

Dahle, Mari Færnes
Frantzen, Ingrid Louise
Kvale, David Kabichev
Vesterdal, Embla Sol Grimsrud

Kan familieytelser brukes som et verktøy mot avtagende fertilitet?

En empirisk studie av sammenhengen mellom familieytelser og fertilitet i OECD-land.

Bacheloroppgave i samfunnsøkonomi

Veileder: Nielsen, David Sørli

Mai 2023



Dahle, Mari Færsnes
Frantzen, Ingrid Louise
Kvale, David Kabichev
Vesterdal, Embla Sol Grimsrud

Kan familieytelser brukes som et verktøy mot avtagende fertilitet?

En empirisk studie av sammenhengen mellom familieytelser og fertilitet i OECD-land.

Bacheloroppgave i samfunnsøkonomi
Veileder: Nielsen, David Sørli
Mai 2023

Norges teknisk-naturvitenskapelige universitet
Fakultet for økonomi
Institutt for samfunnsøkonomi



Kunnskap for en bedre verden

Forord

Bacheloroppgaven er skrevet for Institutt for samfunnsøkonomi på NTNU, som en del av faget SØK2013. Oppgaven er skrevet av to studenter med studieretning økonomi og administrasjon med fordypning i samfunnsøkonomi, en student ved ingeniørvitenskap & IKT og en student ved bachelorprogrammet i samfunnsøkonomi.

Sammendrag

I denne oppgaven analyserer vi hvordan familieytelser påvirker fertiliteten i OECD-land. I mange av disse landene har det i senere år vært en avtagende trend i fertilitet, og formålet med oppgaven er derfor å undersøke om en økning i familieytelser vil kunne bidra til høyere fertilitet. For å besvare problemstillingen gjør vi flere regresjonsanalyser hvor vi undersøker den empiriske sammenhengen mellom fertilitet og familieytelser. Datasettet er hentet fra OECD sine databaser og går fra årene 1999 til 2019. Resultatene våre indikerer at en økning i familieytelser vil ha en positiv effekt på fertiliteten i et land. Den sterkeste sammenhengen finner vi mellom fertilitet og familieytelser i nordiske land. Når vi deler familieytelser inn i pengeoverføringer og tjenester, har pengeoverføringer, med og uten kontrollvariabler, en positiv effekt på fertilitetsraten mens resultatet varierer for tjenester. Om det er pengeoverføringer eller tjenester som gir størst effekt på fertiliteten avhenger også av antall kontrollvariabler, vi får ikke et entydig resultat.

Abstract

In this study, we investigate how family benefits affect the fertility rate in OECD countries. Many of these countries have observed a declining fertility trend in recent years, and the aim of the study is to examine whether an increase in family benefits can contribute to higher fertility rates. In the paper, we perform several regression analyses looking at the empirical relationship between fertility and family benefits. The dataset goes from the year 1999 to 2019 and the source is the OECD database. Our results indicate a positive relationship between total family benefits and fertility rates. We specifically see a strong positive relationship in Nordic countries. In the paper, we also divide family benefits into cash and in-kind benefits. For benefits in cash, we find a positive relationship with fertility both with and without control variables. For in-kind benefits, the result varies depending on the control variables. Whether cash or in-kind benefits have the largest effect on fertility also depends on the control variables we add.

Innhold

Liste over figurer	v
Liste over tabeller	v
1 Innledning	1
1.1 Bakgrunn	1
1.2 Problemstilling	2
1.3 Oppgavens struktur	2
2 Litteraturgjennomgang	4
2.1 Fertilitetsteori	4
2.2 Sentrale forskningsartikler	4
3 Teori	6
3.1 Enkel og multipl regresjonsanalyse	6
3.2 Forutsetninger for modellen	6
3.3 Estimering med faste effekter	7
3.4 Minste kvadraters metode	8
3.5 Hypotesetesting (T-test)	8
3.6 Determinasjonskoeffisient - R^2	9
3.7 Multikollinearitet	10
4 Metode	11
4.1 Datagrunnlag - fertilitet og familieytelser	11
4.1.1 Generelt	11
4.1.2 Fertilitet og familieytelser over tid	12
4.1.3 Forskjellige familieytelsers påvirkning på fertilitet	12
4.1.4 Geografiske forskjeller	13
4.1.5 Bruk av faste effekter	15
4.2 Kontrollvariabler	15

4.2.1	Generelt	15
4.2.2	Valg av kontrollvariabler	15
4.2.3	Korrelasjon mellom variabler	17
4.3	Beskrivelse av regresjonsligningene	17
4.3.1	Generelt	17
4.3.2	Regresjonsligningene	18
4.4	Deskriptiv statistikk	19
4.4.1	Deskriptiv statistikk - alle OECD-land	19
4.4.2	Deskriptiv statistikk - Norden	20
5	Resultater	21
5.1	Alle OECD-land	21
5.1.1	Determinasjonskoeffisient - R^2	23
5.2	Norden	23
5.2.1	Determinasjonskoeffisient - R^2	25
6	Diskusjon	26
6.1	Evaluering av forutsetningene	26
6.2	OECD-land	28
6.2.1	Effekt av totale ytelser	28
6.2.2	Effekt av <i>inkind</i> vs <i>incash</i> -ytelser	29
6.3	Nordiske land	30
7	Konklusjon	32
	Referanser	33
	Vedlegg	35
A	Vedlegg 1 - Oversikt over variabler	35
B	Vedlegg 2 - Oversikt over land	36
C	Vedlegg 3 - Oversikt over forkortelser	37

Liste over figurer

1	Fertilitet (samlet fruktbarhetstall) i Norge fra 2000 til 2022.	1
2	Familieytelser (målt i prosent av BNP) og fertilitet over tid (målt i barn per kvinne).	12
3	OECD - Neste års fertilitet og familieytelser	13
4	Skandinavia - Neste års fertilitet og familieytelser	13
5	Spredningsdiagram kontinenter	14
6	Spredningsdiagram nordiske land	14
7	Korrelasjonsmatrise - med data fra 2014	17
8	Deskriptiv statistikk alle OECD-land	19
9	Deskriptiv statistikk nordiske land	20
10	Resultater fra regresjonsanalyse alle OECD-land	21
11	Regresjonsanalyse nordiske land	24

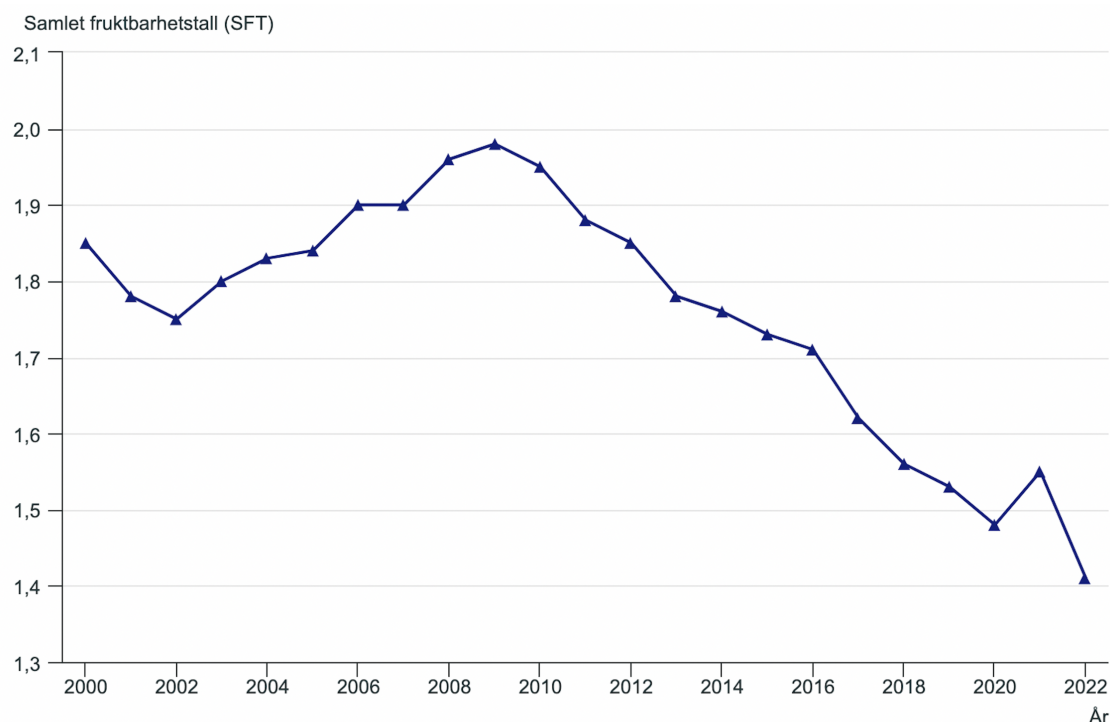
Liste over tabeller

1	Variabler med beskrivelse	35
2	Land med verdensdel	36
3	Forkortelser av variabler	37

1 Innledning

1.1 Bakgrunn

I sin nyttårstale ved inngangen til 2019 oppfordret daværende statsminister Erna Solberg til flere fødsler i Norge (Laland, 2019). Året etter, i 2020, ble fertilitetsraten imidlertid målt til rekordlave 1,48. Den avtagende trenden har vært gjeldende siden 2009, og har fortsatt i 2020 med en fertilitetsrate på 1,41 (Krokedal, 2023). Fertilitetsraten i et land er antall barn forventet født av en kvinne gitt at hun lever gjennom hele sin fruktbare periode (OECD, 2023a). Figuren under viser utviklingen med en stadig fallende fertilitetsrate. (Krokedal, 2023)



Figur 1: Fertilitet (samlet fruktbarhetstall) i Norge fra 2000 til 2022.

