelig informasjon, og som kan sammenlignes på tvers av land hvor omstendigheter kan variere. Det er særlig ønskelig for min analyse å kunne samle hensiktsmessige data så langt tilbake i tid som mulig. Da er det mulig å utføre makroanalyser av implikasjonene av ulikhet på lang sikt. Dette viser seg å være veldig vanskelig, siden måling av ulikhet krever detaljert informasjon om inntektssammensetning i samfunnet. Et eksempel er World Inequality Database (WID) som ble anvendt innledningsvis i denne oppgaven. I WID brukes inntektskonsentrasjonen i toppsjiktet som mål på ulikhet for de eldste observasjonene. Dette skyldes at slike data i større grad foreligger i nasjonale databaser langt tilbake i tid – noen ganger tilbake til 1800-tallet. Persson og Tabellini (1994) benytter seg også av inntektskonsentrasjonen i de øverste 20 prosentene av samfunnet som forklaringsvariabel i sin statistiske analyse av vekst og ulikhet. Dette ulikhetsmålet er ikke uproblematisk å anvende; Selv om inntektskonsentrasjonen i toppsjiktet kan være indikativt for fordelingen i samfunnet, sier ikke dette direkte hvordan inntektsfordelingen i resten av samfunnet faktisk ser ut. Etter

min mening er inntektsfordelingen i lavere sjikt av like stor, om ikke større betydning for mange analyser av økonomisk ulikhet, inkludert dette opplegget.

Det er altså ønskelig at parameteren som anvendes reflekterer inntektsfordelingen i flere kvantiler, noe som også krever at inntektsfordelingen i hele samfunnet er kjent. 20:20 raten sammenligner for eksempel velstanden i de øverste 20 prosentene av den gitte populasjonen mot de laveste 20 prosentene. Tilsvarende sammenligner Palma-raten de øverste 10 prosentene, mot de laveste 40. Andre nevneverdige parametere er Theil-indeksen, Atkinson-indeksen og Hoover-indeksen. Til bedriftsanalyser har man parametere som Galt-score. Disse målene er alle utviklet for å ha en grad av sammenlignbarhet på tvers av land eller andre analyseenheter, da de baserer seg på prosenter og rater, uavhengig av den absolutte forskjellen i ulike land. Dette er fruktbart for en statistisk analyse. Det finnes forskjellige underliggende argumenter for hvorfor parameterne defineres som de gjør, og bruksområdene kan variere. Ingen ulikhetsmål er uten kritikk, men mitt synspunkt er at Gini-koeffisienten er det mest hensiktsmessige målet på ulikhet i en analyse som min.

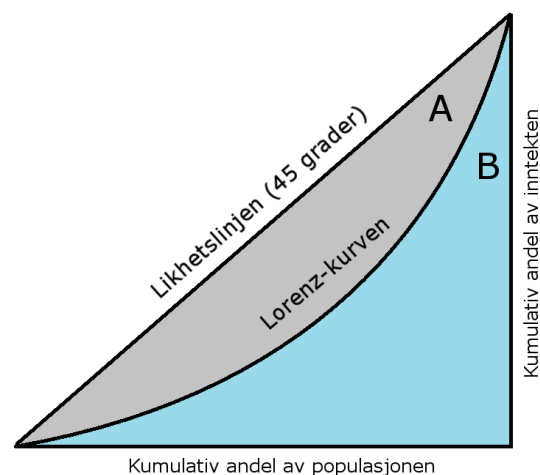
4.2 Gini-koeffisienten

Først og fremst er Gini anvendt flittig i forskning om ulikhet, blant annet i mesteparten av oppleggene som har blitt diskutert så langt i oppgaven. Det er derfor fordelaktig å følge i denne tradisjonen for å øke sammenlignbarheten mellom min, og tidligere analyser. Statistiske databaser som OECD og World Bank benytter seg også av Gini som et primært mål på ulikhet. Gini-koeffisienten tar for seg inntektsfordelingen i hele samfunnet vektet proporsjonalt, uten særlig fokus på hverken det laveste eller høyeste samfunnssjiktet. Tolkningen av Gini, mellom 0 og 1 for alle land er også praktisk for tolkning i statistiske analyser.

Gini-koeffisienten ble utviklet av Corrado Gini gjennom en rekke arbeider tidlig på 1900-tallet, og ble først presentert i boken *Variabilità e Mutabilità* (1912). Gini-koeffisienten estimeres på bakgrunn av Lorenz-kurven som illustrerer den kumulative andelen av et lands velstand som tilfaller ulike deler av populasjonen, fra de fattigste til de rikeste, i et XY-diagram. På X-aksen er den kumulative andelen av landets innbyggere fra fattigst til rikest. Den

kumulative andelen av landets inntekt i hver del av populasjonen illustreres langs Y-aksen. Følgelig vil en helt jevn fordeling innebære en rett 45 graders linje, kalt likhetslinjen, hvor hele populasjonen har lik inntektsfordeling. Lorenz-kurven er en observert kurve under likhetslinjen som illustrerer den faktiske inntektsfordelingen der hvor det ikke er absolutt likhet. Gini-koeffisienten er raten av forskjell mellom Lorenz-kurven og likhetslinjen. Illustrert i figur 2 ved $A/(A+B)$. Såfremt det ikke forekommer negativ inntekt, som gjeld, vil Gini-koeffisienten alltid være positiv, og mellom 0 (absolutt likhet) og 1 (absolutt ulikhet). Disse egenskapene er svært nyttige for statistiske analyser og gir en høy grad av sammenlignbarhet. Dette forutsetter at underliggende data som anvendes for å beregne Gini-koeffisientene benytter forholdsvis lik datainnsamlingsmetode.

Selv om Gini-koeffisienten har blitt et av de mest populære og varige målene på økonomisk ulikhet er ikke parameteren uten kritikk. Gini-koeffisienten utelater mye informasjon om den underliggende inntektsfordelingen; To forskjellige fordelinger av velstanden kan gi samme Gini-koeffisient. Det er derfor også en vanlig praksis at man publiserer inntektsfordelingene dekomponert i ulike kvantiler sammen med Gini-koeffisienten i statistiske databaser. Jeg mener fortsatt at i min analyse er Gini-koeffisienten et langt mer informativt mål på ulikhet enn å se på enkelte, eller få kvantiler. Kritikken som er reist av Gini-koeffisienten, er også reist for mange andre ulikhetsmål. Mitt synspunkt er at ingen ulikhetsmål er perfekte. Mange av dem har ulike fordeler og er egnet til forskjellige formål.



Figur 2: Gini-koeffisienten

4.3 Gjennomgang av ulikhetsdata

I dag foreligger data om ulikhet i mange internasjonale og nasjonale databaser. Mange datasett som dekker ulike grupper av land, har blitt satt sammen av forskere. Til tross for dette er estimering av ulikhet fremdeles preget av en rekke problemer. Først og fremst er utfordringen å samle tilstrekkelige økonomiske data til å beregne et ulikhetsmål som Gini. For tidligere tidsperioder har dette ofte ikke latt seg gjøre, da de tilgjengelige økonomiske dataene simpelthen ikke har vært tilstrekkelig detaljerte. Det er også vedvarende utfordringer ved innsamling av ulikhetsdata i dag, og det er vanlig å benytte seg av forskjellig metodikk i ulike land. PovcalNet, Verdensbanken sin databank om ulikhet, ser på konsumdata i utviklingsland, og inntektsdata i avanserte økonomier. Siden inntektsdata for husholdninger også samles gjennom spørreundersøkelser, har det vært tilfelle at forskjellige nasjonale og internasjonale databaser bruker forskjellige survey-spørsmål og metodikk som gjør dataene mindre sammenlignbare.

Det kanskje største enkelte kildetilfanget til data om ulikhet internasjonalt er World Income Inequality Database (WIID). WIID har samlet store mengder av ulikhetsdata fra forskjellige kilder, og har blitt anvendt i flere statistiske analyser som jeg har nevnt i denne oppgaven. Det er klare svakheter i noen av datakildene med tanke på sammenlignbarhet, siden det underliggende datagrunnlaget har varierende innsamlingsmetodikk. WIID-data kan derfor være utfordrende å bruke i en statistisk analyse.

Tidligere i denne oppgaven har også World Income and Wealth Inequality Database (WID) blitt nevnt. WID skiller seg særlig ut i mengden data som foreligger langt tilbake i tid for noen land. Databasen er derfor godt egnet til lengre analyser, men den eksterne sammenlignbarheten av data for forskjellige land er begrenset.

Luxemburg Income Survey (LIS) er trolig det beste eksempelet på en database med tilstrekkelige utfyllende data til estimering av ulikhet med harmonisert metodikk for mange land. I FN (2018) sin review av forskjellige datasett om ulikhet skrives det at «[LIS] er den eneste kilden [som blir diskutert i notatet] som gir ulikhetsstatistikk kalkulert og utformet ved bruk av et uniformt sett av definisjoner, og basert på grundig harmoniserte mikrodata for å maksimere

sammenlignbarhet». Husholdningsdataene fra LIS har blitt innsamlet og publisert i uregelmessige intervaller fra 1980-tallet for 50 land. Hovedproblemet ved bruk av LIS er manglende observasjoner mellom tidsintervallene.

4.4 Standardized World Income Inequality Database

Til denne analysen vil jeg anvende Standardized World Income Inequality Database (SWIID) som kilde for ulikhetsdata til panelet mitt. SWIID er et ambisiøst forsøk, utført av professor Frederick Solt, på å harmonisere mangfoldige ulikhetsdata. Hensikten med SWIID er å gjøre rede for problemene vedrørende datamangel og svak datasammenlignbarhet i databaser om ulikhet.

Observasjonene i SWIID-datasettet er ikke samlet fra en uniform datakilde og innsamlingsmetodikk, men er beregnet ved bruk av imputasjoner fra en rekke forskjellige datakilder for ulikhet. Grunnen er at ulike datasett dekker forskjellige tidsperioder og land, hvor alle har forskjellig grad av manglende data. Ved å sammenligne ulike datakilder kan man med større sikkerhet fylle inn manglende verdier og sette sammen et helhetlig datasett. Blant de ulike kildene som brukes er LIS, WIID, OECD sin database, og PovcalNet. LIS sine data og definisjoner benyttes som referansepunkt for imputeringsprosessen. Data fra andre kilder blir etter beste evne harmonisert med LIS-datasettet for å fylle inn manglende dataobservasjoner (Solt, 2019).

Det er grunnleggende usikkerheter ved datagenereringsprosessen i SWIID. Det er umulig å avgjøre sikkert om imputerte verdier korrekt predikerer manglende LIS-data, eller at den eksterne sammenlignbarheten mellom ulike datakilder er stor nok til at disse kan harmoniseres. I erkjennelsen av dette mener jeg at SWIID-datasettet bør brukes med varsomhet. Mitt synspunkt er fortsatt at utfordringene med SWIID er mindre i mitt panel, sammenlignet med opplegg med større, mer varierte landutvalg. Ulike datakilder og datamengder foreligger for imputering av manglende data for forskjellige land i SWIID. Generelt sett er det vestlige land som har flest dataobservasjoner, og fra mer sammenlignbare datakilder. Det er for eksempel verdt å nevne at 32 av de 50 landene som undersøkes i LIS er OECD-land. OECD sine egne ulikhetsdata, som dekker alle medlemslandene i varierende grad, anvendes også ekstensivt til

imputering i SWIID-datasettet. Som nevnt tidligere har Verdensbanken forskjellig datainnsamlingsmetodikk for høyinntektsland og lavinntektsland, som gjør at dataene har mindre sammenlignbarhet på ulike velstandsnivåer. Disse dataene benyttes også for imputering i SWIID. Jeg mener at SWIID-data hvor estimatene er basert på et rikere datagrunnlag med større sammenlignbarhet kan forventes å være mer pålitelige. Dette er en viktig grunn til at OECD-landene undersøkes i min analyse. Hadde landpanelet mitt vært større, og omfattet et mer variert utvalg av land med svakere datagrunnlag, hadde jeg hatt større reserverasjoner med å bruke SWIID-datasettet.

På tross av disse utfordringene er det klare fordeler ved å anvende SWIID-datasettet. Andre databaser for ulikhet står ovenfor lignende utfordringer som SWIID vedrørende dataharmonisering. Selv om SWIID ikke har noen egne primærdata, mener jeg at det er en fordel at SWIID forholder seg til mange ulike datakilder samtidig, når data harmoniseres. Dette gjelder særlig når datagrunnlaget er stort. Jeg mener at dette øker, heller enn svekker datasammenlignbarheten på tvers av land i SWIID-datasettet, relativt til de fleste andre datakilder. Unntaket er LIS-datasettet, men LIS lider som nevnt av stor datamangel langs tidsaksen. SWIID har flere observasjoner for ulike land og tidsperioder enn noen andre databaser som har blitt diskutert i denne oppgaven. Dette er fordelaktig for denne analysen, ettersom datamangel skaper vanskeligheter for estimeringsprosessen i panelanalysen. Svak datasammenlignbarhet i panelet kan så tvil om resultatene i sin helhet.

I FN (2018) sitt notat om forskjellige datakilder for ulikhet demonstreres det også at SWIID har en stor grad av sammenlignbarhet med resultatene fra andre databaser, på tross av usikkerheten i datasettet. Dette kommer jeg også til å demonstrere senere i denne analysen, i del 7.

På grunn av usikkerheten i imputasjonsprosessen er 100 multiple imputasjoner estimert for hver land- og tidsobservasjon i SWIID-datasettet. SWIID sin brukermanual oppfordrer til å utføre «multipel imputasjon»-estimering av dataene for å gjøre bedre rede for den underliggende usikkerheten og variansen som følger av dette. Da estimeres modeller separat for hver av de 100 imputerte observasjonene. Deretter legges resultatene sammen og

standardfeilene kombineres for hver modell. Den ferdige analysen tar da høyde for variansen i de imputerte estimatene slik at inferenstesting blir mer presis. Prosessen følger reglene lagt frem av Rubin (1987). «Multippel imputasjon»-estimering utføres i Stata med «mi estimate»-kommandoen. Med mindre jeg eksplisitt informerer om noe annet, vil alle regresjonsanalyser i denne oppgaven benytte seg av denne metodikken.

Selv om jeg utfører denne analysen med SWIID-data, og mener at datasettet kan anvendes med stor sikkerhet for OECD-landene, så må det understrekes at det hadde vært ønskelig å anvende et datasett hvor en uniform datainnsamling har vært tatt i bruk. Se for eksempel Jenkins (2015) for en kritisk gjennomgang av SWIID, og problemene rundt å anvende dette datasettet. Se også Solt (2019) for en mer ekstensiv gjennomgang av metodikken bak SWIID-datasettet.

5.1 Diskusjon av metode

I den formelle analysen har jeg besluttet å utføre en regresjon med paneldata for å undersøke den statistiske sammenhengen mellom økonomisk ulikhet og økonomisk vekst i OECD-landene. Panelanalyser er analyser tuftet på datasett som varierer både i et tverrsnitt og over tid. I dette tilfellet de 36 landene i OECD, med data på relevante variabler så langt tilbake som datakildene tillater, som i grove trekk er tilbake til starten av 80-tallet for de fleste land.

For opplegget mitt er det flere fordeler ved å utføre en panelanalyse. Fremfor alt åpner paneldata for muligheten til å kontrollere for uobserverte landspesifikke forhold. Dette kommer jeg tilbake til senere i dette avsnittet. Ytterligere så ville datagrunnlaget mitt vært forholdsvis lite dersom analysen kun dekket de 36 landene over én periode, eller ett land over tid. En større mengde observasjoner gjør det lettere å estimere sammenhenger mellom variablene. Teoretisk er også en panelanalyse bedre egnet i opplegget mitt. Det er rimelig å anta at en mulig sammenheng mellom økonomisk ulikhet på økonomisk vekst vil virke over et lengre tidsspenn. Mange landspesifikke forhold kan også gjøre at observerte sammenhenger innad i et land ikke nødvendigvis har ekstern gyldighet. Både et større utvalg av land, og en lengre tidsserie er

derfor hensiktsmessig å undersøke. Av nevnte årsaker mener jeg at en panelanalyse er langt bedre egnet i mitt opplegg, fremfor en tverrsnittsanalyse eller tidsserieanalyse.

I denne oppgaven vil jeg anvende en fasteffektmodell og en førstedifferansemodell på en rekke modellspesifikasjoner. Senere i dette avsnittet, og forut for den formelle analysen, vil jeg gjennomgå grunnlaget for valget av disse estimatorene.

Grunnleggende sett er antakelsen i en regresjonsanalyse at det er en statistisk signifikant, vanligvis lineær sammenheng mellom forklaringsvariabel, i dette tilfellet graden av ulikhet, og avhengig variabel, som er økonomisk vekst. Forut for analysen forutsetter man at kausalretningen er etablert, hvor min grunnleggende antakelse er at ulikheten påvirker veksten. I et panel med flere forklaringsvariabler eller kontroller, som i min analyse, kan sammenhengen uttrykkes ved:

$$y_{it} = \beta_1 + \beta_2 x_{it2} + \beta_3 x_{it3} + \dots + \beta_k x_{itk} + \varepsilon_{it}$$

Som på vektorform tilsvarer:

$$y_{it} = x'_{it}\beta + \varepsilon$$

y er avhengig, og x er uavhengige variabler. k er antallet uavhengige variabler inkludert i den gitte modellen. i er betegnelsen for hver observasjonsenhet i tverrsnittet ($i=1,2,\dots,n$) og t er betegnelsen for hver tidsperiode ($t=1,2,\dots,T$). β er koeffisienter for sammenhengen mellom forklaringsvariablene og avhengig variabel, og ε er restleddet.

Hensikten med en regresjonsanalyse er vanligvis å estimere β hvor observerte forklaringsvariabler og avhengig variabel er best mulig tilpasset den antatte underliggende sammenhengen mellom variablene. Inferenstesting avgjør om denne sammenhengen er signifikant. Funksjonen for estimert β for gitt x i et panel med OLS er:

$$\hat{\beta} = (\sum_{i=1}^N \sum_{t=1}^T x_{it} x'_{it})^{-1} \sum_{i=1}^N \sum_{t=1}^T x_{it} y_{it}$$

I denne undersøkelsen er min oppfatning at en slik estimator ikke innfrir de nødvendige forutsetningene som kreves for å estimere sammenhengen forventningsrett og konsistent. Dette skyldes det underliggende endogenitetsproblemet som ofte forekommer i økonomiske

analyser. I neste delavsnitt vil jeg gjøre rede for hva endogenitetsproblemet innebærer, og hvilke teknikker som finnes for å redusere endogenitetsproblemet, som er av relevans for dette opplegget.

5.2 Endogenitet

Endogenitet er en av de mest grunnleggende, og mest utfordrende problemene som et økonometrisk opplegg står ovenfor. Kort sagt er endogenitet en situasjon der hvor forklaringsvariabelen korrelerer med feilledet, som bryter med Gauss-Markov-teoremet for «Best Linear Unbiased Estimator». Estimatoren er da ikke forventningsrett eller konsistent. Hovedkildene til endogenitet er målefeil, samtidighet, revers kausalitet og utelatte variabler, som må gjøres rede for forut for den formelle analysen.

5.2a Målefeil

Målefeil i avhengig variabel y gir tap av effisiens, men et langt større problem er målefeil i uavhengig variabel x , som kan medføre forventningsskeive resultater. Målefeil er alltid en problematikk som kan forekomme i undersøkelser av sammenhenger hvor datainnsamling er vanskelig. I økonomifaget finnes det mange variabler som ikke har en eksakt, universelt akseptert definisjon, og hvor datainnsamlingen enten er mangelfull eller inkonsistent. Dette gjelder også for økonomisk. Problemene med datakilder om ulikhet diskuterte jeg i avsnitt 4. Der understeker jeg at ingen ulikhetsmål er perfekte, og problemer med ulike datakilder diskuteres, inkludert SWIID som brukes i denne analysen. Jeg gir også et eksempel på utfordringen med målefeil i ulikhetsdata i del 3 – DS1996-datasettet ble ansett av noen forskere for å være en klar forbedring fra datakildene som ble brukt i tidligere opplegg, men på samme tid var det andre forskere som stilte seg kritisk til datasettet på bakgrunn av svak sammenlignbarhet (Atkinson og Brandolini, 2001).

Utfordringer vedrørende målefeil er etter min mening et åpent spørsmål der hvor det er utfordringer med datainnsamling og operasjonalisering av parametere, som så ofte er tilfelle

når samfunnsspørsmål undersøkes. Det må alltid diskuteres om datakvalitet og operasjonalisering av parametere muligvis kan forbedres i fremtidige opplegg. Historiske erfaringer i studiet av ulikhetsdata viser at feltet alltid har vært preget av problemer vedrørende datakvalitet. Dette reiser spørsmål om oppleggenes gyldighet. Som diskutert i tidligere mener jeg at SWIID-data har klare fordeler i mitt panel sammenlignet med andre databaser. Valget om å bruke SWIID-datasettet for OECD-landene er i stor grad gjort for å redusere risikoen for målefeil.

5.2b Samtidighet, revers kausalitet og utelatt variabel-skeivhet

Samtidighet er i korte trekk når avhengig og uavhengig variabel påvirker hverandre simultant, og revers kausalitet er en situasjon der hvor den virkelige sammenhengen mellom variablene er motsatt av modellens antakelser. Kausaleffekten peker altså i motsatt retning av antatt. Revers kausalitet og samtidighet er nært relaterte konsepter, og vil diskuteres om hverandre i dette delavsnittet.

$x \rightarrow y$ «Vanlig» sammenheng

$x \leftarrow y$ Revers kausalitet

$x \leftrightarrow y$ Samtidighet

Rent teoretisk er det problematisk om man ikke tilstrekkelig kan etablere kausalitet i undersøkelsen av årsak og virkning. Mer praktisk er det også en utfordring for selve undersøkelsens gyldighet, da samtidighet og revers kausalitet leder til forventnings skeivhet og inkonsistens. Problematikken med samtidighet og revers kausalitet kan skildres ved:

$y = \beta x + \varepsilon$ y er avhengig av x

$x = \gamma y + \nu$ x er avhengig av y

Derav:

$x = \gamma(\beta x + \varepsilon) + \nu$ x er avhengig av feilledet ε .

Problemer med utelatte variabler forekommer når en uobservert variabel z korrelerer med både x og y , og forårsaker deler av den observerte sammenhengen mellom variablene:

$x \rightarrow y$ «Vanlig» sammenheng

$z \rightarrow x, z \rightarrow y$ Utelatt variabel

Sett at det er en sammenheng mellom x og z , og y forklares ved:

$$y = \alpha + \beta x + \delta z + \varepsilon$$

Hvis z er utelatt fra modellen blir estimert β :

$$E(\hat{\beta}) = \beta + \delta (\sum_{i=1}^N \sum_{t=1}^T x_{it} x_{it}')^{-1} \sum_{i=1}^N \sum_{t=1}^T x_{it} z_{it}$$

Den estimerte koeffisienten tilsvarer da sammenhengen mellom x og y og den delen av sammenhengen mellom z og y som også er felles mellom x og z .

Endogenitet er en svært stor utfordring, uten noen endelig løsning. Særlig i studiet av økonomiske variabler i et stort samfunnsperspektiv, og på tvers av land, er det liten tvil om at mange variabler påvirker hverandre. Det finnes argumenter for hvorfor veksten i et land er en pådriver for endringer i ulikheten. Dette var grunnlaget for Kuznets sin hypotese om utviklingen av ulikhet, ledet av økonomisk vekst. Det er også slik at mange faktorer, både politiske, historiske og kulturelle, tenkelig kan forklare både forekomsten av ulikhet i forskjellige samfunn, og veksten i landene. For utelatte variabler-problematikken er en naturlig løsning å inkludere kontrollvariabler. Kontrollene som er valgt til denne analysen vil bli gjennomgått i del 6, men det finnes utallige tilfeller hvor de utelatte variablene ikke lett lar seg måle eller inkludere i statistiske modeller.

5.2c Instrumentvariabler

Dersom det ikke lar seg gjøre å korrigere for endogenitetsproblemet ved bruk av kontrollvariabler, er en typisk løsning å ta i bruk instrumentvariabler. Dette er også et bøtemiddel mot simultanitet og revers kausalitet. Den grunnleggende forutsetningen for en instrumentvariabel er at instrumentet korrelerer med både avhengig variabel og uavhengig

variabel, uten å korrelere med feilleddet. Et svakt instrument kan gjøre mer skade enn nytte for analysen, og instrumentvariabler bør derfor brukes med varsomhet.

Tatt i betraktning de mange måtene landparametere kan ha innvirkning på hverandre er det vanskelig å finne et godt instrument i slike statistiske undersøkelser. Instrumentvariabler vil ikke anvendes i denne analysen, da det er min oppfatning at det ikke foreligger et egnet instrument for ulikhet til mitt panel. Jeg vil derfor ikke ytterligere diskutere instrumentvariabler i denne oppgaven, men jeg understreker at et gyldig instrument for ulikhet kan forbedre fremtidige dataanalyser.

Et eksempel hvor endogenitetsproblemet blir forsøkt adressert med en instrumentvariabel er Easterly (2007). Andelen landareal i ulike samfunn som er egnet til hvetedyrking mot andelen som er egnet til sukkerdyrking brukes som instrument for ulikhet. Argumentet er at land med større relativ sukkerkultivering preges av en historie av plantasjevirkosomhet og utnyttelse, med konsekvenser for graden av nåtidig ulikhet. Hva enn instrumentets gyldighet angår, er mitt synspunkt at hvete-til-sukker-raten er et lite relevant instrument for min analyse. OECD-landene har ikke vært særlig preget av utnyttelse gjennom plantasjevirkosomhet.

5.3 Fasteffekt og førstedifferanse

Paneldata åpner for flere verktøy til korreksjon av endogenitet. Blant disse er fasteffekt- og førstedifferansemodellene, som vil bli benyttet i min analyse. Det er rimelig å anta at ulike land har mange grunnleggende forskjeller seg imellom som kan tenkes å være tidskonstante – eksempler er forskjellige historiske opplevelser og politisk kultur. Slike forhold kan påvirke både graden av ulikhet og vekst i landene. I et panel kan dette uttrykkes ved:

$$y_{it} = \alpha_i + x'_{it}\beta + u_{it}$$

Hvor α_i er konstantleddet for hvert land i analysen. u_{it} er feilleddet, som varierer langs analyseenhet og langs tidsaksen. Ved modelltransformasjoner som fasteffekt- og

førstedifferansemodeller kan den tidskonstante komponenten fjernes. Dersom uobserverte variabler som skaper endogenitet er tidskonstante, kan dette løse endogenitetsproblemet.

$$\hat{\beta}_{FE} = (\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)')^{-1} \sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) \quad \text{Fasteffekt}$$

$$\hat{\beta}_{FD} = (\sum_{i=1}^N \sum_{t=1}^T \Delta x_{it} \Delta x_{it}')^{-1} \sum_{i=1}^N \sum_{t=1}^T \Delta x_{it} \Delta y_{it} \quad \text{Førstedifferanse}$$

Hvor \bar{y} og \bar{x} er gjennomsnittet av variablene langs tidsaksen. For alle variablene i førstedifferansemodellen betegner Δ at variabelen er differensiert langs tidsaksen, eksempelvis er $\Delta x_{it} = (x_{it} - x_{it-1})$. I begge disse tilfellene fjernes den konstante, uobserverte komponenten fra modellen ettersom α_i per definisjon er lik for alle tidsperioder, og for gjennomsnittet av periodene. Metodikken er litt forskjellig i de to modellene.

Fasteffektmodellen undersøker avviket fra gjennomsnittet i variablene.

Førstedifferansemodellen undersøker endringer i konstantene mellom hver tidsperiode.

Fasteffekt-estimatoren regnes generelt for å være mer effisient enn førstedifferanse-estimatoren (Veerbek, 2017 s. 389), men den er mer sensitiv til brudd på modellen sine underliggende antakelser. Dersom det er stor seriekorrelasjon i variablene, som hvis utvalget har enhetsrot, kan FE-modellen få store problemer. I dette tilfellet er FD-modellen å foretrekke, da den bedre håndterer problemer med ikke-stasjonærhet. Dersom det er seriekorrelasjon i variablene, men i mindre grad, er det vanskelig å si hvilken estimator som er best egnet (Woolridge, 2009, s487-488). Faren for enhetsrot kan være stor med økonomiske makrodata, og dette kan lede til spuriøse sammenhenger eller skape problemer for inferenstesting. Derfor er mitt synspunkt at førstedifferansemodellen kan tenkes å være bedre egnet i dette opplegget. Samtidig kan det være vanskelig å avgjøre hvilken av de to modellene som er mest effisient, og avvik mellom FD- og FE-estimatorene kan indikere feilspesifikasjoner i modellen. Derfor kommer jeg til å presentere resultatene fra analysen med begge estimatorene.