Bak uttalelsene til Solberg lå det en tydelig bekymring over utviklingen med fallende tall for fertilitet. Det at det stadig fødes færre barn i Norge kan være problematisk hvis man ønsker å opprettholde den norske velferdsmodellen på dagens nivå. Som Solberg påpekte; på sikt resulterer dette i at, relativt sett, færre unge skal bære en stadig tyngre velferdsstat på sine skuldre.

I utgangspunktet kan man se for seg at flere forhold knyttet til økonomi og politikk påvirker fertiliteten. Fra politisk hold er det mulig å prioritere økonomisk trygghet for barnefamilier gjennom å redusere kostnadene ved å ha barn. Av familiepolitiske tiltak regnes blant annet "betaling" til foreldre for å ta vare på egne barn, subsidiert barnepass, kontantoverføringer til foreldre, samt helsetjenester (Hart og Kravdal, 2020). I statsbudsjettet for 2023 har maksprisen i barnehager blitt redusert for alle

barn i hele landet. Videre tas det sikte på å gjøre barnehage gratis i Finnmark og deler av Troms fra august 2023 (Stortinget, 2022). Med dette som bakteppe later det til at regjeringen, uansett parti, støtter ideen om at lav fertilitet rettferdiggjør politiske tiltak.

Det er på ingen måte kun i Norge at det fødes færre barn enn før. I løpet av de siste tiårene har fruktbarheten sunket i store deler av verden. Fruktbarheten ligger under det såkalte reproduksjonsnivået i bortimot alle industrialiserte land (Rønsen, 2005). Reproduksjonsnivået er det fruktbarhetstallet som er nødvendig for å opprettholde et konstant folketall, dersom vi ser bort fra inn- og utvandring. Dette tallet er anslått til 2,1 (Rønsen, 2005). Synkende fødselstall har resultert i en mer pronatalistisk politikk. Dette fremkommer av FN-rapporten "Policy responses to low fertility. How effective are they?" (Sobotka mfl., 2019) Pronatalistisk politikk har som hovedformål å øke fødselstallene. Tilbake i 1986 oppga myndighetene i 19 land at de førte en slik politikk. I 2015 var det hele 55 land som rapporterte at de hadde som mål å øke fødselstallene – hvorav 27 europeiske land og 18 asiatiske land.

1.2 Problemstilling

Med dette som utgangspunkt, synes vi det er interessant og aktuelt å se nærmere på hvordan familieytelser påvirker fertiliteten i industrialiserte land. Denne oppgaven tar derfor for seg følgende problemstilling:

- **Vil en økning i familieytelser gi høyere fertilitet i OECD-land?**

Dette kan være en bred problemstilling, og i forsøket på å svare på spørsmålet er det mulig å ha flere innfallsvinkler. Siden bakgrunnen for oppgaven er et politisk fokus og interesse for fertilitetspolitikk i Norge, ønsker vi spesifikt å skille ut de nordiske landene for å se om resultatene blir det samme. Vi ønsker også å finne ut om det har noe å si om familieytelsene gis som pengestøtte eller i form av tjenester. For å spesifisere at dette er fokusområdene velger vi å avgrense problemstillingen med følgende spørsmål.

- Påvirker familieytelser fertiliteten i OECD-land?
- Påvirkes fertiliteten av hvilke typer familieytelser som gis?
- Er det noen forskjeller i de nordiske landene sammenlignet med alle OECD-landene?

1.3 Oppgavens struktur

Denne studien er skrevet på empirisk grunnlag med data fra OECD-land. Først blir eksisterende litteratur om fertilitet og familiepolitikk gjennomgått, etterfulgt

av en presentasjon av teorien brukt i analysen. Deretter beskrives metodene som er brukt i oppgaven, inkludert en beskrivelse av dataene og modellene som er brukt. Videre blir resultatene presentert, etterfulgt av diskusjon av disse funnene. Til slutt konkluderes artikkelen.

2 Litteraturgjennomgang

I litteratursøket har vi primært brukt databasene *Oria* og *Google Scholar*. I litteraturgjennomgangen presenteres først det teoretiske rammeverket med to sentrale teorier innenfor fertilitet og økonomisk utvikling. Deretter presenteres relevante studier, både med fokus på nordiske land og OECD-land, som ser på sammenhengen mellom familieytelser og fertilitet.

2.1 Fertilitetsteori

Fertilitet har i lang tid vært gjenstand for interesse i vitenskapen. Allerede på 1700-tallet skrev Thomas Malthus i "An Essay on the Principle of Poulation" at befolkningen vokser raskere enn matressursene og at konsekvensen vil bli matmangel og reduksjon i velstanden (Malthus, 1872). I nyere tid har denne tilnærmingen til fertilitet blitt sterkt utfordret. Teorien samsvarer ikke med den demografiske utviklingen i Europa. Dette har gjort at andre teoretiske tilnærminger har tatt over.

Gary Becker har i senere år vært en sentral bidragsyter på dette området, og han har lagt mye av grunnlaget for fertilitetsteori innenfor samfunnsøkonomien. I «A Treasie on the Family» (Becker, 1991) studerer han sammenhengen mellom inntekt og antall barn. Becker betrakter barn som et normale gode, som kan analyseres på samme måte som forbruksvarer, hvor økt konsum gir økt nytte. Basert på denne antagelsen vil økt inntekt isolert sett gi økt etterspørsel etter barn.

Becker fortsetter analysen med å forklare at vi ikke kun kan se på kvantiteten av barn. På samme måte som flere barn vil gi økt nytte, vil også høyere 'kvalitet' på barna gi høyere nytte. Et barn vil for eksempel få høyere 'kvalitet' hvis foreldre investerer i barnets utdanning. Det å øke 'kvaliteten' på et barn innebærer økte kostnader, og til en gitt inntekt vil en investering i 'kvaliteten' per barn gjøre at man total sett må få færre barn. Det er derfor en negativ korrelasjon mellom kvalitet og kvantitet av barn. Dette kan være med på å forklare en utvikling i industrialiserte land hvor fertilitetsraten har blitt lavere. Inntektsøkning og økning i velstand trenger ikke å være ensbetydende med et ønske om flere barn, men heller et ønske om å investere mer i de enkelte barna.

2.2 Sentrale forskningsartikler

Den negative fertilitetstrenden vi har observert de siste tiårene har vekket en økende interesse blant en rekke forskere, og mye av litteraturen undersøker på hvilke måter familiepolitiske tiltak påvirker fertilitetsatferd.

Rønsen og Skrede har ved flere anledninger undersøkt fertilitetsnivået i Norden sam-

menlignet med andre europeiske land (Rønsen og Skrede, 2008, Rønsen og Skrede, 2010). Fra 1980-tallet og utover ble det observert en økning i fertilitetsraten i samtlige nordiske land. Parallelt med dette var trenden fallende i mange av de andre europeiske landene. Det finnes ikke et entydig svar på hvorfor de nordiske fertilitetstrendene har stått i såpass kontrast til andre europeiske land. Likevel påpeker Rønsen og Skrede at generøsiteten som preger den nordiske familiepolitikken - kombinert med lønnskompenserte foreldrepermisjoner og offentlig subsidiert barnepass - kan være forklarende faktorer på hvorfor vi observerer et slikt mønster i de nordiske landene.

Imidlertid hevder forfatterne (Rønsen og Skrede, 2010) at det isolert sett ikke er tilstrekkelig med generøse støtteordninger, men at det også må legges til rette for et familievennlig og likedelt arbeidsmarked. De understreker videre at å føre en politikk som fremmer likestilling mellom kjønnene, og dermed også gjør det enklere å kombinere barn og arbeid, kan virke stimulerende på fertilitetsraten.

Anne Helene Gauthier og Jan Hatzius publiserte i 1997 en studie hvor de ved hjelp av en regresjonsanalyse studerer sammenhengen mellom familieytelser og fertilitet i 22 industriland i tidsperioden 1970 til 1990 (Gauthier og Hatzius, 1997). De finner en positiv sammenheng, men presiserer at effekten av familieytelser er begrenset. Gauthier og Hatzius bruker en logaritmisk funksjon, og antar i tillegg en tidsforskyvning der familieytelser eventuelt påvirker neste års fertilitet.

I 2013 gjorde Luci-Greulich og Thévenon en lignende undersøkelse av hvordan fertilitetstrender responderer på familiepolitiske tiltak. De ser på 18 OECD-land i årene 1982 til 2007 (Luci-Greulich og Thévenon, 2013). Tiltakene som undersøkes er kontantoverføringer, foreldrepermisjon og barnepass. Forfatterne bruker en regresjonsanalyse med faste effekter, og bruker flere kontrollvariabler. De finner at alle de tre tiltakene har en positiv effekt på fertiliteten, men i varierende grad. Den sterkeste positive sammenhengen finner de mellom kontantoverføringer i det første året etter fødsel og tilbud av barnepasstjenester de tre første leveårene.

3 Teori

I dette kapitlet presenteres den relevante økonometriske teorien for oppgaven. Teorien blir i resten av kapitlene brukt for å vurdere styrker og utfordringer ved datasettet og regresjonsanalysen.