Det må merkes at estimatorene kun baserer seg på «within»-dimensjonen i data – altså hva som forklarer endringer i avhengig variabel innad i hvert land. Koeffisienten som estimeres er gjennomsnittet av sammenhengen for hvert av landene i panelet, noe som er verd å ta i betraktning i tolkningen av resultatet. Siden panelet består av OECD-landene, som er relativt

sammenlignbare, mener jeg at dette er mindre problematisk enn i et større panel, hvor within-variasjonen kan tenkes å være større for ulike land.

Dynamiske modellspesifikasjoner hvor det kontrolleres for tidligere verdier av avhengig variabel, er vanlige i økonomiske analyser. Dette har også blitt gjort i tidligere analyser av økonomisk ulikhet og vekst. Dynamiske paneler utelukker fasteffektmodellen og førstedifferansemodellen, og åpner for andre endogenitetsproblemer som må gjøres rede for. I en slik modellspesifikasjon kan man eksempelvis ta i bruk en førstedifferensiert instrument-estimator som Arellano-Bond-estimatoren (Arellano og Bond, 1996). Eksempelvis Forbes (2000), benyttet seg av en slik metode.

Arellano-Bond-estimatoren ble sterkt vurdert i dette opplegget, men jeg gikk bort fra dette, noe som vil begrunnes i del 6. Der vil jeg også ytterligere gjøre rede for valgene jeg har foretatt for å gjøre rede for endogenitetsproblemet i denne analysen.

6.1 Presentasjon av datasett

Jeg har valgt å analysere data på bakgrunn av 5-årige tidsintervaller. Dette er gjort av flere grunner. Islam (1995) argumenterer for at årlige tidsintervaller er et uegnet tidsperspektiv i statistiske vekstanalyser på grunn av den store forekomsten av kortsiktige forstyrrelser. Mye kortsiktig variasjon følger av eksterne sjokk som er vanskelig å gjøre tilstrekkelig rede for i en statistisk modell. Den økonomiske veksttrenden over en lengre tidsperiode er derfor et bedre utgangspunkt for analysen min. Det er også slik at samtlige av undersøkelsene jeg har diskutert tidligere i denne oppgaven har benyttet seg av denne metodikken, formodentlig av tilsvarende grunner som Islam legger frem. Det er hensiktsmessig å fortsette denne fremgangsmåten for å øke sammenlignbarheten mellom min, og tidligere undersøkelser.

6.2 Valg av land til analysen

OECD-landene er valgt av flere årsaker. Som diskutert i del 4 mener jeg at SWIID-datasettet er mer pålitelig for avanserte økonomier. Økonomiske data til andre variabler i analysen foreligger også i større grad for høyinntektsland, siden nasjonale statistiske databaser i Vesten ofte har utført datainnsamling til relevante variabler i veldig lang tid.

Videre er det også grunner til å tenke at sammenhengen jeg undersøker, mellom økonomisk vekst og økonomisk ulikhet, ikke nødvendigvis er lik for land i veldig forskjellige utviklingsstadier, som nevnt tidligere. Eksempelvis Barro (2000) og Brueckner og Lederman (2018) finner ulike sammenhenger mellom vekst og ulikhet for land med forskjellige økonomiske velstandsnivåer. Det er derfor mulig at et datasett med større variasjon i landene vil ha svakere sammenhenger, og resultatene for veldig ulike land kan peke i forskjellige retninger. Derfor mener jeg at det er hensiktsmessig å ta utgangspunkt i en gruppe av land som befinner seg på relativt likt velstandsnivå. Risikoen for problemer med utelatte variabler reduseres også når landene som er valgt i analysen er henholdsvis like. OECD består av velstående, vestlige, demokratiske og kapitalistiske land, som gjør at dette er et godt grunnlag for sammenligning med ceteris paribus-antakelse om avhengig og uavhengig variabel. Amparo Castelló -Climent (2010) argumenterer tilsvarende, og gir også OECD-landene særlig oppmerksomhet i sin analyse av denne tematikken.

Det må fortsatt understrekes at et mindre panel av mer sammenlignbare land har ulemper. Selv om jeg mener at den interne sammenlignbarheten mellom landene mine er forholdsvis stor, så skjer dette på bekostning av ekstern sammenlignbarhet; Resultatene i analysen min kan ikke nødvendigvis generaliseres til andre land som er svært forskjellige fra OECD.

6.3 Valg av tidsperiode

Panelet jeg har satt sammen inneholder data på relevante variabler for alle OECD-landene for et så langt tidsperspektiv som lar seg gjøre. Tidsperioden i analysen avgjøres først og fremst av datatilgjengelighet. Derfor går analysen fra 1980, med noen unntak hvor det også foreligger

observasjoner om økonomisk ulikhet på 70-tallet. Analysen min strekker seg ikke lengre frem i tid enn til 2015 av samme grunn. Jeg erkjenner at valget av tidsperiode er noe arbitrært, og det hadde vært ønskelig med en enda lengre dataanalyse. Flere observasjoner langs tidsaksen ville beriket den statistiske analysen. Jeg mener fortsatt at 1980-tallet er et godt startpunkt for undersøkelsen. Av graf 1 innledningsvis ser man at 1980-tallet markerer en økning i ulikhetstrenden i Vesten. Derfor kan denne perioden forstås som starten av den «moderne» trenden i ulikhet som har vedvart til i dag. Til sammenligning var ulikhetstrenden fallende i etterkrigstiden.

Det må da også merkes at noen av landene som i dag er i OECD har opplevd store økonomiske og politiske endringer siden 1980-tallet. Øst-Europa er mest nevneverdig. Disse landene gjennomgikk en dramatisk omveltning ved overgangen fra kommunistisk planøkonomi til liberalt demokrati. Siden den historiske erfaringen i Øst-Europa ikke lett kan sammenlignes med resten av OECD, vil disse landene bli gitt særlig oppmerksomhet i analysen. Jeg kommer til å gjøre rede for dette på to måter. Jeg kontrollerer for Øst-Europa sin særegenhet, som også ble gjort av Scholl og Klasen (2018). Analysen utføres også både med og uten Øst-Europa i panelet.

Analysen dekomponeres også i to kortere tidsperioder, for tidsspennet frem til 2000 og for 1990-2015. Dette gjøres av flere grunner. Kuznets illustrerer muligheten for at sammenhengen mellom ulikhet og vekst ikke er konstant ettersom velstandsnivået i land øker. Siden landene i panelet har en så og si kontinuerlig vekst gjennom hele analyseperioden er det derfor mulig at observerte sammenhenger er ulike for tidligere og senere deler av tidsserien. Dette er verd å undersøke. Tidligere undersøkelser av sammenhengen mellom ulikhet og vekst har også tatt for seg ulike tidsspenn, ofte på grunn av datatilgjengeligheten på tidspunktet hvor analysene ble utført. Mye innflytelsesrikt arbeid på dette forskningsfeltet ble utført på 90-tallet og tidlig på 2000-tallet, jeg vurderer det som hensiktsmessig å utføre en analyse som dekker det samme tidsspennet. Ulikhetsvariabelen er også ubalansert i panelet mitt – noen land, særlig i Øst-Europa, har ikke ulikhetsdata før 90-tallet. Derfor mener jeg det er interessant å foreta en regresjonsanalyse fra 1990 og til slutten av panelet. Da er mengden av observasjoner jevnere per land, og færre data mangler. Dette kan styrke estimatene sin treffsikkerhet.

6.4 Presentasjon av variabler

Valget av variabler til analysen er til dels basert på tidligere forskning på feltet. Dette skyldes både at det teoretiske grunnlaget for de gitte kontrollvariablene er henholdsvis godt, og fordi sammenlignbarheten mellom mine, og tidligere resultater da blir større. Jeg vil også inkludere kontrollvariabler som etter min mening burde gjøres rede for, uten at dette har vært etablert praksis i tidligere forskning.

Avhengig og uavhengig variabel er økonomisk vekst og økonomisk ulikhet. Det kontrolleres for utdanningsnivået og det generelle velstandsnivået i landene. En mye anvendt kontroll som også brukes i denne analysen er prisnivået på investeringer, relativt til USA. Som nevnt kontrollerer jeg også for de østeuropeiske landene. En ytterligere kontroll i panelet mitt som ikke generelt har blitt anvendt i analysene jeg har diskutert så langt, er graden av internasjonal handel som et mål på globalisering. Det benyttes også kontrollvariabler for de ulike tidsperiodene. Alle kontrollvariablene vil begrunnes ytterligere i neste delavsnitt.

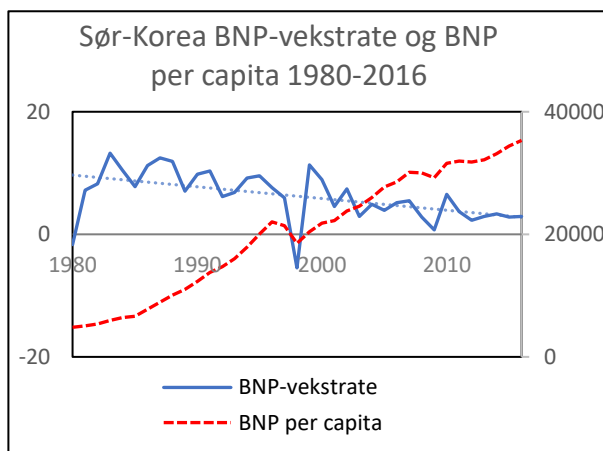
6.5 Begrunnelse av kontrollvariabler

Utdanningsnivå, velstandsnivå og prisnivået på investeringer er kontroller som har blitt mye anvendt i vekstlitteraturen, og i sammenhengen mellom vekst og ulikhet spesifikt. Alle disse kontrollene ble inkludert i Perotti (1996) sin undersøkelse, som senere ble replisert av Forbes (2000). Forbes sin analyse har deretter blitt en slags mal for flere senere undersøkelser, med ytterligere spesifikasjoner. Eksempelvis Halter m.fl. (2014) og Scholl og Klasen (2018) baserer sin modell på Forbes sin undersøkelse. Dollar og Kraay (2001) forsøker også å gjøre sin undersøkelse sammenlignbar med arbeidet gjort av Forbes ved å introdusere sammenlignbare variabler. Selv om Brueckner og Lederman (2018) benytter seg av en veldig original modellspekifisering, refereres det mye til tidligere forskning, og særlig til Forbes sin analyse.

Prisnivået på investeringer relativt til USA er en kontroll av markedsvidninger i ulike land. Ideen er at prisnivået på kapitalinvesteringer følger av forskjellige grader av statlig

inngripen og regulering, som kan ha veksthemmende konsekvenser, da kapitalinvesteringer blir dyrere.

Det generelle velstandsnivået i ulike land, målt i BNP per innbygger, er en hensiktsmessig kontroll i analysen, ettersom velstandsnivået kan prege graden av vekst; Land i tidlige utviklingsstadier kan oppleve større, mer eksplosiv vekst enn avanserte økonomier, i en transitorisk fase mot konvergens. I dette datasettet er Sør-Korea et godt eksempel, vist i graf 3.



Graf 3: Sør-Korea BNP-vekstrate og BNP per capita 1980-2015

Utdanningsnivået i landene er en kontroll for forskjeller i humankapitalen i landene. Det er mulig at folk i land med lavere ulikhet også har større generell tilgang på utdanning, enten på grunn av større generell kjøpekraft, eller på grunn av offentlige utdanningstilbud. Min antakelse forut for den formelle analysen er at land med lavere ulikhet generelt har høyere utdanningsnivå, og at høyere utdanningsnivå er positivt for veksten. Dette er derfor en klar grunn til å kontrollere for utdanning. I sin undersøkelse ser Perotti (1996) på sammenhengen mellom valg av utdanning og grad av ulikhet, og etablerer en mulig sammenheng mellom de to variablene, som underbygger denne antakelsen.

Videre introduserer jeg en kontrollvariabel for internasjonal handel, åpenhet og globalisering. Som mål på dette bruker jeg Handel-til-BNP-raten, som er import + eksport som prosentandel av BNP. Internasjonal handel er en tilsynelatende lite anvendt kontroll i litteraturen. Blant undersøkelsene som har blitt diskutert så langt er det kun Dollar og Kraay (2001) som gir betydelig oppmerksomhet til internasjonal handel. Jeg mener at det er et solid teoretisk fundament for hvorfor internasjonal handel øker den økonomiske veksten i et land. Samtidig predikerer økonomisk handelsteori at globalisering skaper tapere og vinnere innad i landene. Med økt internasjonal handel er forventningen at næringer med komparative fortrinn tjener mer, mens andre produsenter i landet får det vanskeligere på grunn av tilspisset

konkurransen. Nettogevinsten forstås som positiv for økonomisk vekst. Et mye diskutert tema er også negativt press på etablerte velferdsordninger som følge av internasjonal handel. Mange vestlige land har for eksempel relativt høye skatteinivåer, lønninger og sjenerøse trygdeordninger. Økt internasjonal handel impliserer større konkurranse med land hvor dette ikke er tilfelle. Større konkurransedyktighet kan skje på bekostning statlig velferd, negativt lønnspress eller tap av arbeidsplasser. Det er mulig at slike forhold kan påvirke graden av ulikhet i landene. Ettersom mengden internasjonal handel har økt i flere tiår mener jeg at internasjonal handel burde bli gitt større oppmerksomhet som kontrollvariabel i statistiske analyser av ulikhet.

Scholl og Klasen (2018) finner at forholdet mellom vekst og ulikhet i de østeuropeiske landene i overgang fra kommunisme til kapitalisme skiller seg veldig fra den generelle sammenhengen mellom ulikhet og vekst som man finner i andre land. De skriver at forholdene i Øst-Europa er særegne og kan ikke uten videre sammenlignes med den vanlige vekst-strukturen i land uten en lignende historisk opplevelse. Som diskutert tidligere er jeg enig i dette poenget. Mitt synspunkt er at den politiske transformasjonen, den transitoriske veksten og økningen i ulikhet i Øst-Europa på 1990-tallet er substansielt forskjellig fra noe man observerer i de andre OECD-landene. Derfor kontrollerer jeg for dette i analysen.

Mine modellspesifikasjoner og valg av data vil avvike noe fra tidligere undersøkelser. Dette skyldes i stor grad spesifikke særtrekk ved OECD-landene, hvor tidligere undersøkelser ofte har hatt større land-utvalg med mer variasjon

6.6 Data til variabler

Som diskutert tidligere er data om ulikhet hentet fra den nyeste versjonen av SWIID-datasettet (2019). Det er Gini-verdiene for disponibel inntekt som er valgt til analysen. Gini-verdiene, som går fra 0-1, er deretter multiplisert med 100. Etter min mening gir dette en mer hensiktsmessig tolkning av koeffisientene i regresjonsanalysen, ettersom Gini-målet da går fra 0-100, så én enhets økning tilsvarer ett prosentpoeng endring i ulikheten, målt i Gini.