3.1 Enkel og multippel regresjonsanalyse

Regresjonsanalyse er en statistisk metode som brukes for å undersøke sammenhengen mellom en avhengig variabel og en eller flere uavhengige variabler. Ved en enkel regresjonsanalyse har vi kun en uavhengig variabel og skriver ligningen på følgende form (Wooldridge, 2015, s.20):

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i \quad (1)$$

Ved en multippel regresjonsanalyse har vi flere uavhengige variabler (Wooldridge, 2015, s.69):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad (2)$$

Hvor Y er den avhengige variabelen, β_0 er konstantleddet, og X_1 til X_k de uavhengige variablene. Parameterne β_1 til β_k viser økningen i den avhengige variabelen hvis de uavhengige variablene øker med én enhet. Restleddet, u , er variasjoner i modellen som ikke er fanget opp i de uavhengige variablene.

I tilfellet med paneldata ser vi på data for forskjellige entiteter, i , gjort over tid, t . Vi skriver derfor ligningene på følgende form:

Ved enkel regresjonsanalyse:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + u_{it} \quad (3)$$

Og ved multippel regresjonsanalyse:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + u_{it} \quad (4)$$

3.2 Forutsetninger for modellen

I regresjonsanalysen kommer vi i hovedsak til å bruke multippel regresjonsanalyse. Vi presenterer her derfor forutsetningene for multippel regresjonsanalyse (Wooldridge, 2015):

-
- **MLR1:** Parameterne i modellen, b_0, b_1, \dots, b_k , er lineære.
 - **MLR2:** Tilfeldig utvalg. Vi har n antall observasjoner som representerer et tilfeldig utvalg av populasjonen.
 - **MLR3:** Ikke perfekt multikollinearitet mellom variablene. Variablene kan være korrelert med hverandre, men denne korrelasjonen kan ikke være lik 1.
 - **MLR4:** Den forventede verdien av restleddet er 0 gitt verdien av hver av de uavhengige variablene. $E(u|x_1, x_2, \dots, x_n) = 0$ For at forutsetningen skal holde kan det ikke være en variabel i restleddet u som påvirker en av de uavhengige variablene x_i samtidig som den påvirker y .
 - **MLR5:** Homoskedastisitet, $var(u|x_1, x_2, \dots, x_n) = \sigma^2$. Variasjonen i restleddet er konstant gitt de uavhengige variablene.
 - **MLR6:** Normalitet, restleddet u er normalt distribuert.

Disse forutsetningene er i stor grad de samme i en enkel regresjonsanalyse. Med unntak av forutsetning 3. I en enkel regresjonsanalyse gir det ikke mening å snakke om multikollinearitet. Isteden vil forutsetningen være at det er en variasjon i observasjonene i den uavhengige variabelen, alle observasjonene kan ikke være like. (Wooldridge, 2015, s.42)

3.3 Estimering med faste effekter

I regresjonsanalyser hvor man bruker paneldata har entitetene individuell karakteristik som kan påvirke resultatet av regresjonsanalysen. For å illustrere dette kan et ekstra ledd, a_i , legges til i regresjonsligningen fra delkapittel 3.1 (ligning 4):

$$y_{it} = \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + a_i + u_{it} \quad (5)$$

Variabelen a_i fanger opp alle uobserverte, tidskonstante faktorer som kan påvirke y_{it} . Siden variabelen ikke forandres over tid, kun mellom forskjellige entiteter, kalles den ofte en fast effekt (Wooldridge, 2015, s.440). Leddet u_{it} kalles den tidsvarierende feilen, siden den representerer uobserverte faktorer som forandres over tiden t .

Når vi bruker en modell med faste effekter fjerner vi den faste effekten, a_i , og uavhengige variabler som er tidskonstante.

Dette gjøres ved å først finne snitt verdien for hver entitet, i , over tid:

$$\bar{y}_{it} = \beta_1 \bar{x}_{it1} + \beta_2 \bar{x}_{it2} + \dots + \beta_k \bar{x}_{itk} + a_i + \bar{u}_{it} \quad (6)$$

Siden a_i er tidskonstant, vil den være lik i ligning (6), og ligning (7). Hvis man subtraherer ligning (7) fra ligning (6), ender man opp med følgende uttrykk:

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it1} + \beta_2 \ddot{x}_{it2} + \dots + \beta_k \ddot{x}_{itk} + \ddot{u}_{it} \quad (7)$$

Her er \ddot{y}_{it} tidsdifferensiert data for y , og det samme gjelder for \ddot{x}_{it1} og \ddot{u}_{it} .

Ved å bruke en estimering med faste effekter har vi altså fjernet tidsuavhengig forskjeller mellom entitetene (i vårt tilfelle land). Dette gjør at vi kan fokusere på hvordan en endring i en uavhengig variabel i et land vil påvirke den avhengige variabelen (i vårt tilfelle fertilitet), uten at tidskonstante forskjeller som allerede eksisterer mellom landene, blir tatt med i beregningen (Wooldridge, 2015, s. 463).

3.4 Minste kvadraters metode

Regresjonsligningen estimeres ved å bruke minste kvadraters metode. I minste kvadraters metode ser man på avstanden mellom observasjonspunktene, og den estimerte funksjonen. Man lar så den estimerte funksjonen være den funksjonen som minimerer summen av den kvadrerte avstanden fra observasjonene til funksjonen.

3.5 Hypotesetesting (T-test)

Når man gjennomfører en hypotesetest setter man opp en nullhypotese H_0 , og en alternativ hypotese H_1 .

Det er mulig å bruke hypotesetesting i mange typer vitenskapelige undersøkelser. I vårt tilfelle bruker vi hypotesetesting for å sjekke om parameterne i regresjonsanalysen, $\beta_0, \beta_1, \dots, \beta_k$, er forskjellige fra null.

Vi skiller mellom en ensidig og tosidig test. I begge tilfellene har vi en nullhypotese:

$$B_i = 0, i = 1, 2, \dots, k \quad (8)$$

Ved en ensidig test vil den alternative hypotesen (gitt nullhypotesen vi har i ligning 9) enten være:

$$B_i > 0 \quad (9)$$

eller:

$$B_i < 0 \quad (10)$$

Ved en tosidig test vil den alternative hypotesen være:

$$B_i \neq 0 \quad (11)$$

Det finnes flere forskjellige typer hypotesetester man kan bruke. I vår oppgave har vi valgt å bruke T-test. T-testen er en statistisk hypotesetest som baserer seg på Students t-fordeling. Målet ved å bruke en T-test er å teste om gjennomsnittsverdien i et normalt fordelt datasett er signifikant ulikt fra nullhypotesen.

T-verdien finner man på følgende måte:

$$t = \frac{\bar{B}_i - B_i}{s/\sqrt{n}} \quad (12)$$

B_i er den reelle verdien for parameteren, mens \bar{B}_i er et estimat man har gjort ut ifra utvalget (observasjonene) man har. s/\sqrt{n} er standardavviket. Hvis man ønsker å teste nullhypotesen $B_i = 0$ har vi følgende ligning for t-verdien:

$$t = \frac{\bar{B}_i}{s/\sqrt{n}} \quad (13)$$

Forenklet kan man si at om t-verdien er større, så vil beviset mot nullhypotesene være sterkere. Høyere standardavvik gjør det derfor vanskeligere å forkaste nullhypotesen mens en \bar{B}_i som er langt fra nullhypotesen gjør det enklere å forkaste nullhypotesen. T-verdien blir videre brukt i en t-fordeling, hvor man ønsker å finne p-verdien som er assosiert med den t-verdien. P-verdien viser sannsynligheten for å få en t-verdi som er lik eller større (i absoluttverdi) enn det man har fått gitt at nullhypotesen er sann. Hvis p-verdien er lavere enn signifikansnivået, ofte satt til 0.05, kan man forkaste nullhypotesen. (Wooldridge, 2015, s. 739)

3.6 Determinasjonskoeffisient - R^2

Determinasjonskoeffisienten er et statistisk mål på hvor stor andel av variansen i den avhengige variabelen som kan bli forklart ved bruken av de uavhengige variablene. Determinasjonskoeffisienten vil alltid befinne seg mellom 0 og 1, hvor en verdi lik 0 vil si at regresjonslinjen ikke passer dataen i det hele tatt, mens en verdi av 1 sier at regresjonslinjen passer dataen helt perfekt. (Wooldridge, 2015, s.35). I formelen under er SS_{res} , summen av kvadrerte residualer, mens SS_{tot} er den totale variasjonen i den avhengige variabelen.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (14)$$

SS_{res} viser til hvor mye av variasjonen som ikke kan forklares med regresjonsmodellen, mens SS_{tot} viser til den totale variasjonen i den avhengige variabelen. Selv om determinasjonskoeffisienten kan si noe om hvordan en regresjonsmodell presterer, så kan man ikke bruke tallet som en fasit på hvor god modellen er. Determinasjonskoeffisienten har for eksempel en tendens til å øke når man legger til flere uavhengige

variabler selv om de uavhengige variablene man legger til ikke har noen god forklaringssevne.