Økonomiske data er hentet fra Penn World Table (2015). Dette gjelder BNP-vekstraten, BNP per innbygger, og data om import og eksport, samt data for prisnivå på investeringer, kalt prisnivå på kapitalformasjon. 5-årige gjennomsnitt av årlig reell vekstrate i BNP anvendes som avhengig variabel i analysen. Som nevnt tidligere er hensikten å redusere variasjonen som følger av årlige svingninger, for å bedre analysere den underliggende veksttrenden på lengre sikt.

Utdanningsdata er hentet fra den nyeste versjonen av «Barro-Lee Educational Attainment Data» av Robert Barro og Jong-Wha Lee (2013). Dette datasettet har også vært datakilden for tilsvarende parametere i tidligere analyser. I datasettet er utdanningsdata tilgjengelig i femårige intervaller for flere dekomponerte størrelser, som prosentandelen av befolkningen med sekundær eller tertiærutdanning i gjennomsnitt for hvert land, og utdanningsnivå for menn og kvinner separat. Flere tidligere undersøkelser som Perotti og Forbes har separate kontroller for utdanning blant menn og kvinner. En korrelasjonsanalyse i mitt datasett gir en koeffisient på 0,9 mellom utdanningsnivået for kjønnene – trolig fordi OECD-landene er blant de mest likestilte samfunnene i verden. Dette er vist i tabell 8 i appendiks. Det er derfor min oppfatning at det ikke tjener noen hensikt, og kan være problematisk for analysen, å kontrollere for utdanningsnivået til menn og kvinner separat når variablene korrelerer så tett. Derfor vil ikke dette gjøres i min analyse.

I mange tidligere analyser har man hovedsakelig kontrollert for andelen i befolkningen med sekundær utdanning som mål på utdanningsnivå. I min undersøkelse vil det kontrolleres for prosentandelen med sekundær og tertiær utdanning som to separate variabler. Dette skyldes at andelen med sekundær utdanning er rimelig høy i alle vestlige land. «Inflasjon» i utdanning i avanserte økonomier gjør at andelen av befolkning med sekundær utdanning alene kanskje ikke er et tilstrekkelig mål på forskjeller i humankapitalen i landene. Det er liten korrelasjon mellom andelen med sekundær og tertiær utdanning i OECD, så de to variablene kan inkluderes i analysen uten problemer med multikollinearitet, og kan muligvis belyse svært ulike forhold.

De østeuropeiske landene er definert i en dummy-variabel som alle landene i OECD som har en sovjetisk kommunistisk fortid, enten de var en del av Sovjetunionen, som de baltiske landene, eller var underlagt sovjetisk-dominerte ettpartiregimer, som Polen, Ungarn og Tsjekia. Jeg inkluderer en interaksjonsvariabel for østeuropeiske land i analysen ved å multiplisere den østeuropeiske dummy-variabelen med Gini-variabelen.

Jeg anvender også dummyvariabler for hver tidsperiode i analysen for å kontrollere for eksterne periodiske sjokk som kan påvirke økonomiske anliggender i alle landene i panelet. Et eksempel er Den store resesjonen.

I den grunnleggende modellen vil jeg først kontrollere for BNP per innbygger, prisnivå på investeringer, utdanning og tidsdummyene. Jeg gjør dette fordi disse variablene er konvensjonelle kontroller som har blitt benyttet flittig i tidligere undersøkelser som Forbes (2000). Deretter introduserer jeg interaksjonsleddet for Øst-Europa i modellen for å undersøke om dette har implikasjoner for resultatene, slik som Scholl og Klasen (2018) finner. Til slutt kontrollerer jeg for internasjonal handel ettersom dette er en ny kontrollvariabel som jeg introduserer til denne analysen, og som ikke har vært anvendt tidligere.

6.7 Håndtering av endogenitetsproblemet

I del 5 diskuterte jeg ulike kilder til endogenitet, og hvordan disse kan håndteres. For å redusere sannsynligheten for utelatt variabel-skeivhet benytter som nevnt fasteffekt- og førstedifferansemodeller. Jeg anvender også etter beste evne hensiktsmessige kontrollvariabler.

For å redusere risikoen for simultanitet og revers kausalitet utfører jeg også en annen modellspesifikasjon. Som nevnt er avhengig variabel gjennomsnittet av økonomisk vekst i hver 5-årsperiode. Forklaringsvariablene i panelet mitt er på sin side målt i starten av hver femårsperiode. Ideen er at de uavhengige variablene i starten av periodene kan påvirke BNP-veksten i den kommende perioden, mens gjennomsnittsveksten gjennom perioden ikke kan påvirke de uavhengige variablene i starten av perioden. Det er ikke slik at denne spesifikasjonen

løser endogenitetsproblemet, men i teorien mener jeg at risikoen reduseres, og usikkerheten rundt kausalretningen blir mindre. Dette har også blitt gjort i flere tidligere analyser, noe som øker sammenlignbarheten mellom mine, og tidligere resultater.

6.8 Dynamisk modell?

Forut for analysen har jeg også undersøkt om modellen skal ha en dynamisk lag-struktur, hvor det kontrolleres for tidligere verdier av avhengig variabel. Dersom avhengig og uavhengige variabler i analysen påvirkes av tidligere økonomisk vekst er dette en kilde til endogenitet. En dynamisk spesifisering er et alternativ som ble vurdert ekstensivt i den innledende fasen av modellkonstruksjonen min.

Jeg gikk til sist bort fra å benytte en dynamisk lag-struktur. Dette gjøres av flere grunner. I innledende regresjoner undersøkte jeg lag-strukturen i panelet mitt med en Arellano-Bond-modell. Jeg fant ikke en signifikant sammenheng mellom tidligere, og nåværende BNP-vekstrate. Denne modellen er vedlagt i appendiks i tabell 9. Resultatet er overraskende siden tidligere undersøkelser har funnet en signifikant sammenheng. Dette kan skyldes særtrekk ved panelet mitt. Innledningsvis, i graf 2 viser jeg at den økonomiske vekst-trenden har vært fallende siden etterkrigstiden i OECD-landene. Man skulle derfor kanskje forvente en signifikant, negativ sammenheng mellom nåværende og forutgående vekst. I graf 4 i del 7 illustrerer jeg den samme veksttrenden for OECD i perioden 1980-2015. Det viser seg at veksttrenden nesten er utjevnet i dette tidsløpet. En grunn til dette kan være de store økonomiske fluktuasjonene i analyseperioden. Trenden er derfor ikke så entydig i mitt utvalg. Det virker mindre hensiktsmessig å benytte Arellano-Bond-estimatoren når det ikke er sikkert at en dynamisk modell er passende til panelet mitt.

Inklusjonen av lagget avhengig variabel bringer også med seg en rekke metodiske problemer og utfordringer for analysen, som instrumentprolifisering som kan føre til overfitting. Se Roodman (2009) for en mer detaljert gjennomgang av disse problemene, hvor Forbes (2000) blir brukt som et eksempel. Etter min mening er det en verdi i å ikke overkomplisere en statistisk analyse, så langt det lar seg gjøre. Jeg utelukker ikke at en

dynamisk lag-struktur kunne hatt fordeler i denne analysen, men jeg velger å ikke benytte en slik modell på grunn av usikkerheten rundt spørsmålet. Jeg gjør oppmerksom på at dette kan være en svakhet i analysen min.

6.9 Presentasjon av modell

På bakgrunn av variablene jeg har gjort rede for over så har jeg to grunnleggende modeller til analysen som formaliseres i dette delavsnittet.

Fasteffektmodell (FE):

$$\begin{aligned}
 BNPvekst_{it} - \overline{BNPvekst}_i &= \beta_1(Gini_{it} - \overline{Gini}_i) \\
 &+ \beta_2(SekundærUtdanning_{it} - \overline{SekundærUtdanning}_i) \\
 &+ \beta_3(TertiærUtdanning_{it} - \overline{TertiærUtdanning}_i) \\
 &+ \beta_4(PLI_{it} - \overline{PLI}_i) \\
 &+ \beta_5([Gini_{it} \times Dummy_i] - [\overline{Gini}_i \times Dummy_i]) \\
 &+ \beta_6(Handel_{it} - \overline{Handel}_i) \\
 &+ (u_{it} - \bar{u}_i)
 \end{aligned}$$

Førstedifferansemodell (FD):

$$\begin{aligned}
 \Delta BNPvekst_{it} &= \beta_1 \Delta Gini_{it} \\
 &+ \beta_2 \Delta SekundærUtdanning_{it} \\
 &+ \beta_3 \Delta TertiærUtdanning_{it} \\
 &+ \beta_4 \Delta PLI_{it} \\
 &+ \beta_5 \Delta (Gini_{it} \times Dummy) \\
 &+ \beta_6 \Delta Handel_{it} \\
 &+ \Delta u_{it}
 \end{aligned}$$

Dummyvariabelen representerer de østeuropeiske landene i OECD. Først utfører jeg en analyse uten denne variabelen og handelsvariabelen, ettersom modellen da er henholdsvis lik grunnmodellen til Forbes. Deretter utvider jeg analysen til å også inkludere interaksjonsvariabelen for Øst-Europa og kontrollen for internasjonal handel. Avhengig variabel,

BNP-vekst, er et femårig gjennomsnitt, og de uavhengige variablene er målt i starten av femårsperioden. Begge modellene har også dummyvariabler for hver tidsperiode.

7.1 Deskriptiv statistikk

Forut for den formelle analysen vil jeg gjennomgå litt deskriptiv statistikk om datasettet og de ulike variablene.

Gini-variabelen estimeres på bakgrunn av 100 forskjellige sett av imputasjoner, og jeg kommer ikke til å gjennomgå all variasjonen på tvers av disse, men i tabell 1 ser man deskriptiv statistikk for imputasjon 1 og imputasjon 100 ($m=1$ og $m=100$). Som man ser er variasjonen dem imellom henholdsvis liten, men fortsatt av betydning, som understreker den underliggende usikkerheten i SWIID-estimatene.

Tabell 1: Gini-data fra SWIID for imputasjon 1 og 100				
m=1				
Antall	Gjennomsnitt	Std. avvik	Min.	Maks.
312	30.36078	6.462311	18.23392	54.70512
m=100				
Antall	Gjennomsnitt	Std. avvik	Min.	Maks.
312	30.38316	6.414979	17.74519	51.64902

Tabell 1: Gini-data fra SWIID for imputasjon 1 og 100

Jeg gjør oppmerksom på at jeg til den deskriptive statistikken ikke har benyttet multipl imputasjon-kommandoene i Stata, da disse er mindre egnet til deskriptiv statistikk og til korrelasjonsmatriser. I korrelasjonsmatrisene i tabell 2 og 4 er det gjennomsnittet av de 100 imputasjonene som analyseres og diskuteres i relasjon til de andre variablene i analysen. Dette gir et godt estimat av størrelsen på sammenhengen mellom Gini-variabelen og andre data i panelet mitt, men gjør ikke rede for den underliggende variansen i SWIID sine ulikhetsdata. Derfor bør ikke signifikansnivået på korrelasjonen mellom SWIID sine Gini-data og andre variabler bli gitt for mye oppmerksomhet, da disse kan være upresise. Dette har ikke implikasjoner for tolkningen av korrelasjonene mellom andre variabler i matrisene, og har ingen

videre konsekvenser for den kommende regresjonsanalysen, hvor jeg benytter meg av multippel imputasjon-kommandoer.

For å belyse graden av sammenlignbarhet mellom SWIID-data og primærdata for ulikhet i datasettet mitt har jeg utført en korrelasjonsanalyse av SWIID sine Gini-data mot ulikhetsdataene fra OECD, Verdensbanken og Luxembourg Income Survey, vist i tabell 2. Det viser seg å være en svært høy grad av sammenlignbarhet mellom SWIID og de andre datakildene, og mellom de andre datakildene generelt. Korrelasjonskoeffisienten er på over 0,9 for alle de ulike Gini-dataene. Ettersom de forskjellige datakildene dekker ulike tidsspenn og utvalg av land mener jeg at den interne likheten mellom disse datakildene, og deres likhet til SWIID-data, underbygger gyldigheten av å anvende SWIID.

Tabell 2: Korrelasjonsmatrise mellom Gini-data fra SWIID, OECD, World Bank, LIS

	SWIID	OECD	VB	LIS
SWIID	1			
OECD	0.951*	1		
WB	0.936*	0.975*	1	
LIS	0.945*	0.992*	0.981*	1
Observasjoner:	312	73	83	71

* $p < 0.010$, Merk: SWIID-data i korrelasjonsmatrisen er gjennomsnittet av de 100 imputasjonene.

Tabell 2: Korrelasjonsmatrise mellom Gini-data fra SWIID, OECD, World Bank, LIS

I tabell 3 er deskriptiv statistikk for andre variabler i analysen. Deskriptiv statistikk fordelt på hver tidsperiode er også vedlagt i tabell 12 i appendiks.

Tabell 3: Deskriptiv statistikk for andre variabler i analysen							
Variabel	Definisjon	Datakilde	Antall	Gjennomsnitt	Std. avvik	Min.	Maks.
BNP-vekst	5-årig gjennomsnitt av årlig reell BNP-vekstrate	Penn World Table	336	2.838396	2.587617	-14.87353	12.84494
BNP per capita	BNP/befolkning	Penn World Table	336	24394.81	12650.94	1971.66	73262.68
Internasjonal handel	Import + Eksport prosent av BNP	Penn World Table	300	71.1789	42.69478	8.042179	326.1383
Tertiær-utdanning	Prosent av befolkning med tertiær utdanning	Barro-Lee	324	15.38945	9.618543	1.400103	53.94054
Sekundær-utdanning	Prosent av befolkning med sekundær utdanning	Barro-Lee	324	46.38362	15.7868	7.442836	88.9875
Prisnivå på investeringer	Prisnivå på kapitalformasjon/ investeringer relativt til USA	Penn World Table	336	.5760144	.2727228	.0739387	1.374739

Tabell 3: Deskriptiv statistikk for andre variabler i analysen

I tabell 4 presenterer jeg en korrelasjonsmatrise for variablene i panelet mitt. Jeg kommer ikke til å vie mye tid til resultatene i tabellen, men analysen kaster lys på flere forhold som er verd å merke seg forut for den formelle analysen. Mest interessant for denne oppgaven er den positive korrelasjonen mellom vekst og ulikhet. Tabellen viser og en negativ korrelasjon mellom ulikheten og BNP per capita, som antyder at landene med høyere velstandsnivå generelt har lavere ulikhet, ettersom store Gini-verdier betyr høy ulikhet. Videre vises en negativ korrelasjon mellom ulikhet og andelen med sekundær utdanning som er ti ganger større enn den observerte korrelasjon mellom tertiær utdanning og ulikhet. Prisnivå på investeringer, som er en kontroll for markedsvridninger, korrelerer negativt med ulikheten, altså antydes det at vridningene er større i land med lavere ulikhet. Det er også en negativ korrelasjon mellom handel og ulikhet som antyder, som tidligere antatt, at ulikheten er større i land med mer

internasjonal handel. Det kommer og til syne at det er en negativ korrelasjon mellom de østeuropeiske landene og graden av ulikhet. Det er også en negativ korrelasjon mellom BNP per capita og BNP vekstraten, som antyder at landene som allerede har et høyere velstandsnivå generelt opplever lavere vekst. Andelen med sekundær utdanning korrelerer, overraskende nok, negativt med BNP-veksten, mens handel ikke korrelerer signifikant med veksten i det hele tatt. Handel korrelerer derimot positivt med BNP per capita som indikerer at de rikere landene er mer globaliserte. Prisnivå på investeringer korrelerer også negativt med BNP-veksten og antyder dermed at markedsvridninger har en negativ sammenheng med den økonomiske veksten.

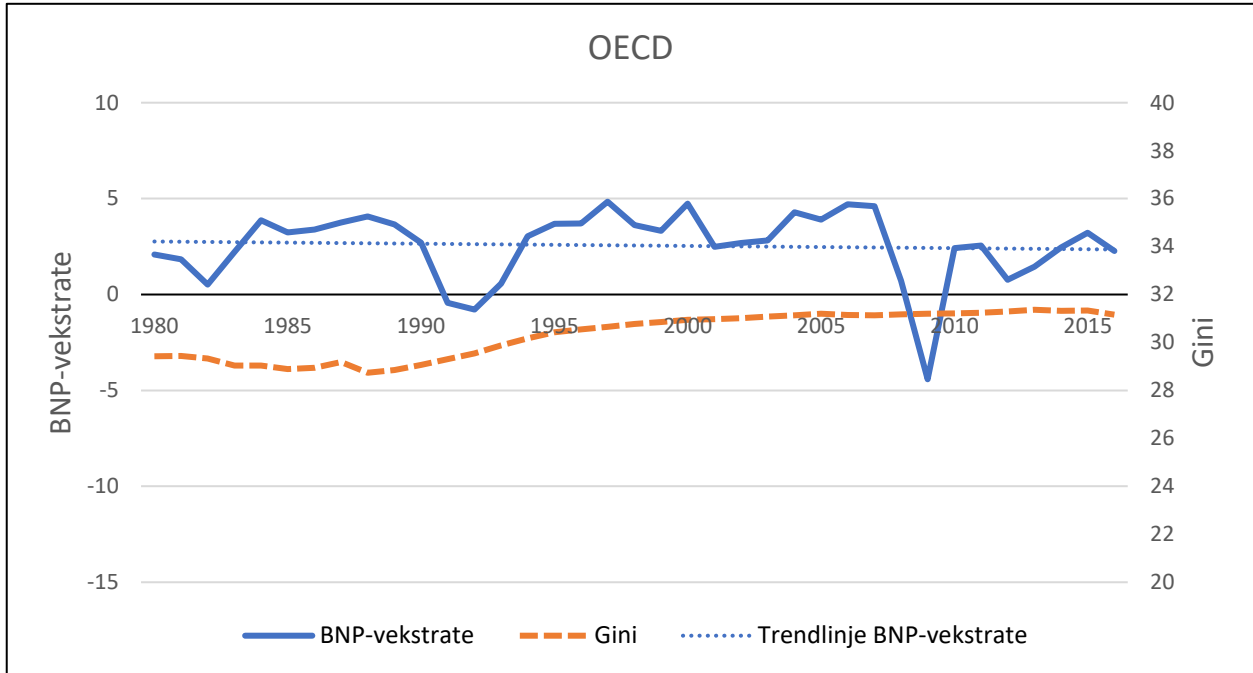
Tabell 4: Korrelasjonsmatrise for variabler i analysen

	BNP-vekst	Gini	BNP per capita	Ter. utd.	Sek. utd.	PI	Handel	Øst-Europa (dummy)
BNP-vekst	1							
Gini	0.176***	1						
BNPc	-0.255***	-0.264***	1					
Ter. utd.	-0.0973	-0.0292	0.648***	1				
Sek. utd.	-0.160***	-0.275***	0.377***	0.361***	1			
PI	-0.301***	-0.157***	0.767***	0.596***	0.439***	1		
Handel	-0.0513	-0.264***	0.347***	0.157***	0.390***	0.293***	1	
Ø. Europa	-0.167***	-0.153**	-0.299***	-0.129**	0.408***	-0.136**	0.283***	1

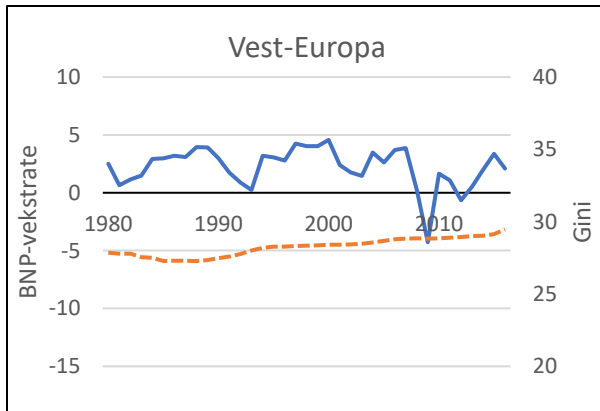
* p<0.10, ** p<0.05, *** p<0.010, Merk: SWIID-data i korrelasjonsmatrisen er gjennomsnittet av de 100 imputasjonene.

Tabell 4: Korrelasjonsmatrise for variabler i analysen

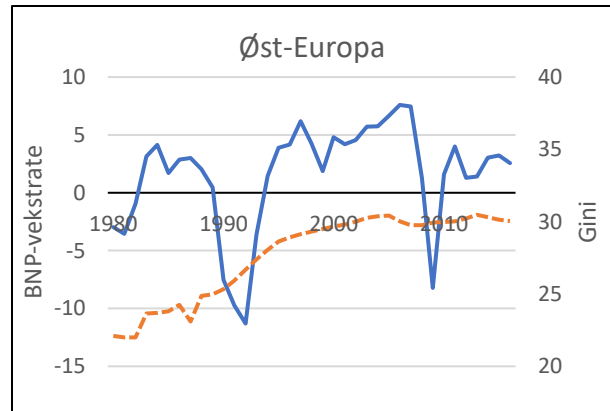
7.2 Grafisk oversikt



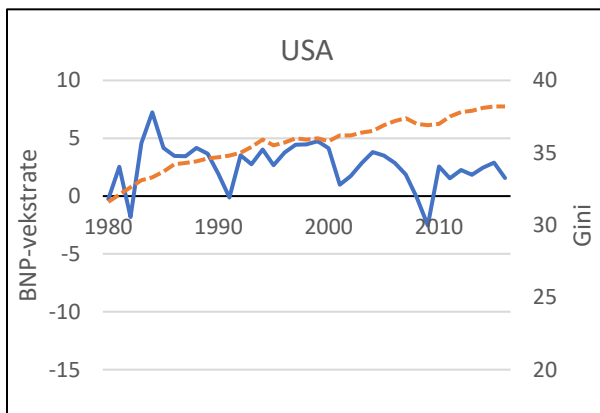
Graf 4: Gjennomsnittlig vekstrate og ulikhet i OECD 1980-2015



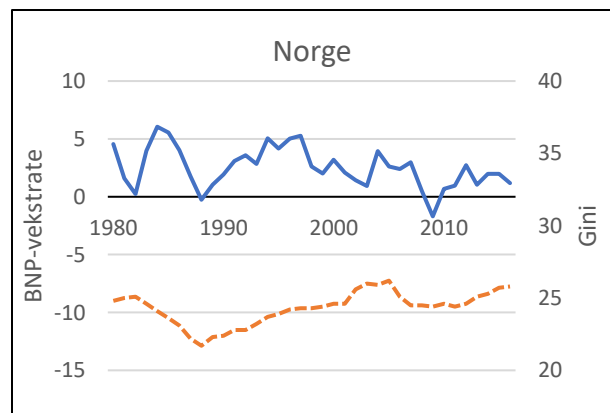
Graf 5: Vest-Europa



Graf 6: Øst-Europa



Graf 7: USA



Graf 8: Norge

I grafene 4, 5, 6, 7 og 8 fremstiller jeg årlig endring i økonomisk ulikhet og økonomisk vekst for hele OECD, og for flere undergrupper. Data som benyttes i grafene kommer fra panelet som jeg har satt sammen, og gir flere grunnleggende innsikter som er hensiktsmessige å bemerke forut for analysen. Først må det nevnes at for Øst-Europa er datatilgjengeligheten større frem til 90-tallet, og forut for dette er det i første rekke Polen og Ungarn som utgjør størstedelen av datagrunnlaget for graf 6.

For hele OECD og for alle undergruppene observeres det en økning i ulikheten over tid, muligvis med unntak av i Norge, hvor det observeres både opp- og nedganger i ulikheten, med liten forskjell i Gini fra starten til slutten av tidsintervallet. Den klareste økningen i ulikheten observeres i Øst-Europa, som går fra å ha minst økonomisk ulikhet i 1980, til et ulikhetsnivå på linje med Vest-Europa i 2015. USA har høyest ulikhet i starten av perioden, men også en betydelig økning i ulikheten deretter. Til sammenligning er den gjennomsnittlige økningen i Vest-Europa og OECD generelt betydelig mindre.

Veksttrenden er mer tvetydig. Trenden er svakt fallende, som nevnt i del 6. Lite overraskende ser man en klar nedgang i veksten sent på 2000-tallet i alle grafene, rundt Den store resesjonen. Det er også generelt økonomisk nedgang tidlig på 90-tallet, og sent på 80-tallet i Norge sitt tilfelle. Mest bemerkelsesverdig er vekstsituasjonen i Øst-Europa ved avviklingen av kommunisme på 90-tallet, hvor det først inntreffer en betraktelig økonomisk nedgang, og deretter høy, vedvarende økonomisk vekst som også sammenfaller med økningen i ulikheten. Deretter opplever de østeuropeiske landene en dypere økonomisk nedgang enn resten av OECD under finanskrisen.

8. Diskusjon av teorier for sammenhengen mellom ulikhet og vekst

Det er vanlig praksis i statistiske undersøkelser å presentere ulike hypoteser for hvordan uavhengig variabel påvirker avhengig variabel. Jeg kommer ikke til å gjøre dette formelt i denne oppgaven, ettersom formålet ikke er å avgjøre hva som forklarer en mulig årsakssammenheng mellom de to variablene, men heller kartlegge hvorvidt det er en sammenheng, og om den er positiv eller negativ. Denne undersøkelsen er derfor mer utforskende enn forklarende, og det ligger utenfor denne oppgaven sitt siktemål å avgjøre nøyaktig hva resultatene av analysen skyldes. Det er fremdeles av interesse å gjøre rede for forskjellige teorier som har blitt fremlagt i tidligere undersøkelser av økonomisk ulikhet og vekst.

Både Alesina og Rodrik (1994) og Persson og Tabellini (1994) fant en sammenheng mellom lavere ulikhet og høyere vekst. Begge artiklene foreslo at dette kan skyldes at større ulikhet vil kunne føre til flere krav om offentlig omfordelingspolitikk som kan virke veksthemmende. Barro (2000) legger også frem denne hypotesen. Det må merkes at disse undersøkelsene så på et større utvalg av land enn min analyse. Jeg stiller meg skeptisk til denne forklaringen i mitt panel. Selv om andre faktorer som historie og kultur kan være av betydning, er min vurdering at offentlig politikk er en av de største pådriverne for endringer i ulikheten i avanserte økonomier som OECD-landene. I tabell 10 i appendiks utfører jeg en korrelasjonsanalyse av ulikhet og størrelsen på offentlig sektor i OECD. Sammenhengen er stor og negativ. Dette antyder at OECD-land med lavere ulikhet generelt utfører mer omfordelingspolitikk enn dem med høy ulikhet, heller enn motsatt. Se også eksempelvis Alesina m. fl. (2018) hvor det fremkommer at vesteuropeere støtter redistribusjonell politikk i større grad enn amerikanere. Fra graf 5 og 7 ser man også at Vest-Europa har lavere økonomisk ulikhet enn USA.

Argumentet følger fremdeles av en av økonomifagets grunnteser – markedsforstyrrende inngripen kan virke veksthemmende. Det ovennevnte resonnementet kan i den sammenhengen bli snudd på hodet – reduksjon i ulikheten skjer gjennom veksthemmende økonomisk inngripen, slik at lavere ulikhet oppnås bekostning av veksten. Ideen er da at de landene som har avstått fra å gripe for mye inn i økonomien for å jevne ut ulikheten, har opplevd større

økonomisk vekst på sikt. Forbes (2000) antyder dette i sin analyse; Der hvor Alesina og Rodrik og Persson og Tabellini fant en sammenheng mellom lavere ulikhet og høyere vekst, fant hun det motsatte. Hun skriver i sin artikkel, basert på sine resultater, at det finnes en trade-off mellom vekst og ulikhet.

Dollar og Kraay (2001) utbroderer de forutgående poengene i sin analyse. De finner at økonomisk vekst i proporsjonel grad har tilfalt de fattigste i samfunnet, og at indikatorer på økt økonomisk vekst, som stor internasjonal handel og velutviklede finansmarkeder, ikke har noen innvirkning på inntektssammensetningen. Tilsvarende finner de liten sammenheng mellom offentlige tiltak som er ment til å utjevne ulikheten, og inntektssammensetningen i landene. Konklusjonen deres er at vekst er til gevinst for de fattige, mens offentlige velferdstiltak ikke har den ønskede utjevningseffekten.

Relatert til dette reflekterer Barro (2000) omkring at økt omfordeling fra rik til fattig kan redusere spareraten i samfunnet. Dette forklares ved at spareraten er høyere i det økonomiske toppsjiktet slik at økt omfordeling fra rik til fattig kan redusere samlet privat sparing. Dette er av betydning ettersom fremtredende økonomiske vekst-modeller, som Solow-modellen, vektlegger høy sparerate for langsiktig vekst. Tanken er at høyere sparing fører til flere vekstfremmende investeringer.