3.7 Multikollinearitet

Multikollinearitet er når det er en lineær sammenheng mellom to uavhengige variabler i en regresjonsmodell. (Wooldridge, 2015, s.90) Dette vil si at en uavhengig variabel i regresjonsanalysen kan bli predikert ved å bruke en annen uavhengig variabel i analysen. Multikollinearitet, spesielt ved nær perfekt korrelasjon, kan gjøre det vanskelig å finne gode koeffisienter i regresjonsanalysen siden koeffisientene vil være avhengig av hverandre.

I denne oppgaven er en korrelasjonsmatrise brukt for å se nærmere på multikollinearitet, mellom ulike variabler. En korrelasjonsmatrise vil gi korrelasjonskoeffisienten mellom hvert par av variabler i datasettet. En korrelasjonskoeffisient som er 0 vil indikere at det ikke er noen korrelasjon mellom variablene, mens 1 vil si at det er en perfekt positiv korrelasjon, og -1 gir en perfekt negativ korrelasjon. Formelen for å regne ut korrelasjonskoeffisienten er gitt ved:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (15)$$

4 Metode

Hensikten med kapitlet er å presentere, forklare og se på utfordringer med datagrunnlaget til regresjonsanalysen. Først presenteres data for fertilitet og familieytelser, deretter kontrollvariabler og oppsettet av ligningene. Til slutt ser vi nærmere på den deskriptive statistikken for alle variablene. I dette kapitlet, og i de neste kapitlene, brukes ofte det engelske begrepet for en variabel. Se vedlegg 1 for en mer grundig beskrivelse av hvert begrep på norsk.

4.1 Datagrunnlag - fertilitet og familieytelser

4.1.1 Generelt

Datasettet er hentet fra OECD sine databaser (OECD, 2023b). OECD er en organisasjon som arbeider for økonomisk vekst og bærekraftig utvikling (OECD, 2023i). Kjennetegnene som assosieres med OECD-landene er høy levestandard, høy utdanning, demokratisk styresett, stor økonomisk frihet, og velutviklet teknologi og industri. Dataen er derfor ikke representativ for hele verden. I vedlegg 2 er OECD-landene brukt i denne oppgaven listet opp med *verdensdelen* landet tilhører. Alle OECD-land som ble medlem før 2018 er inkludert i datasettet.

I oppgaven undersøker vi sammenhengen mellom fertilitet og familieytelser (som kan deles inn i *incash* og *inkind*). OECD definerer *incash*-ytelser som offentlige barnere-laterte pengeoverføringer, inkludert barnetrygd, foreldrepenger under permisjon og økonomiske støtteordninger for enslige forsørgere (OECD, 2023b). I noen land er barnetrygden behovsprøvd, og størrelsen på utbetalingene kan variere med barnets alder. Ytelser i *inkind* er offentlige utgifter til tjenester for familier med barn. Dette inkluderer blant annet finansiering og subsidiering av barnehager og økonomisk støtte til barnepass gjennom øremerkede utbetalinger til foreldre (OECD, 2023b). Både *incash*-ytelser og *inkind*-ytelser er målt som prosentandel brukt av BNP.

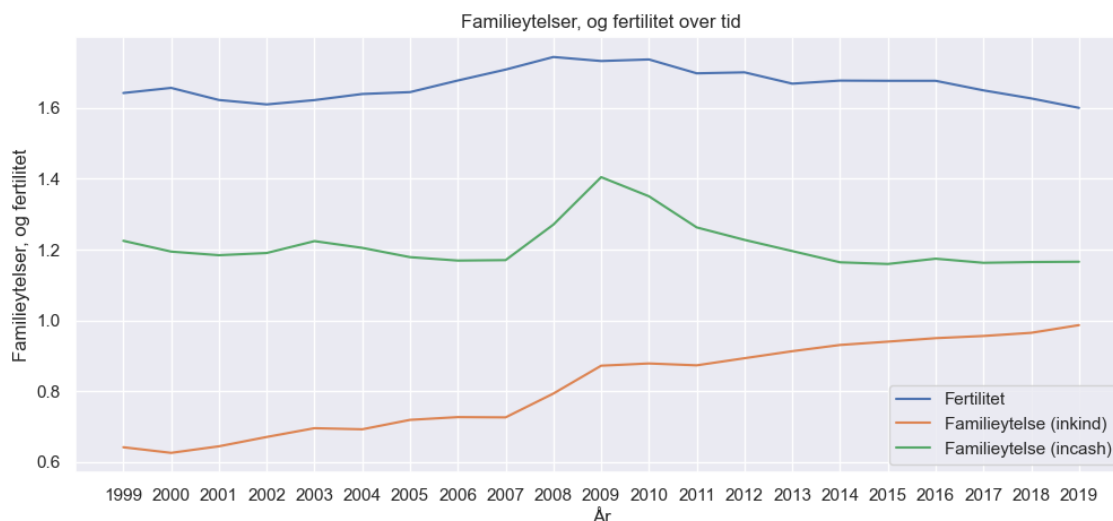
Tidsrommet vi har valgt å fokusere på i denne oppgaven er fra 1999 til 2019. Vi stopper i 2019 fordi det er en del land som mangler data etter dette året, i tillegg ønsker vi ikke at endringer i variablene som skyldes COVID-19 skal påvirke resultatene. Vi har startet i år 1999 fordi OECD hadde en medlemsvekst i årene før årtusenskiftet (OECD, 2023g). Ved å starte i 1999 får vi derfor med et datagrunnlag for flere land.

Fra man tar en avgjørelse om å få barn til barnet blir født vil det ta en viss tid. Derfor velger vi å se på sammenhengen mellom familieytelser i år t og fertilitet i år $t + 1$. Dette samsvarer med regresjonsligningen i artikkelen til Gauthier og Hatzius. Hvis et land har annonserte planer for hvordan utviklingen i familieytelser vil være i årene fremover kan det likevel hende at sammenhengen heller ligger mellom familieytelser i år t og fertilitet i år t . I presentasjonen av data og regresjonsanalysen ser

vi i hovedsak på sammenhengen mellom familieytelser og neste års fertilitet. Vi tar likevel også hensyn til at det kan være en eventuell umiddelbar effekt ved å i tillegg ha med en regresjonsligning for dette.

4.1.2 Fertilitet og familieytelser over tid

Vi begynner med å se på gjennomsnittlig fertilitet og familieytelser i OECD-landene fra 1999 til 2019:

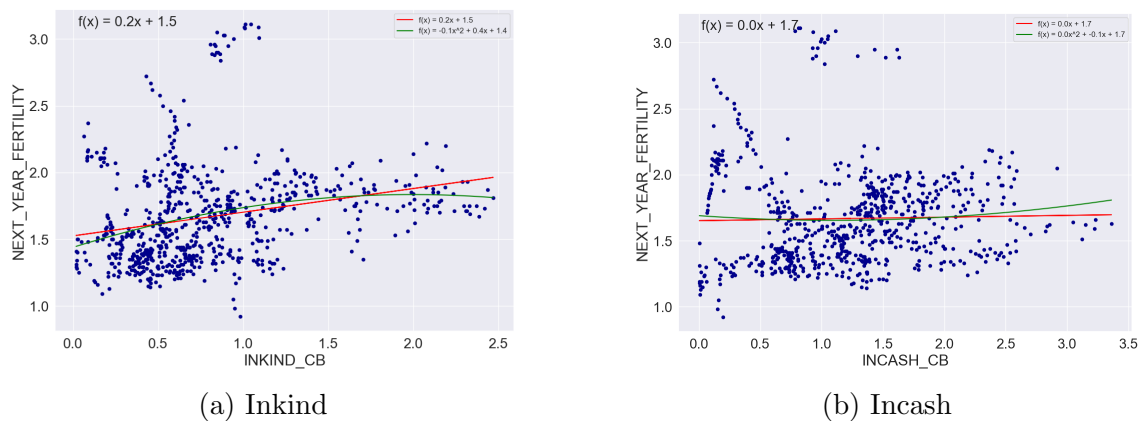


Figur 2: Familieytelser (målt i prosent av BNP) og fertilitet over tid (målt i barn per kvinne).

Familieytelser er målt i prosentandel brukt av BNP. Familieytelser er delt inn i støtte gjennom tjenester (heretter omtalt som *inkind*) og kontantutbetalinger (omtalt som *incash*). I grafen ser vi at *incash*-støtten ligger på et høyere nivå enn *inkind* i hele tidsperioden. Fertiliteten har holdt seg på et stabilt nivå de siste 20 årene, med en liten nedgang fra 1999 til 2019. *Inkind*-ytelser har steget siden årtusenskiftet. *Incash*-ytelser har variert litt; man ser en oppgang etter finanskrisen, mens nivået i dag er cirka det samme som det var i starten av 2000-tallet. Grunnen til at *incash* betalingene gikk opp etter finanskrisen kan være at BNP falt i dette tidsrommet, og at familieytelsene som prosentandel av BNP derfor ble større.