Et motstridende argument finnes i keynesiansk økonomisk teori. Keynes (1936 s. 93-103) kritiserer den klassiske terminologien hvor sparing tilsvarer investering, som deretter leder til økt økonomisk vekst. Keynesiansk analyse gir langt større oppmerksomhet til betydningen av konsum (131-147), som per definisjon er det motsatte av sparing. I Keynes sin teori regnes konsum for å være positivt for sysselsettingen og øker multiplikatoreffekten av økonomisk stimuli, med positive virkninger utover kilden til stimulansen. Innen disse termene vil økt konsum motivere til investeringer, med positive implikasjoner for fremtidig økonomisk vekst. På den måten regnes konsum for å være en kilde til økonomisk vekst i Keynesiansk teori, og ikke sparing som forstått i klassisk økonomisk teori. Joseph Stiglitz (2016) argumenterer i tråd med denne tradisjonen. Stiglitz skriver at fattige bruker en større andel av inntekten sin på

konsum enn hva rike gjør. En jevnere inntektsfordeling øker privat konsum, og derav aggregert etterspørsel, som virker stimulerende på økonomien.

Stiglitz diskuterer videre hvordan lavere ulikhet leder til høyere investeringer i humankapital som helse og utdanning, som også kan virke positivt på økonomisk vekst.

Perotti (1996) foreslår i sin analyse at økt økonomisk ulikhet kan virke veksthemmende gjennom større sosial uro i samfunnet. Stor ulikhet kan lede til mindre tillit til institusjoner og økt kriminalitet, «rent-seeking» og kupp, som vil virke negativt på veksten. Videre kan dette lede til usikkerhet som reduserer mengden av investeringer i økonomien. Noen av disse forholdene er mer aktuelle for økonomier i utvikling, men dette er en forklaring som også går igjen i andre artikler, som Barro (2000) sin undersøkelse.

Zak og Knack (2001), som undersøker implikasjonene av tillit i samfunnet, finner en sammenheng mellom høyere økonomisk ulikhet, og mindre tillit. Videre finner de en positiv sammenheng mellom tillit og økonomisk vekst. Det er derfor mulig at høyere ulikhet reduserer den økonomiske veksten, som følge av mindre tillit i samfunnet.

Piketty (2013) finner at ulikheten i verden er økende, og advarer om at dette kan forårsake økonomisk ustabilitet. Han skriver også at økt ulikhet forårsaker en større inntektskonsentrasjon i et fåtall sine hender, i økende grad avgjort av arv heller enn evne. Dette kan ha negative konsekvenser for den meritokratiske dynamikken i kjernen av kapitalistisk konkurranse (377-429). Om Piketty sin teori er korrekt, kan man kanskje forvente at land med stor økonomisk ulikhet har lavere økonomisk vekst.

9.1 Presentasjon av resultater

Alle resultatene fra regresjonene jeg har utført presenteres og diskuteres i de kommende avsnittene. Alle tabellene følger samme format med førstedifferanse i første kolonne, og fasteffekt i andre kolonne. *p=0,1 **p=0,05 ***p=0,01. Standardfeil er vedlagt i parenteser. Hver av analysene er utført med robuste standardfeil for å gjøre rede for mulig heteroskedastisitet.

I tabell 5 er resultatene fra den grunnleggende analysen med alle land for hele tidsspennet i datasettet. Modellen er estimert i tre omganger hvor flere variabler har blitt inkludert. I modell (1) er resultatene fra en grunnleggende modell, rimelig lik den man finner i Forbes sin analyse. I modell (2) har jeg inkludert interaksjonsleddet for Øst-Europa, som i Scholl og Klasen (2018) sitt opplegg. I modell (3) har jeg inkludert handelsvariabelen i modellen for å undersøke om dette er en hensiktsmessig kontroll i slike opplegg, noe som ikke har vært tilfelle i tidligere analyser.

	(1)		(2)		(3)	
	FD	FE	FD	FE	FD	FE
Gini	0.247** (0.109)	0.223** (0.101)	0.0712 (0.0833)	0.0564 (0.0800)	0.0712 (0.0833)	0.0555 (0.0803)
BNP capita	-0.000311*** (0.000107)	-0.000169** (0.0000706)	-0.000244** (0.0000998)	-0.000106* (0.0000585)	-0.000243** (0.0000998)	-0.000105* (0.0000583)
Sek. utd.	0.108*** (0.0320)	0.0380 (0.0250)	0.0704** (0.0350)	-0.00820 (0.0235)	0.0695* (0.0358)	-0.00727 (0.0238)
Ter. utd.	0.0866 (0.0661)	0.0624 (0.0695)	0.0587 (0.0623)	0.0173 (0.0639)	0.0574 (0.0595)	0.0194 (0.0638)
PI	-1.050 (1.337)	1.378 (1.919)	-1.845 (1.312)	0.158 (1.947)	-1.808 (1.518)	0.0690 (1.964)
Ø. Europa			0.828** (0.350)	0.826** (0.321)	0.831** (0.358)	0.837** (0.340)
Handel					0.00134 (0.0186)	-0.00267 (0.0131)
Konstant	-0.536 (0.483)	-1.374 (3.609)	-0.816*** (0.293)	1.239 (3.216)	-0.815*** (0.290)	1.307 (3.154)
N	239	275	239	275	239	275
F	7.52	11.73	11.00	9.64	10.25	8.94

Tabell 5: Førstedifferanse- og fasteffektmodeller for hele OECD

I dette opplegget har min antakelse vært at kausalretningen peker i retning av at det er ulikheten som påvirker veksten. Som diskutert tidligere har jeg gjort forsøk på å etablere kausalitet i modellspesifikasjonen min, men på tross av dette mener jeg at kausalretningen på sammenhengene må vurderes med varsomhet. Estimerte koeffisienter i analysen må derfor i første rekke forstås som korrelasjoner mellom variablene.

Den grunnleggende modellen (1) i tabell 5 viser en positiv og signifikant koeffisient for sammenhengen mellom ulikhet og økonomisk vekst. Gitt at modellen min er korrekt spesifisert er tolkningen av dette at landene med høyere inntektsulikhet i starten av femårsperiodene, tilsynelatende opplever større gjennomsnittlig økonomisk vekst i den kommende perioden.

Koeffisienten på ulikhetsvariabelen er svært lik i førstedifferansemodellen og fasteffektmodellen, i overkant av 0,2. Siden Gini-verdiene i datasettet mitt går fra 0-100 er tolkningen av dette at ett prosentpoeng økning i ulikheten, målt i Gini, gir en økning på 0,2% for den økonomiske veksten, ceteris paribus. Dette er etter min mening en rimelig stor koeffisient, ettersom tabell 3 viser at gjennomsnittlig vekst er på om lag 2,8%, og tatt i betraktning hvor liten endring i ulikheten det her er snakk om. En forholdsvis liten endring i graden av ulikhet kunne ifølge modellen gitt betraktelig høyere økonomisk vekst.

I modell (2) hvor det introduseres et interaksjonsledd for de østeuropeiske landene, er ikke lengre ulikhetsvariabelen signifikant, og størrelsen på koeffisienten reduseres drastisk. Interaksjonsleddet mellom ulikhet og dummyvariabelen for Øst-Europa er derimot i aller høyeste grad signifikant og positiv. Tolkningen min av modell (2) er at den observerte sammenhengen mellom ulikhet og vekst egentlig skyldes særtrekk ved de østeuropeiske landene, som diskutert tidligere. Dette resultatet fant også Scholl og Klasen (2018) i sin analyse. Basert på resultatene i tabellen virker det som om ikke det er en generell sammenheng mellom graden av ulikhet og den økonomiske veksten i OECD-landene, men at det er en spesifikk sammenheng mellom disse variablene for de østeuropeiske landene.

I modell (3) introduserer jeg en kontroll for internasjonal handel, som har veldig liten innvirkning på resultatene, og gir redusert F-verdi for modellen som helhet. Dette kan antyde at denne modellspesifikasjonen er svakere. Min umiddelbare tolkning er at handelsvariabelen ikke

tjener noen hensikt i analysen og den blir ikke inkludert i senere analyser i denne oppgaven. Etter min oppfatning er dette et interessant funn, ettersom min forventning var at graden av internasjonal handel har innvirkning på både graden av økonomisk vekst og ulikhet.

Selv om kontrollvariablene i analysen er av mindre interesse for denne oppgaven er det også verd å merke seg at BNP per capita er den eneste variabelen som er signifikant på tvers av modellspesifikasjonene, med negativt fortegn.

Andelen i befolkningen med sekundær utdanning er signifikant med positivt fortegn i førstedifferansemodellen, men ikke fasteffektmodellen, i modell (2).

I del 6 var min forhåndsantakelse at andelen med tertiær utdanning kanskje ville være en like god- eller bedre kontroll enn sekundær utdanning i en analyse av avanserte økonomier. Ifølge resultatene mine ser dette ikke ut til å stemme, og variabelen gir ikke signifikant utslag på økonomisk vekst i tabell 5.

Det er interessant at prisnivå på investeringer ser ut til å ha liten sammenheng med økonomisk vekst i noen av modellene i tabell 5, tatt i betraktning hvordan denne kontrollen har blitt anvendt mye i økonomiske vekstmodeller som et mål på markedsvidringer; Man skulle kanskje trodd at dette ville hatt negative implikasjoner for den økonomiske veksten.

9.2 Modell uten Øst-Europa

En grunnleggende antakelse for analysen min er at OECD-landene er en særlig sammenlignbar gruppe av land. Resultatene i tabell 5 antyder at Øst-Europa har gjennomgått en utvikling som er så forskjellig fra de resterende landene i OECD at sammenligningsgrunnlaget er svakere enn for det øvrige panelet.

Jeg underbygger dette poenget i en supplerende regresjon, vist i tabell 11 i appendiks. I den supplerende regresjonen introduserer jeg ytterligere interaksjonsvariabler for andre landgrupper som jeg mener kan være av interesse. Dette er henholdsvis Vest-Europa, Norden og «angliske» land, definert som USA, Canada, Australia og New Zealand. Dette er gjort for å undersøke om andre gruppe-spesifikke forhold er av tilsvarende betydning som

sammenhengen jeg finner for Øst-Europa. Den supplerende regresjonen gir ingen signifikante resultater for noen av disse andre kontrollene. Dette underbygger min antakelse om at Øst-Europa skiller seg betraktelig fra de øvrige landene i panelet.

På grunn av resultatene i tabell 5 og tabell 11 mener jeg at det er hensiktsmessig å utføre den grunnleggende regresjonsanalysen på ny, uten Øst-Europa, som vist i tabell 6. Som i de tidligere resultatene er koeffisienten for ulikhet ikke signifikant for utvalget utenfor Øst-Europa, selv om koeffisienten er positiv. Prisnivået på investeringer er signifikant med negativt fortegn i førstedifferansmodellen men ikke i fasteffektmodellen. Det er derfor en mulighet at denne variabelen er av betydning på generell basis, men i mindre grad for Øst-Europa, siden variabelen ikke var signifikant i tabell 5. Dette er av liten betydning for min analyse, men jeg mener at dette illustrerer hvorfor det er hensiktsmessig å utføre flere analyser med ulike landutvalg inkludert i panelet. Det er mulig at ulike forhold virker inn i de forskjellige landene, noe som krever ulike kontroller. En standard modellspesifikasjon er ikke nødvendigvis tilstrekkelig for å analysere statistiske sammenhenger på tvers av forskjellige utvalg.

Det er nevneverdig at BNP per capita, som er signifikant med negativt fortegn i alle de andre tabellene, ikke er signifikant for fasteffektmodellen uten Øst-Europa. En mulig forklaring er at panelet stort sett omfatter land med rimelig høyt velstandsnivå når ikke Øst-Europa er inkludert. Det er mulig at en regresjonsanalyse uten Øst-Europa har vanskeligere for å fange opp sammenhengen mellom BNP per capita og økonomisk vekst siden variasjonen på variabelen er mindre i utvalget. I denne oppgaven har jeg argumentert for hvorfor det er gunstig for analysen å inkludere mest mulig like land, men dette eksempelet illustrerer kanskje også det motsatte; Det kan være til fordel å analysere ulike land for å fange opp større variasjon i forskjellige variabler.

Tabell 6: Førstedifferanse- og fasteffektmodeller uten Øst-Europa		
	FD	FE
Gini	0.0882 (0.0792)	0.0808 (0.0789)
BNP capita	-0.000114*** (0.0000385)	-0.0000401 (0.0000322)
Sek. utd.	0.0375* (0.0211)	-0.00307 (0.0139)
Ter. utd.	0.00155 (0.0456)	-0.0513 (0.0527)
PI	-3.210*** (1.231)	-1.915 (1.643)
Konstant	-0.970*** (0.220)	4.101 (2.965)
<i>N</i>	202	230
<i>F</i>	15.44	12.88

Tabell 6: Førstedifferanse- og fasteffektmodeller uten Øst-Europa

9.3 Modell med forskjellige tidsperioder

Den siste tabellen, 7 går over to ulike tidsspenn. Modell (1) går fra de første observasjonene, til 2000, og modell (2) går fra 1990-2015. Disse regresjonene har færre tidsobservasjoner, som kan skape utfordringer for estimeringsprosessen. På samme tid er det slik at panelet er mer balansert etter 1990-tallet som jeg diskuterte tidligere i analysen, som også kan være til fordel for treffsikkerheten til estimatene.

I (1) er den mest betydelige forskjellen fra tidligere at koeffisientene endrer størrelse og signifikans for den østeuropeiske variabelen. I tabell 5 var koeffisientene nesten identiske på tvers av FD- og FE-modellen. For tidsperioden frem til 2000 er koeffisienten større, og signifikant i FD-modellen, men mindre og insignifikant i FE-modellen. Jeg kan foreslå ulike mulige grunner til dette. I dette underutvalget er det få tidsobservasjoner for Øst-Europa; Data for mange av disse landene foreligger ikke før 1990. Dette kan ha implikasjoner for estimerte sammenhenger. Graf 6 antyder at Øst-Europa opplevde høyere vekst på 2000-tallet enn på 90-

tallet. Siden denne modellen avslutter i 1999 kan det hende at en statistisk sammenheng mellom høyere ulikhet og høyere vekst i mindre grad observeres for Øst-Europa.

Der hvor de fleste variablene blir mindre signifikante sammenlignet med tabell 5 blir variabelen for tertiær utdanning signifikant for denne perioden. Uten å gi dette en spesifikk tolkning, så illustrerer det mitt tidligere poeng om hvor sårbare resultatene i en regresjonsanalyse kan være for endringer i utvalget.

Modell (2) går fra 1990 til 2015 og igjen er den mest nevneverdige forskjellen at koeffisienten for interaksjonsleddet for Øst-Europa endrer seg. I denne modellen øker koeffisienten i fasteffekt-modellen og den reduseres i førstedifferansemodellen, men er signifikant i begge tilfeller. Den større signifikansen kan skyldes at antallet observasjoner er større for Øst-Europa i dette datasettet, og panelet er mer balansert på tvers av land for dette underutvalget.

Tabell 7: Førstedifferanse- og fasteffektmodeller, År>2000 og 1990-2015				
	(1)		(2)	
	FD	FE	FD	FE
Gini	0.0578 (0.124)	0.0588 (0.0769)	0.0299 (0.117)	-0.0109 (0.110)
BNP capita	-0.000487*** (0.000176)	-0.000464*** (0.000158)	-0.000248** (0.000108)	-0.000125* (0.0000694)
Sek. utd.	0.0620 (0.0642)	-0.0195 (0.0339)	0.0827** (0.0413)	0.0423 (0.0408)
Ter. utd.	0.148 (0.133)	0.252** (0.103)	0.0713 (0.0718)	0.0184 (0.0808)
PI	-3.137 (2.455)	1.014 (2.391)	-1.936 (1.406)	-2.382 (1.742)
Ø. Europa	1.137* (0.584)	0.742 (0.572)	0.882** (0.367)	1.137*** (0.313)
Konstant	-0.125 (0.570)	4.709 (3.890)	0.415 (0.450)	-1.740 (4.922)
<i>N</i>	131	167	171	179
<i>F</i>	10.20	9.65	6.12	8.05

Tabell 7: Førstedifferanse- og fasteffektmodeller, 1980-2000 og 1990-2015

9.4 Sammenligning med tidligere forskning

Denne analysen har hatt til hensikt å bidra til arbeidet som har blitt gjort for å undersøke den statistiske sammenhengen mellom økonomisk ulikhet og økonomisk vekst. Jeg har satt sammen en modell med variabler som bygger på tidligere analyser, slik at resultatene bedre kan sammenlignes. Det må nevnes at modellen min skiller seg i forskjellig grad fra alle tidligere undersøkelser, men et sentralt spørsmål i denne oppgaven er om tidligere resultater står seg i lys av nye data.

Alesina og Rodrik (1994) og Persson og Tabellini (1994) sine undersøkelser skapte stor debatt om sammenhengen mellom ulikhet og økonomisk vekst. De fant at lavere ulikhet var positivt for veksten. En lignende sammenheng observeres ikke i noen av mine modellspesifikasjoner. Jeg har diskutert hvorfor jeg mener at disse studiene har metodiske svakheter. Tverrsnittsanalysene deres reiser spørsmål om endogenitet. Resultatene i analysene kan skyldes uobserverte, landspesifikke forhold som korrelerer med lavere ulikhet, og høyere vekst. I begge undersøkelsene diskuteres det at land med høyere ulikhet opplever større krav om veksthemmende omfordelingspolitikk. Mine empiriske funn lar seg vanskelig forene med denne forklaringen.

Min undersøkelse har kanskje flest likhetstrekk med Forbes (2000) sin undersøkelse. En vesentlig forskjell er at jeg ikke identifiserte en dynamisk lag-struktur, og utførte derfor en førstedifferanse- og fasteffektanalyse. Jeg benytter meg ikke av en Arellano-Bond-estimatoren slik som Forbes gjør. Jeg har anvendt et annet datasett for ulikhet enn tidligere analyser, og litt forskjellige kontroller, som diskutert i del 6. Land-utvalget til panelet og lengden på tidsserien er også forskjellig. Uansett er resultatene i den grunnleggende modellen i tabell 5 sammenlignbare med dem som Forbes finner; Der er det en, signifikant positiv sammenheng mellom høyere ulikhet og graden av økonomisk vekst. Som i Scholl og Klasen (2018) sin analyse finner jeg fortsatt at denne sammenhengen tilsynelatende kun skyldes særegenheter i Øst-Europa. I panelet mitt som ender før år 2000 er ikke interaksjonsvariabelen for Øst-Europa

signifikant på tvers av modellspesifikasjoner. Siden Forbes utførte sin analyse dette året, kan dette forklare hvorfor hun ikke identifiserer en lignende sammenheng i sitt opplegg.

Barro (2000) og Brueckner og Lederman (2018) fant at det var en signifikant sammenheng mellom høyere ulikhet og økonomisk vekst i høyinntektsland. Siden OECD består av avanserte økonomier med høyt velstandsnivå er det interessant at jeg ikke finner en lignende sammenheng i mitt opplegg unntatt i Øst-Europa.

Kuznets (1955) sin hypotese var at ulikheten er fallende for avanserte økonomier med høyt velstandsnivå. I tabell 4 observeres en korrelasjon mellom lavere ulikhet og høyere BNP per innbygger, i overensstemmelse med hypotesen. Ingen videre empiriske eller statistiske funn i analysen min underbygger den generelle gyldigheten av Kuznets-kurven. Mine data antyder at ulikheten har vært voksende i nyere tid for hele OECD.

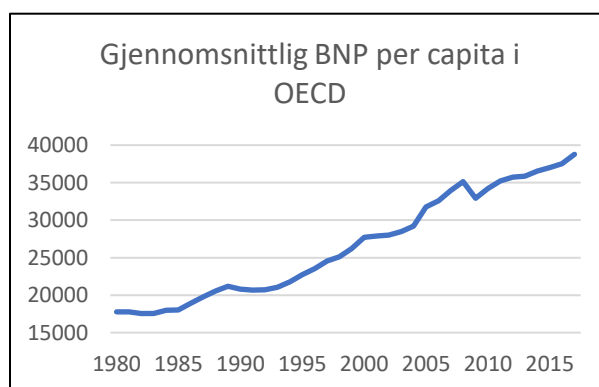
De resulterende koeffisientene for noen variabler i analysen min endrer seg for ulike modellspesifikasjoner, som forskjellige land-utvalg, tidsperioder og kontroller. Jeg mener at dette er av interesse for den øvrige litteraturen som har blitt diskutert i denne oppgaven. Ulike modellspesifikasjoner har blitt benyttet i mange av disse analysene, med tilsvarende variasjon i resultatene som fremkommer. Det er lærdom å hente fra dette; Ulike resultater i tidligere undersøkelser kan muligens skyldes metodiske forskjeller. Jeg mener at dette kan så tvil om den eksterne gyldigheten av resultatene fra forskjellige forskningsopplegg. Derfor er mitt synspunkt at det grunnleggende spørsmålet om sammenhengen mellom ulikhet og vekst fortsatt er preget av stor usikkerhet, på tross av de mange analysene som har blitt utført.

9.5 Diskusjon av resultatene

Resultatene mine må tolkes med forbehold om underliggende utfordringer ved etablering av kausalitet og endogenitetsproblemer. Sett at resultatene mine er pålitelige, er funnet mitt i første rekke at graden av ulikhet i OECD-landene generelt ikke har en hverken positiv eller negativ sammenheng med økonomisk vekst. Unntaket er Øst-Europa, som ser ut til å ha opplevd betydelig økonomisk fremgang samtidig som ulikheten har økt. Dette er formodentlig

på grunn av særtrekk ved disse landene, som skyldes overgangen fra kommunistisk planøkonomi til vestlig kapitalisme. Koeffisienten for den østeuropeiske variabelen er forholdsvis stor i alle modellene, og på om lag 0,8 i den grunnleggende FD- og FE-modellen i tabell 5. Resultatet impliserer kanskje at «noe» ulikhet må aksepteres for å oppnå økonomisk vekst, dersom alternativet er sovjetisk planøkonomi. Resultatet har tvilsom ekstern gyldighet, og kan for eksempel utelukkende skyldes særtrekk ved sovjetisk kommunisme, heller enn graden av økonomisk ulikhet generelt. Jeg henviser da til graf 8 som viser at Norge har hatt veldig lav økonomisk ulikhet relativt til gjennomsnittet i OECD, uten de samme negative konsekvensene som sovjetisk kommunisme medførte. Uten å diskutere dette for dyptgående kan jeg foreslå en mulig forklaring i lys av teoriene presentert i del 8. Zak og Knack (2001) diskuterer at lavere ulikhet kan være positivt for økonomisk vekst, fordi lavere ulikhet øker tilliten i samfunnet. De diskuterer Skandinavia i sin undersøkelse, og finner at disse landene har særlig høy tillit. Det er mulig at ulikhetsutjevningen i Øst-Europa ikke hadde positive implikasjoner for tilliten i samfunnet, fordi omfordelingen fant sted i undertrykkende ettpartiregimer. Jeg nevner dette som et eksempel på hvordan forskjellige politiske forhold kan tenkes å ha innvirkning på sammenhengen som analyseres.

Innledningsvis i del 2 viser jeg at ulikheten har økt i Vesten i moderne tid. Samtidig er den økonomiske veksten lavere enn i etterkrigstiden. Et spørsmål jeg reiste er om disse trendene henger sammen. Mine resultater tatt i betraktning, så ser ikke dette ut til å være tilfelle. BNP per capita er den eneste variabelen i analysen min som stort sett er signifikant på tvers av modellspesifikasjoner. Dette resultatet kan tilby en mulig forklaring på den fallende veksttrenden – selv om trenden har vært negativ i moderne tid, har BNP per capita i OECD vært entydig økende, vist i graf 9. Siden land med høyere BNP per capita ser ut til å oppleve mindre økonomisk vekst, er dette en mulig forklaring på den negative veksttrenden fra graf 2.



Graf 9: Gjennomsnittlig BNP per Capita i OECD 1980-2015

Forut for analysen antok jeg at flere av variablene i panelet mitt, som ulikhet, utdanningsnivå og internasjonal handel, ville ha innvirkning på økonomisk vekst. Sett at resultatene mine er til å stole på, så ser dette kanskje ikke ut til å være tilfelle. Dette resultatet kan implisere at ulike politiske vekststrategier har mindre betydning for økonomisk vekst i avanserte økonomier enn man kanskje skulle ønske. Velstandsnivået i landene er tilsynelatende den mest fremtredende faktoren som henger sammen med veksten i panelet mitt. Det kan derfor hende at avanserte økonomier simpelthen må «ta til takke» med lavere økonomisk vekst enn i tidligere tidsperioder.

Det grunnleggende funnet mitt er at graden av økonomisk ulikhet ikke har en sammenheng med graden av økonomisk vekst. Det er vanskelig å si hvilke, om noen, praktiske implikasjoner dette resultatet burde ha. Funnet må nødvendigvis tolkes med forbehold om at kausalitet ikke nødvendigvis er tilstrekkelig etablert. Det kan også være risiko for andre kilder til endogenitet i analysen. Dersom modellen er korrekt spesifisert er en mulig følge av resultatet mitt at ulike land står fritt til å prioritere ulikhetsutjevning i forskjellig grad, tuftet på egne ønsker og preferanser. Jeg understreker at analysen min kun har sett på økonomiske konsekvenser av ulikhet. Funnene mine utelukker ikke at ulikhet kan ha andre implikasjoner for et samfunn. Siden jeg ikke finner statistisk belegg for å si at økonomisk ulikhet påvirker økonomisk vekst hverken positivt eller negativt, så mener jeg at normative spørsmål vedrørende økonomisk ulikhet fortjener større plass i den offentlige diskusjonen av temaet. Tematikken bør særlig bli gitt mer oppmerksomhet fremover, hvis ulikheten fortsetter å vokse.

9.6 Veien videre

I denne oppgaven har jeg diskutert utfordringene ved mitt eget opplegg, og forskningsfeltet generelt. Mitt synspunkt er at noen problemer er uunngåelige i et opplegg som dette, og jeg mener at ingen enkelt analyse kan avgjøre et så omfattende spørsmål som implikasjonene av økonomisk ulikhet på økonomisk vekst. Min oppfordring er at forskningen innen denne delen av fagfeltet stadig videreføres fordi kunnskapsakkumulasjon vil gi oss bredere, dypere og kvalitativt bedre forståelse for problematikken omkring temaet. I lys av mitt eget opplegg, og

begrensningene ved dette, mener jeg at det er hensiktsmessig å formulere hva jeg mener ytterligere kan forbedre statistiske analyser av økonomisk vekst og økonomisk ulikhet i fremtiden.

Min vurdering er at datatilgjengelighet er en av de største utfordringene ved forskning på ulikhet. I fremtiden oppfordrer jeg til mer ekstensiv, årlig innsamling av data om ulikhet for flest mulig land, med uniform metodikk, etter Luxembourg Income Survey sitt eksempel. På samme tid applauderer jeg arbeidet som har blitt gjort med SWIID for å harmonisere ulikhetsdata. For eldre tidsperioder mener jeg at denne innfallsvinkelen har stort potensiale for å øke datamengden og sammenlignbarheten. Det lar seg ikke nødvendigvis gjøre å samle nye rådata av LIS sin kvalitet langt tilbake i tid.

Jeg mener at lengre datasett med flere land og større sammenlignbarhet kan ha store positive implikasjoner for fremtidige analyser. Et lengre panel kan gjøre det mulig å utføre mer langsiktige analyser enn i mitt opplegg. Jeg har analysert data over femårige intervaller, i en spesifikk periode med mange særtrekk, som avviklingen av planøkonomi i Øst-Europa på 1990-tallet og Den store resesjonen på 2000-tallet. Slike forhold kan medføre at observasjonene i panelet mitt avviker fra den generelle økonomiske trenden. For eksempel Piketty sin analyse av ulikhet, og teoriene han formulerer, omhandler utviklingen av ulikhet over et tidsspenn på flere hundre år. Implikasjonene av endringer i ulikheten kan kanskje ikke analyseres tilstrekkelig i et relativt kort panel som mitt.

Jeg mener også at det kan være av stor verdi å undersøke faktorene som forårsaker ulikhet, og forskjellige politiske strategier for økonomisk utjevning. I forrige delavsnitt sammenligner jeg Norge og Øst-Europa, hvor det ser ut til at Norge har lyktes i å jevne ut ulikheten uten de negative konsekvensene som sovjetisk kommunisme forårsaket. En antakelse i mitt opplegg er at OECD-landene er rimelig sammenlignbare, men ulike politiske strategier for reduksjon av ulikhet kan ha forskjellige implikasjoner for den kommende økonomiske veksten. Mer dyptgående analyser av faktorene som påvirker graden av ulikhet i forskjellige land kan også bidra til å berike senere analyser av typen som jeg har utført. For eksempel kan slike

undersøkelser bidra til å finne nye kontrollvariabler eller instrumenter som forbedrer estimatorene og reduserer endogenitetsproblemet.