4.1.3 Forskjellige familieytelsers påvirkning på fertilitet

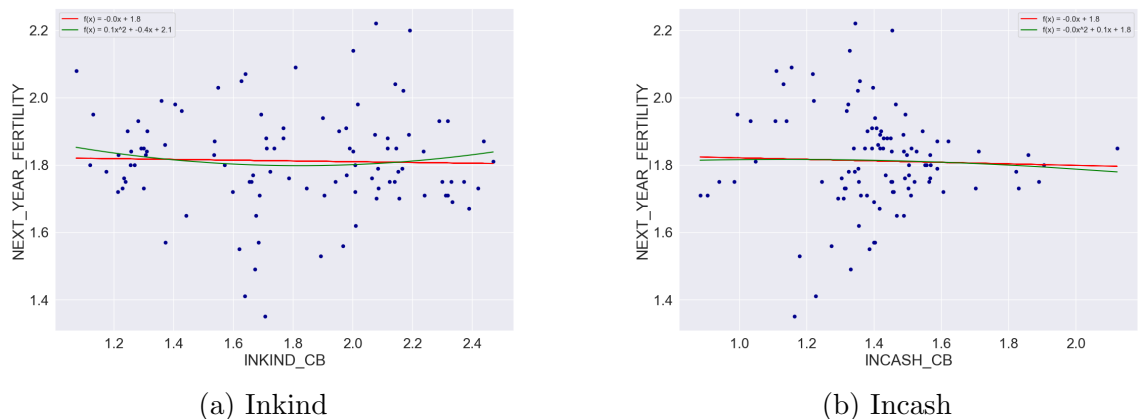
Videre ønsker man å se nærmere på den generelle sammenhengen mellom fertilitet og de ulike familieytelsene. Vi bruker da et spredningsdiagram:



Figur 3: OECD - Neste års fertilitet og familieytelser

Det ser ut til å være en sammenheng, både for *incash* og *inkind*, hvor land med veldig lav støtte gjerne har høy fertilitet. Deretter har vi en trend med økt støtte og synkende fertilitet frem til et vendepunkt hvor fertiliteten starter å øke med økte familieytelser. Når vi derimot tegner opp en trendlinje med bruk av en andregrads-ligning klarer vi ikke å fange opp denne tendensen. Dette kan skyldes at det er flere observasjoner for land med litt støtte og lav fertilitet enn det er for land med høy støtte og høyere fertilitet og land med lav støtte og høy fertilitet.

Videre kan vi undersøke dataen for de nordiske landene nærmere, som gir følgende figur:



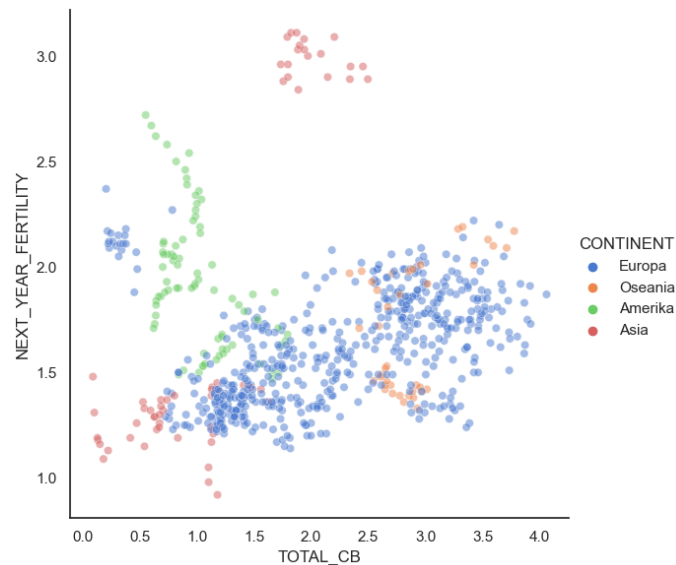
Figur 4: Skandinavia - Neste års fertilitet og familieytelser

De nordiske landene er likere enn OECD-landene, og man ser derfor at observasjonene er mer konsentrert enn når man ser på alle OECD-landene. I tillegg er det vanskeligere å finne en tydelig trend enn det er i dataene for hele OECD.

4.1.4 Geografiske forskjeller

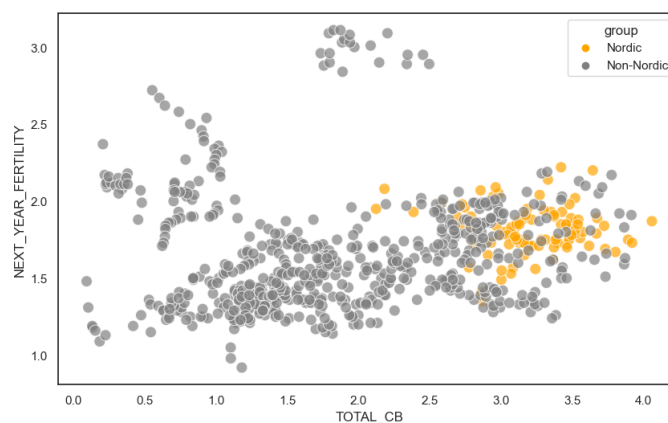
Det er interessant å se nærmere på hvordan geografiske forskjeller påvirker dataen. Selv om det er fellestrekk mellom OECD-landene som tidligere beskrevet i oppgaven,

så er det også store geografiske og kulturelle forskjeller. Vi har valgt å dele landene inn i fire områder; Europa, Oseania, Asia og Amerika. Forskjellene kan illustreres i et spredningsdiagram:



Figur 5: Spredningsdiagram kontinenter

Det er flest europeiske land i datasettet, og selv om familiestøtten er relativt stor i disse landene er ikke fertiliteten spesielt høy. Asia skiller seg ut ved at dataen er veldig spredt. Her finner vi både noen av landene med lavest fertilitet, Japan og Sør-Korea, og høyest fertilitet, Israel. Dette kan være et resultat av at Asia er et stort kontinent, og at det er store kulturelle forskjeller mellom disse tre landene. Amerikanske land har relativt høy fertilitet i forhold til familiestøtten de gir, mens Oseania i stor grad ser lik ut som Europa.



Figur 6: Spredningsdiagram nordiske land

I figur 6 ser vi at de nordiske landene skiller seg ut ved at en høyere prosentandel av BNP brukes på familieytelser enn i de andre landene. I figur 6 ser vi også tydeligere

at det er mindre spredning i observasjonene vi har for nordiske land (både når det gjelder andel av BNP brukt på familieytelser og fertilitetsrate) sammenlignet med alle OECD-land.

4.1.5 Bruk av faste effekter

Etter å ha studert spredningsdiagrammene virker det fornuftig å bruke faste effekter i regresjonsanalysen. Den store spredningen i fertilitet kan skyldes at det er store forskjeller i forhold som kultur, religion og tankesett mellom landene. Dette vil være faktorer som mest sannsynlig ikke har store variasjoner innad i landene over tid. Ved å bruke faste effekter unngår vi at resultatene i regresjonsanalysen skyldes en slik type heterogeniteten mellom landene.

4.2 Kontrollvariabler

4.2.1 Generelt

Hvis man kun gjør en regresjonsanalyse hvor man ser på sammenhengen mellom *incash*, *inkind* og fertilitet, risikerer man å finne en såkalt spuriøs sammenheng. En sammenheng mellom den uavhengige variabelen og den avhengige variabelen som egentlig skyldes en variabel som ligger i restleddet (Ward, 2013). I vårt tilfelle risikerer vi altså at det er en annen variabel som både påvirker nivået på familieytelser og fertiliteten. For å unngå dette skiller vi ut variabler som finnes i restleddet og inkluderer disse som kontrollvariabler i regresjonsanalysen. I analysen vår er vi ikke interessert i hvordan kontrollvariablene påvirker fertiliteten i seg selv, de er inkludert for at de ikke skal påvirke en eventuell sammenheng mellom familieytelser og fertilitet.

4.2.2 Valg av kontrollvariabler

Kontrollvariablene vi velger å inkludere er *unemployment_rate_woman*, *unemployment_rate_all* (OECD, 2023j), *infant_mortality* (OECD, 2023f), *gdp_per_capita* (OECD, 2023e), *working_age_population* (OECD, 2023k), *gender_wage_gap* (OECD, 2023d), *foreign_born_population* (OECD, 2023c) og *percent_with_teretatry_education* (OECD, 2023h). Se vedlegg 1 for en mer detaljert beskrivelse. Alle observasjonene er hentet fra OECD sine databaser.

Vi har valgt å ta med *unemployment_rate_woman* og *unemployment_rate_all* fordi arbeidsledighet innebærer økonomisk usikkerhet. Av den grunn ser vi for oss at arbeidsledighet kan ha en negativ innvirkning på fertilitet. Arbeidsledighet kan også ha en negativ virkning på familieytelser ved at en økt andel av BNP må brukes på ledighetstrygd. Begge disse kontrollvariablene er også med i studien Luci-Greulich

og Thévenon.

Barnedødelighet, *infant_mortality*, er inkludert fordi det kan hende at høy barnedødelighet kan føre til høyere fertilitet. I tillegg vil økt barnedødelighet, med alt annet likt, gi lavere familieytelser hvis nivået på familieytelser avhenger av antall barn som blir født.

I tillegg ser vi for oss at hvor stor andel av befolkningen som tilhører arbeidsstyrken, vil være av betydning. Dette sier oss noe om alderssammensetningen i befolkningen, som igjen har betydning for både bruk av offentlige inntekter og fertilitet. Hvis *working_age_population* er lav, betyr dette at mange er over 64 år og/eller mange under 15 år, altså mange som ikke er i fruktbar alder. Lav *working_age_population* vil derfor innebære lavere fertilitet. Det virker også rimelig å anta at en lavere *working_age_population* vil føre til at en større andel av BNP brukes på enten skole eller eldreomsorg som derfor vil redusere andelen BNP brukt på familieytelser.

Rønsen og Skrede har i sin forskning uttalt at et likestilt arbeidsmarked kan virke positivt på fertilitetsraten. Likestilling i arbeidsmarkedet er et vidt begrep, og vi har valgt å kun ta med kontrollvariabelen *gender_wage_gap*, selv om også andre variabler kunne vært inkludert. Basert på funnene til Rønsen og Skrede ønsker vi å luke ut en sammenheng der mindre likestilling, høyere *gender_wage_gap*, gir lavere fertilitet. Samtidig som høyere *gender_wage_gap* også har en innvirkning på familieytelser målt i andel av BNP. Hvis vi antar at kvinner oftere har omsorgen for barnet (i tilfeller med en forsørger) og at familieytelser for eksempel er behovstrygd, vil høyere *gender_wage_gap* føre til høyere familieytelser.

En variabel for *foreign_born_population* er også tatt med. Det kan hende at land med gode offentlige tjenester tiltrekker seg flere innvandrere med barn, altså at innvandringen er større i land hvor familieytelsene er høyere. Høyere innvandring gir ikke i seg selv høyere fertilitet, selv om populasjonen øker. Det kan derfor være at høyere innvandring gir økte familieytelser uten at fertiliteten påvirkes. Vi ønsker å luke ut denne effekten.

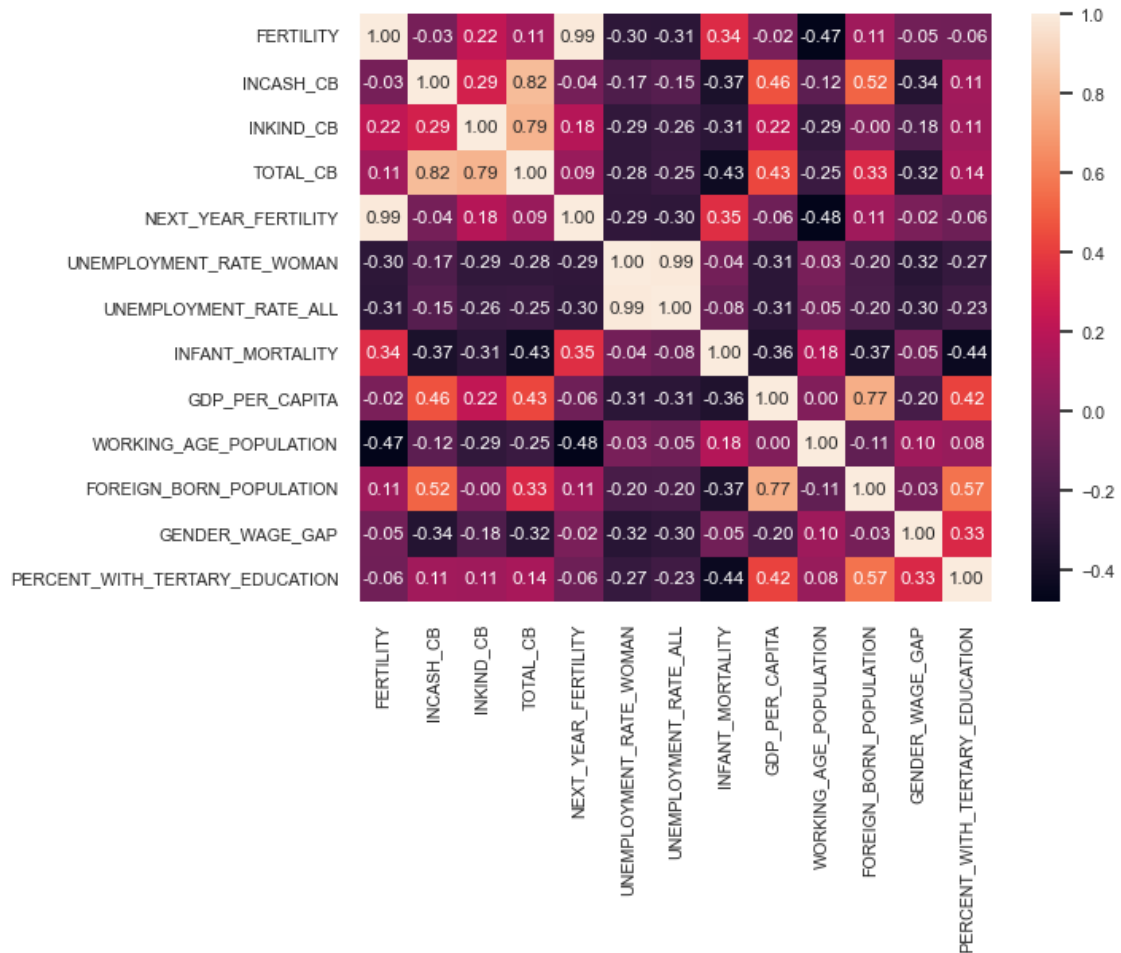
Dessuten inkluderer vi en variabel som representerer utdanningsnivået i befolkningen, *percent_with_tertiary_education*. Det kan tenkes at lengre utdanningsløp påvirker fertiliteten gjennom en økning i gjennomsnittsalderen for førstegangsfødende. Dette kan igjen påvirke totalt antall barn hver kvinne får. En økning i *percent_with_tertiary_education* kan også føre til økt offentlig investering i utdanning som, med alt annet likt, gir lavere prosentandel brukt på familieytelser.

Til slutt legger vi til en tidstrend-variabel. Dette er en variabel som Luci-Greulich og Thévenon også har inkludert i sin regresjonsanalyse. Vi inkluderer denne variabelen for å unngå at vi får med en effekt på familieytelser og fertilitet som skyldes en generell endring over tid i disse variablene. Det kan for eksempel være at det har

vært en tidstrend med en synkende fertilitet eller tidstrend med økte ytelser.

4.2.3 Korrelasjon mellom variabler

Vi ønsker også å undersøke korrelasjonen mellom variablene. Vi velger å se på korrelasjonen i 2014 fordi vi mangler minst data fra dette året.



Figur 7: Korrelasjonsmatrise - med data fra 2014

Korrelasjonsmatrisen viser en nær perfekt positiv korrelasjon (0,99) mellom total arbeidsledighet og arbeidsledigheten av kvinner som prosent av arbeidsstyrken. Derfor velger vi å ta bort sistnevnte (*Unemployment_rate_woman*).

4.3 Beskrivelse av regresjonsligningene

4.3.1 Generelt

I regresjonsanalysen bruker vi 6 forskjellige ligninger for å beskrive sammenhengen mellom familieytelser og fertilitet. Vi har med flere ligninger for å kunne få et mer

nøyaktig bilde av den reelle sammenhengen mellom familieytelser og fertilitet. Ved å ha flere ligninger, får vi både sett på effekten av totale familieytelser på fertilitet, delt inn i *incash* og *inkind*, og sett på hvordan dette påvirker fertilitet. I tillegg har vi muligheten til å studere hvordan resultatet endrer seg når vi legger til flere kontrollvariabler. Flere kontrollvariabler gjør at vi i større grad kan finne den reelle effekten av *incash* og *inkind* ytelser, men dette fører samtidig til at antall observasjoner reduseres betraktelig. Derfor har vi i noen av ligningene ikke tatt med alle kontrollvariablene. I tillegg har vi en ligning hvor vi bruker *current_year_fertility* isteden for *next_year_fertility* for å se om det er mulig å fange opp en umiddelbar effekt av familieytelser.

I regresjonsligningene og i fremstillingen av resultatene forkortes noen av variabelnavnene. Se vedlegg 3 for en oversikt over variabelnavn og forkortelser.

4.3.2 Regresjonsligningene

Vi begynner med en enkel lineær regresjonsmodell på formatet:

$$\text{nextyearfertility}_{it} = \beta_0 + \beta_1 \cdot \text{total_cb}_{it} + u_{it} \quad (16)$$

Deretter utvider vi til multippel lineær regresjon ved å dele familieytelser inn i to kategorier:

$$\text{nextyearfertility}_{it} = \beta_0 + \beta_1 \cdot \text{incash_cb}_{it} + \beta_2 \cdot \text{inkind_cb}_{it} + u_{it} \quad (17)$$

Deretter legges kontrollvariabler til. Begynner med to modeller hvor arbeidsledighetsraten, BNP per innbygger og utenlandsfødt populasjon er lagt inn som kontrollvariabler. Det som skiller de to modellene fra hverandre er at ligning nummer 19 deler familieytelser i *incash* og *inkind*-støtte.

$$\begin{aligned} \text{nextyearfertility}_{it} = & \beta_{0it} + \beta_1 \cdot \text{total_cb}_{it} \\ & + \beta_2 \cdot \text{unemployment_rate}_{it} + \beta_3 \cdot \text{gdp_per_capita}_{it} \\ & + \beta_4 \cdot \text{foreign_born_pop}_{it} + u_{it} \end{aligned} \quad (18)$$

$$\begin{aligned} \text{nextyearfertility}_{it} = & \beta_{0it} + \beta_1 \cdot \text{incash_cb}_{it} + \beta_2 \cdot \text{inkind_cb}_{it} \\ & + \beta_3 \cdot \text{unemployment_rate}_{it} + \beta_4 \cdot \text{gdp_per_capita}_{it} \\ & + \beta_5 \cdot \text{foreign_born_pop}_{it} + u_{it} \end{aligned} \quad (19)$$