10. Konklusjon

I denne oppgaven har jeg tatt for meg problemstillingen «Er det en sammenheng mellom ulikhet og økonomisk vekst i OECD-landene?». Jeg har undersøkt dette i en panelanalyse med en fasteffekt- og førstedifferansemodell. Tidsperioden for analysen går i første rekke fra 1980 til 2015. Undersøkelsen min bygger på tidligere forskning på temaet, og resultatene diskuteres i lys av den øvrige litteraturen. Den nyeste versjonen av Standardized World Income Inequality Database fra 2019 anvendes som datakilde for ulikhet i panelet mitt. Jeg analyserer et mindre utvalg av land enn mange tidligere undersøkelser, men argumenterer for fordelene ved dette utvalget. Påliteligheten til SWIID-datasettet er større for de 36 landene i OECD, og den interne sammenlignbarheten mellom landene er bedre enn for mange andre land-utvalg. Dette har positive konsekvenser for analysen sin treffsikkerhet, og reduserer risikoen for endogenitetsproblemer. Flere andre modellspesifikasjoner utføres også for å redusere denne risikoen, men endogenitet kan fortsatt ikke utelukkes i analysen. Jeg understreker derfor at resultatene mine bør tolkes med varsomhet.

I lys av utfordringene med denne analysen, foreslår jeg flere bidrag til forskningsfeltet som kan forbedre fremtidige undersøkelser.

Hovedresultatet mitt er at det ikke observeres en generell sammenheng mellom økonomisk ulikhet og økonomisk vekst i OECD-landene. Jeg finner samtidig at det er en sammenheng mellom større ulikhet og høyere vekst i Øst-Europa. Ingen lignende sammenheng observeres for andre grupper i panelet. Dette resultatet diskuteres ekstensivt og jeg argumenterer for hvorfor funnet har liten ekstern sammenlignbarhet. Overgangen fra sovjetisk planøkonomi til vestlig kapitalisme er vesentlig forskjellig fra historiske erfaringer i andre OECD-land.

Jeg undersøker også hvorvidt internasjonal handel er en hensiktsmessig kontroll i analyser av ulikhet og vekst, men resultatene mine antyder at dette ikke er tilfelle.

Forutsatt at funnene mine er pålitelige, antydes det at økonomisk ulikhet ikke har hverken positive eller negative konsekvenser for økonomisk vekst. Jeg mener uansett at ulikhet er et samfunnsspørsmål med stor betydning, som ikke vil bli mindre aktuelt i årene foran oss. En mulig tolkning av resultatene mine er at ulikhet også bør undersøkes og diskuteres i lys av andre faktorer enn økonomisk vekst.

Et sentralt spørsmål i denne oppgaven har vært hvorvidt tidligere resultater står seg i lys av nye data. Mine funn skiller seg fra mange tidligere analyser som har hatt stor innflytelse på forskningsfeltet. Persson og Tabellini (1994), og Alesina og Rodrik (1994) fant for 25 år siden at høy ulikhet er negativt for økonomisk vekst. Ingen av mine modellspesifikasjoner antyder at dette er tilfelle. Tilsvarende bestrider mine resultater funnene til Forbes (2000). Hennes konklusjon var at det finnes en «trade-off» mellom ulikhet og økonomisk vekst. Mine analyseresultater stemmer overens med funnene til Scholl og Klasen (2018). De observerer også en særegen sammenheng mellom ulikhet og vekst i Øst-Europa.

Jeg utfører analysen med ulike land-utvalg og tidsperioder. Det observeres variasjon i resultatene mine som følge av dette. Av denne grunnen foreslår jeg at forskjellig metodikk og datapaneler kan være årsaken til ulike funn i tidligere undersøkelser. Derfor mener jeg at det fortsatt er stor usikkerhet vedrørende sammenhengen mellom ulikhet og vekst. Dette er en hovedårsak til at jeg mener at økonomisk ulikhet bør bli gitt stor oppmerksomhet i økonomiske analyser i tiden fremover.

Kilder

- Alesina, A. og Rodrik, D. (1994). Distributive Politics and Economic Growth. *The Quarterly Journal of Economics*, 109(2), s. 465-490.
- Alesina, A., Stantcheva S. og Teso E. (2018). Intergenerational Mobility and Preferences for Redistribution. *American Economic Review*, 108(2), s. 521-554.
- Arellano, M. og Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58, issue 2, s. 277-297.
- Atkinson, A. B. og Brandolini, A. (2001). Promise and pitfalls in the use of 'secondary' data-sets: income inequality in OECD countries as a case study. *Journal of Economic Literature*, 39(3), s. 771-799.
- Atkinson, A. B., Piketty, T. og Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature, American Economic Association*, 49(1), s. 3-71.
- Barro, R. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5(1), s. 5-32.
- Barro, R. og Lee, J. W. (2013). A New Data Set of Educational Attainment in the World, 1950-2010. *Journal of Development Economics*, 104, s. 184-198.
- Brueckner, M. og Lederman, D. (2018). Inequality and economic growth: the role of initial income. *Journal of Economic Growth*, 23(3), s. 341-366.
- Castelló-Climent, A. (2010). Inequality and growth in advanced economies: An empirical investigation. *The Journal of Economic Inequality*, 8(3), s. 293-321.
- Deininger, K. og Squire, L. (1996). A new data set measuring income inequality. *The World Bank economic review*. 10(3), s. 565-591.
- Dollar, D. og Kraay, A. (2002). Growth is good for the poor. *Journal of Economic Growth*, 7, s. 195.
- Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics*, 84(2), s. 755-776.

Feenstra, R. C., Inklaar R. og Timmer M. P. (2015), The Next Generation of the Penn World Table *American Economic Review*, 105(10), 3150-3182, Tilgjengelig på: www.ggd.net/pwt [hentet 01.05.19].

Forbes, K. J. (2000). A Reassessment of the Relationship between Inequality and Growth. *American Economic Review*, 90 (4), s. 869-887.

Gini, C. (1912). Variabilità e Mutuabilità. Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche. C. Cuppini, Bologna.

Halter, D. M. Oechslin og J. Zweimuller (2014). Inequality and Growth: The Neglected Time Dimension. *Journal of Economic Growth*, 19, s. 81-104.

Islam, Nazrul, (1995), Growth Empirics: A Panel Data Approach, *The Quarterly Journal of Economics*, 110(4), s. 1127-1170.

Jenkins, S. P. (2015). World income inequality databases: An assessment of WIID and SWIID. *The journal of Economic Inequality*, 13(4), s. 629-671.

Keynes J. M. (1936) *Allmenn teori om sysselsetting, rente og penger*. Oversatt fra engelsk av R. L. Eliassen (2011), Oslo: Vidarforlaget AS

Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review*, 45(1), s. 1-28. Tilgjengelig på <http://www.jstor.org/stable/1811581> [hentet 01.05.19]

Perotti, R. (1996). Growth, income distribution, and democracy: What the data say. *Journal of Economic Growth*, 1(2), s. 149-187.

Persson, T. og Tabellini, G. (1994). Is Inequality Harmful for Growth? *The American Economic Review*, 84(3), s. 600-621

Piketty, T. (2013). *Capital in the Twenty-First Century*. London: Belknap.

Roodman, D. (2009). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), s. 135-158.

Rubin, D. B. (1987) *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons

Scholl, N. og Klasen, S. (2018). Re-estimating the relationship between inequality and growth. *Oxford Economic Papers*. s. 1-24. doi: 10.1093/oep/gpy059

Solt, F. (2019). Measuring Income Inequality Across Countries and Over Time: The Standardized World Income Inequality Database. SWIID Version 8.0, February 2019. [online] [hentet 01.05.19].

Stiglitz, J. E. (2016). 'Inequality and Economic Growth' i Jacobs, M. og Mazzucato, M. (red.) *Rethinking Capitalism: Economics and policy for sustainable and inclusive growth*. Chichester: Wiley-Blackwell, s. 134-155.

UN, Department of Economic and Social affairs (2018). *Note on Income Inequality Data*. [online] Tilgjengelig på: <https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2018/11/Note-Income-Inequality-Data.pdf> [hentet 30.05.19]

Veerbek, M. (2017). *A Guide to Modern Econometrics*. John Wiley & Sons, Inc., 5. utgave.

WID - World Inequality Database (2019). [online] Tilgjengelig på: <https://wid.world/> [hentet 13.05.19].

Wooldridge, J. M. (2009) *Introductory Econometrics: A modern approach*. CENGAGE Learning Custom Publishing, 5. Utgave.

Zak, P. J. og Knack, S. (2001) Trust and Growth. *The Economic Journal*, 111(470), s. 295-321.

Appendiks:

Tabell 8: Korellasjonstabell for menn og kvinner sitt utdanningsnivå	
Sekundærutdanning	Tertiærutdanning
0.916*	0.903*
<i>N</i>	288

* $p < 0.001$

Tabell 9: Initiell regresjon for å undersøke lag-struktur	
	Arellano-Bond
Lag1	0.0552 (0.043)
Gini	0.165** (0.073)
BNP capita	0.0569*** (0.000)
Sek. utd.	-0.0351 (0.019)
Ter. utd.	-1.009 (0.042)
PI	-0.000105*** (1.174)
Konstant	-0.257 (2.458)
<i>N</i>	232
<i>F</i>	13.12

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Tabell 10: Korrelasjon mellom offentlig sektor og ulikhet	
	Gini
Offentlig sektor	-0.402*
<i>N</i>	1538

* $p < 0.001$, merk: SWIID-data i korrelasjonsmatrisen er gjennomsnittet av de 100 imputasjonene.

Offentlig sektor er målt som offentlige utgifter som prosentandel av BNP i hvert OECD-land, hentet fra OECD sin statistiske database data.oecd.org

Tabell 11: Førstedifferanse- og fasteffektmodeller for hele OECD med ekstra interaksjonsvariabler		
	FD	FE
Gini	0.00898 (0.203)	-0.0220 (0.212)
BNP capita	0.0763** (0.000102)	0.00547 (0.0000587)
Sek. utd.	0.0680 (0.0359)	0.0309 (0.0277)
Ter. utd.	-1.750 (0.0640)	0.186 (0.0600)
PI	-0.000250** (1.338)	-0.000106* (1.971)
Ø. Europa	0.881** (0.389)	0.888** (0.361)
V. Europa	-0.0143 (0.230)	0.128 (0.229)
N. Europa	0.168 (0.180)	0.00694 (0.105)
Anglikansk	0.305 (0.245)	0.360 (0.234)
Konstant	-0.722* (0.051)	0.340 (0.917)
<i>N</i>	239	275
F	8.09	9.03

Tabell 12: Deskriptiv statistikk for hver tidsperiode

	Antall	Min.	Maks.	Gjennomsnitt	Standardavvik
1970-1975					
BNP-vekstrate	30	1.063514	10.06129	5.107138	1.970092
Gini	15	23.3	52.1	31.15333	7.992663
BNPc	30	1971.66	27383.45	13180.3	5673.227
Sek. utd.	36	7.442836	62.52378	32.04797	13.90629
Ter. utd.	36	1.409546	20.88762	6.829554	4.951208
PI	30	.0739387	.2430063	.1573722	.0345071
Handel	30	8.042179	157.0281	48.71569	30.31687
1975-1980					
BNP-vekstrate	30	-.670522	10.53129	3.437899	2.13475
Gini	19	21.5	50.4	30.59474	7.560973
BNPc	30	4006.849	27182.72	15152.23	5799.77
Sek. utd.	36	11.55778	65.45964	36.5617	13.84757
Ter. utd.	36	1.400103	26.67886	8.932146	5.882085
PI	30	.1014097	.4271853	.277803	.0685903
Handel	30	12.57631	171.64	54.28597	30.69216
1980-1985					
BNP-vekstrate	30	-1.906738	7.485846	2.10484	1.547458
Gini	24	21.1	48.1	29.42083	6.873957
BNPc	30	4830.754	30425.78	17778.4	6693.286
Sek. utd.	36	14.14104	69.9	40.70722	13.34128
Ter. utd.	36	2.710823	29.90905	10.56865	6.734586
PI	30	.184532	.7173054	.441591	.1083027
Handel	30	13.26121	168.6122	60.66462	31.00892
1985-1990					
BNP-vekstrate	30	.507862	10.07508	3.620573	1.874243
Gini	28	20.5	47.9	28.89643	6.551137
BNPc	30	5631.504	32498.3	18038.74	7025.488
Sek. utd.	36	8.181595	67.5421	43.31596	12.76061
Ter. utd.	36	4.332417	37.3579	12.77394	7.108579
PI	30	.1733437	.5466292	.3754941	.0742687
Handel	30	16.57349	201.3535	63.72789	36.66635
1990-1995					
BNP-vekstrate	36	-14.87353	8.478714	.6821522	4.832843
Gini	35	17.3	48	29.06	6.657822
BNPc	36	7207.909	36386.51	20813.89	7714.94
Sek. utd.	36	14.30053	73.5	46.31111	12.33271
Ter. utd.	36	5.093718	43.10262	14.57679	7.978482
PI	36	.213236	.9899698	.6054524	.1961325
Handel	36	19.66722	183.0739	64.57219	35.01353
1995-2000					
BNP-vekstrate	36	1.107574	9.370358	3.833069	1.508529
Gini	36	22	48.2	30.42222	6.27672

BNPc	36	8386.096	40390.87	22699.79	8923.634
Sek. utd.	36	18.78687	76.14561	50.71807	12.72648
Ter. utd.	36	7.648501	44.19622	17.52879	7.89789
PI	36	.3721584	1.185268	.6587635	.2030282
Handel	36	16.7497	184.9795	70.26411	34.85848
2000-2005					
BNP-vekstrate	36	1.023516	7.33682	3.405048	1.657767
Gini	36	22.6	48.5	30.94167	5.982898
BNPc	36	10430.43	57062.16	27728.19	11778.46
Sek. utd.	36	26.70465	88.9875	54.04389	14.18658
Ter. utd.	36	8.3	48.47354	19.36439	8.619049
PI	36	.3727391	.9979446	.5954129	.1457259
Handel	36	20.30605	269.1856	84.24511	45.45008
2005-2010					
BNP-vekstrate	36	-.42047	5.242132	1.904231	1.38977
Gini	36	23.8	46.7	31.20278	5.521775
BNPc	36	12930.96	66904.05	31754.84	13057.43
Sek. utd.	36	30.99609	87.96585	56.38807	14.29106
Ter. utd.	36	7.555349	49.48077	22.47432	8.939496
PI	36	.503284	1.087896	.795519	.1734432
Handel	36	25.50063	297.5617	88.57931	49.4645
2010-2015					
BNP-vekstrate	36	-4.882654	7.60976	1.928102	2.00159
Gini	36	24.6	46.2	31.21389	5.301275
BNPc	36	14507.01	73262.68	34208.15	13031.13
Sek. utd.	36	31.08584	84.33771	57.35856	13.94638
Ter. utd.	36	9.282008	53.94054	25.45645	9.584185
PI	36	.5757652	1.374739	.8805951	.2043905
Handel	36	28.18239	326.1383	96.00167	56.39183