I ligning 5 legger vi til barnedødelighet, andel av befolkningen i arbeidsstyrken, lønnsgap mellom kjønn, andel mellom 25 og 34 som har fullført den høyeste graden mulig for deres alder og en tidstrend. Da er alle kontrollvariablene inkludert.

$$\begin{aligned}
nextyearfertility_{it} = & \beta_{0it} + \beta_1 \cdot incash_cb_{it} + \beta_2 \cdot inkind_cb_{it} + \beta_3 \cdot unemployment_rate_{it} \\
& + \beta_4 \cdot gdp_per_capita_{it} + \beta_5 \cdot foreign_born_pop_{it} + \beta_6 \cdot infant_mortality_{it} \\
& + \beta_7 \cdot working_age_pop_{it} + \beta_8 \cdot gender_wage_gap_{it} + \beta_9 \cdot p_with_ter_educ_{it} \\
& + \beta_{10} \cdot time_trend_t + u_{it}
\end{aligned} \tag{20}$$

Den siste ligningen er lik ligning nummer 5, men vi erstatter *next_year_fertility* med *fertility*:

$$\begin{aligned}
currentyearfertility_{it} = & \beta_{0it} + \beta_1 \cdot incash_cb_{it} + \beta_2 \cdot inkind_cb_{it} + \beta_3 \cdot unemployment_rate_{it} \\
& + \beta_4 \cdot gdp_per_capita + \beta_5 \cdot foreign_born_pop_{it} + \beta_6 \cdot infant_mortality_{it} \\
& + \beta_7 \cdot working_age_pop_{it} + \beta_8 \cdot gender_wage_gap_{it} + \beta_9 \cdot p_with_ter_educ_{it} \\
& + \beta_{10} \cdot time_trend_t + u_{it}
\end{aligned} \tag{21}$$

4.4 Deskriptiv statistikk

Etter å ha presentert alle variablene som er tatt med i regresjonsanalysen ønsker vi å se nærmere på den deskriptive statistikken for å se om det er andre eventuelle utfordringer ved datasettet.

4.4.1 Deskriptiv statistikk - alle OECD-land

Descriptive Statistics					
Variable	Obs	Mean	Std. Dev.	Min	Max
fertility c y	756	1.667	.369	.92	3.11
incash cb	756	1.211	.662	.003	3.362
inkind cb	756	.814	.517	.018	2.471
time trend	756	10	6.059	0	20
gdp per capita	756	34.67	16.229	7.406	119.253
working age pop	756	66.67	2.367	59.502	73.418
infant mortality	740	4.664	2.938	.7	29
unemployment rate	735	7.696	4.059	1.9	27.825
fertility n y	720	1.668	.366	.92	3.11
p with ter educ	687	36.587	12.375	8.713	69.852
foreign born pop	597	13.069	8.448	.5	47.3
gender wage gap	487	16.098	8.012	.384	41.803

Figur 8: Deskriptiv statistikk alle OECD-land

Vi har total 12 variabler hvor to av dem (fertilitet i år t og fertilitet i år $t + 1$) brukes som avhengig variabler og resten brukes som uavhengige.

Dataen er i tabellen sortert etter antall observasjoner. Hvis antall observasjoner er mindre enn 756, mangler vi data for enkelte land i enkelte år. Dette er en av de

største problemene i datasettet ettersom det ikke er sikkert at det er tilfeldig hvilke land som mangler data i enkelte år. Denne utfordringen gjelder ikke for fertilitet i år $t + 1$. Denne variabelen inneholder kun 720 observasjoner fordi vi ikke har med tall fra 2020 i datasettet. For alle landene står derfor *next_year_fertility* tom i 2019.

Det blir i gjennomsnitt brukt en større prosentandel av BNP på *incash*-ytelser enn *inkind*-ytelser, noe som samsvarer med figur 2 i delkapittel 4.1.2. Gjennomsnittet for *incash* er 1,211% og gjennomsnittet for *inkind* er 0,814%. Det er også større spredning i verdiene våre relativt til gjennomsnittsverdien for *incash*-ytelser. Standardavviket ligger på 0,662 for *incash* og 0,517 for *inkind*. Vi har i tillegg en høyere maks prosentandel som blir brukt på *incash*-benefits. De høyeste observasjonene er 3,362% av BNP brukt på *incash* og 2,471% av BNP på *inkind*.

4.4.2 Deskriptiv statistikk - Norden

Descriptive Statistics					
Variable	Obs	Mean	Std. Dev.	Min	Max
fertility c y	105	1.81	.152	1.35	2.22
incash cb	105	1.395	.21	.883	2.123
inkind cb	105	1.802	.38	1.073	2.471
gdp per capita	105	43.424	11.116	24.767	70.253
time trend	105	10	6.084	0	20
infant mortality	105	2.811	.848	.7	5.3
working age pop	105	65.388	1.271	62.064	67.527
unemployment rate	101	5.974	2.063	2.475	10.217
foreign born pop	100	9.496	4.196	2.5	19.5
fertility n y	100	1.812	.151	1.35	2.22
p with ter educ	92	40.401	4.968	28.631	49.008
gender wage gap	86	11.84	5.533	4.368	28.722

Figur 9: Deskriptiv statistikk nordiske land

Ettersom vi har fem land og observasjoner over 21 år, blir totalt antall (maksimum) observasjoner 105. For *next_year_fertility* har vi ikke inkludert data for 2020 og totalt antall observasjoner blir derfor 100. Gjennomsnittlig fertilitet er høyere i de nordiske landene enn for resten av OECD-landene, 1,81 sammenlignet med 1,667 for *current_year_fertility*. Standardavviket på *current_year_fertility* er også lavere. For de nordiske landene er det 0,152 og for alle land 0,369. Ellers er gjennomsnittlig prosentandel bruk på *incash* og *inkind* også høyere i de nordiske landene. Til forskjell fra OECD-landene er det i gjennomsnitt i de nordiske landene brukt mer på *inkind*-ytelser enn *incash*-ytelser.

5 Resultater

I dette kapitlet presenteres resultatene fra regresjonsanalysen. Resultatkapitlet er delt inn i to deler, hvor vi først presenterer resultatene for regresjonsanalysen gjort for alle OECD-land og deretter resultatene for Norden.

Regresjonsanalysen er gjennomført i *Stata*. For å gjøre regresjonsanalyse av paneldata i *Stata* brukes “xtset”-kommandoen for å presisere hvilke variabler som beskriver de ulike entitetene, og hva som er tidsvariabelen. For å gjøre regresjonen, brukes “xtreg” som tilpasser regresjonen til paneldata. *Stata* utfører også en tosidig hypotesetest med en nullhypotese lik ligning 8 i delkapittel 3.5. Hypotesetesten brukes når vi presenterer resultatene, og når vi diskuterer resultatene i neste kapittel. *Stata* gjør hypotesetesting med et signifikansnivå på 1, 5 og 10 prosent. I oppgaven vil vi, i samsvar med det som står beskrevet i teoridelen, ha som utgangspunkt at resultatet er statistisk signifikant dersom man kan forkaste nullhypotesen med et signifikansnivå på 5 prosent.

5.1 Alle OECD-land

Med datasettet fra OECD gjennomfører vi en regresjonsanalyse med de 6 ligningene fra delkapittel 4.3.

	(1)	(2)	(3)	(4)	(5)	(6)
	fertility_n_y	fertility_n_y	fertility_n_y	fertility_n_y	fertility_n_y	fertility_c_y
total_cb	.059*** (.014)		.098*** (.016)			
incash_cb		.101*** (.019)		.126*** (.02)	.082*** (.029)	.102*** (.029)
inkind_cb		0 (.022)		.031 (.033)	.095** (.046)	.136*** (.045)
unemployment_rate			-.007*** (.002)	-.008*** (.002)	-.01*** (.003)	-.007** (.003)
gdp_per_capita			.002*** (.001)	.003*** (.001)	.001 (.003)	.002 (.003)
foreign_born_pop			-.019*** (.003)	-.018*** (.003)	-.021*** (.005)	-.016*** (.005)
infant_mortality					.007 (.013)	-.004 (.013)
working_age_pop					.005 (.007)	-.001 (.007)
gender_wage_gap					.001 (.003)	.002 (.003)
p_with_ter_educ					.013*** (.002)	.012*** (.002)
time_trend					-.009* (.005)	-.011** (.005)
_cons	1.549*** (.028)	1.545*** (.028)	1.713*** (.042)	1.698*** (.042)	1.031* (.547)	1.357** (.537)
Observations	720	720	555	555	344	365
R-squared	.026	.042	.188	.197	.259	.224

Standard errors are in parentheses

*** $p < .01$, ** $p < .05$, * $p < .1$

Figur 10: Resultater fra regresjonsanalyse alle OECD-land

I ligning 1 ser vi en forventet positiv sammenheng mellom totale familieytelser og *next_year_fertility*. For hvert prosentpoeng man øker totale ytelser med (som prosentandel av BNP), er den forventede økningen i *next_year_fertility* 0,059. Ved hjelp av tabellen kan vi konkludere med at det er mulig å forkaste nullhypotesen med et signifikansnivå på 1%. Det ser ut til å være en statistikk signifikant positiv sammenheng mellom ytelser til familier og fertilitet.

I ligning 2 er forskjellen mellom den forventede virkningen på *next_year_fertility* med *incash* og *inkind* ytelser stor. For *incash* har vi en forventet økning i neste års fertilitet på 0,101 hvis *incash* øker med ett prosentpoeng. For *inkind* er den forventede økningen 0 (rundet av med tre desimaler). Vi kan forkaste nullhypotesen om at *incash*-ytelser ikke har en effekt på fertilitet med et signifikansnivå på 1%, mens vi ikke kan forkaste nullhypotesen hverken på 1%, 5% eller 10% signifikansnivå for *inkind*.

I ligning 3 har vi lagt til kontrollvariablene *unemployment_rate_all*, *gdp_per_capita* og *foreign_born_population*. Ved å legge til disse kontrollvariablene har den forventede økningen i *next_year_fertility*, ved en økning på ett prosentpoeng i totale familieytelser, gått fra 0,059 til 0,098. Resultat er statistisk signifikant ved et 1% signifikansnivå.

I ligning 4 har vi de samme kontrollvariablene, men familieytelser er nå delt inn i *inkind* og *incash*. Med kontrollvariablene går den forventede økningen i *next_year_fertility* fra 0,101 til 0,126 ved en økning i *incash* på ett prosentpoeng. Resultatet er fortsatt statistisk signifikant. Når vi ser på effekten av økning i *inkind* ytelser, går den forventede effekten opp fra 0 til 0,031. Resultatet er fortsatt ikke statistisk signifikant, vi kan hverken forkaste nullhypotesen på et signifikansnivå på 1%, 5% eller 10%.

Å inkludere kontrollvariablene i ligning 3 og 4 gjør at totale familieytelser og *inkind* og *incash* ytelser får en sterkere effekt på fertilitet. Ved å inkludere kontrollvariablene har vi separert ut en negativ sammenheng mellom fertilitet og familieytelser som egentlig skyldtes arbeidsledighet og innvandring. I tillegg har vi skilt ut en tilsynelatende positiv, men mindre, sammenheng som skyldtes BNP.

I ligning 5 er alle kontrollvariablene tatt med. Vi har altså lagt til *infant_mortality*, *working_age_population*, *gender_wage_gap*, *percent_with_tertary_education* og en variabel for tidstrend. Hvis vi sammenligner med ligning 4 har den forventede økningen i *next_year_fertility*, ved en økning i andel av BNP bruk på *incash* ytelse på ett prosentpoeng, gått nedover fra 0,126 til 0,082. Vi kan fortsatt forkaste nullhypotesen ved et signifikansnivå på 1%. Den forventede økningen i fertilitet ved en økning i *inkind* har gått oppover fra 0,031 til 0,095. Vi kan nå også forkaste nullhypotesen ved et signifikansnivå på 5%.

I ligning 6 er den avhengige variabelen skiftet ut mens de uavhengige variablene er

de samme. Hvis vi ser på effekten av en økning i *inkind*-ytelser på ett prosentpoeng, har den forventede økningen i fertilitet nå gått fra 0,082 til 0,102. Vi kan fortsatt forkaste nullhypotesen ved et signifikansnivå på 1%. Når det kommer til *inkind*-ytelser, har den forventede effekten økt fra 0,095 til 0,136. Vi kan også her forkaste nullhypotesen ved et signifikansnivå på 5%.

I ligningene 5 og 6 har vi fått skilt ut en god del variabler som har ligget i restleddet. I begge ligningene har dette gjort at effekten av *incash* har blitt mindre mens effekten av *inkind* har blitt større sammenlignet med ligning 4. Totalt antall observasjoner har gått ned. Dette gjør at vi risikerer å bryte forutsetningen om tilfeldig utvalg. Vi ser også at vi får en høyere forventet økning i fertilitet i likning 6 når vi bruker *current_year_fertility* som avhengig variabel.

5.1.1 Determinasjonskoeffisient - R^2

Determinasjonskoeffisienten øker med antall kontrollvariabler vi legger til. Den er også høyere når vi deler familieytelser inn i *inkind* og *incash* enn når vi ser på totalen. R^2 er høyere i ligning 5, hvor *next_year_fertility* er den avhengige variabelen, enn i ligning 6 hvor *current_year_fertility* er den avhengige variabelen.

5.2 Norden

Her gjør vi tilsvarende analyse med de nordiske landene. Vi bruker de samme seks ligningene med de samme variablene, men i denne regresjonen har vi bare beholdt data for Danmark, Finland, Island, Norge og Sverige.

	(1)	(2)	(3)	(4)	(5)	(6)
	fertility_n_y	fertility_n_y	fertility_n_y	fertility_n_y	fertility_n_y	fertility_c_y
total_cb	-.001 (.049)		.296*** (.081)			
incash_cb		.337*** (.077)		.642*** (.123)	.838*** (.152)	1.002*** (.158)
inkind_cb		-.038 (.044)		.088 (.096)	-.059 (.081)	.053 (.081)
unemployment_rate			-.005 (.011)	-.004 (.01)	-.018* (.01)	-.007 (.011)
gdp_per_capita			.009*** (.003)	.015*** (.003)	.01** (.004)	.013*** (.005)
foreign_born_pop			-.07*** (.013)	-.053*** (.013)	-.068*** (.011)	-.067*** (.011)
infant_mortality					.011 (.031)	.023 (.034)
working_age_pop					.085*** (.015)	.064*** (.016)
gender_wage_gap					.013 (.01)	.016 (.01)
p_with_ter_educ					.012** (.005)	.012** (.006)
time_trend					.028** (.011)	.024* (.012)
_cons	1.817*** (.158)	1.407*** (.159)	1.132*** (.228)	.615** (.257)	-5.482*** (.925)	-4.811*** (.984)
Observations	100	100	92	92	71	75
R-squared	0	.231	.353	.441	.813	.784

Standard errors are in parentheses

*** $p < .01$, ** $p < .05$, * $p < .1$

Figur 11: Regresjonsanalyse nordiske land

I ligning 1 er sammenhengen mellom totale familieytelser og *next_year_fertility* svak negativ. En økning i totale familieytelser gir en forventet reduksjon i fertiliteten på 0,001. Ved hypotesetesting med et signifikansnivå på 1%, 5% og 10% kan vi ikke forkaste nullhypotesen vår som er at sammenhengen er lik null.

I ligning 2 finner vi en positiv sammenheng mellom *incash*-ytelser og *next_year_fertility*. En økning i prosentandel brukt på *incash* gir en forventet økning i fertilitet på 0,296. Ved hypotesetesting kan vi forkaste nullhypotesen med en signifikansnivå på 1%. Sammenhengen mellom *inkind* og fertilitet er svak negativ. En økning på ett prosentpoeng i *inkind* gir en forventet reduksjon i fertiliteten på 0,038. Vi kan ikke forkaste nullhypotesen på noen av signifikansnivåene.

I ligning 3 er sammenhengen mellom totale familieytelser og *next_year_fertility* blitt positiv ved å legge til de tre kontrollvariablene. En ett prosentpoengs økning i totale familieytelser gir nå en forventet økning i fertiliteten med 0,269. Vi kan forkaste nullhypotesen med en signifikansnivå på 1%.

I ligning 4 har vi en sterk positiv sammenheng mellom *incash* og *next_year_fertility*. En økning i ett prosentpoeng brukt på *incash*-ytelser gir en forventet økning i fertiliteten med 0,642. Vi kan forkaste nullhypotesen med en signifikansnivå på 1%. Sammenhengen mellom *inkind* og fertilitet er positiv, vi har en forventet økning på 0,088, men vi kan ikke forkaste nullhypotesen.

Det kan se ut til at økningen i effekten av totale ytelser fra ligning 1 til ligning 3, og økningen i effekten av *incash*-ytelser og *incash*-ytelser fra ligning 2 til 4, har samme forklaring som i regresjonsanalysen for alle OECD-land. I ligning 1 og 2 påvirker variablene i restleddet både familieytelsene og fertiliteten. Ved å inkludere kontrollvariablene får vi et bedre bilde av den reelle effekten.

I ligning 5, hvor alle kontrollvariablene er inkludert, har sammenhengen mellom *incash* og *next_year_fertility* blitt enda sterkere. En økning i *incash* med ett prosentpoeng gir en forventet økning i fertiliteten på 0,838. Vi kan forkaste nullhypotesen med et signifikansnivå på 1%. Med *inkind* er sammenhengen svak negativt, en forventet nedgang på 0,059. Vi kan ikke forkaste nullhypotesen.

I ligning 6 hvor den avhengige variabelen har blitt *current_year_fertility* har sammenhengen mellom *incash* og fertilitet igjen økt. En økning i *incash* med ett prosentpoeng gir en forventet økning i fertiliteten på 1,002. Vi kan forkaste nullhypotesen med et signifikansnivå på 1%. Sammenhengen mellom *inkind* og fertilitet har nå blitt svak positiv, forventet oppgang i fertilitet på 0,053, men vi kan fortsatt ikke forkaste nullhypotesen.

Til forskjell fra regresjonsanalysen for alle land har ikke det å inkludere flere kontrollvariabler gjort at *inkind* har fått en statistisk signifikant effekt på fertiliteten. Også for Norden reduseres antall observasjoner når vi inkluderer alle kontrollvariablene.

5.2.1 Determinasjonskoeffisient - R^2

I den første ligningen uten kontrollvariabler er determinasjonskoeffisienten lik 0. Determinasjonskoeffisienten er høyest når alle kontrollvariablene er tatt med, den ligger da på 0,813. Determinasjonskoeffisienten er også høyere når vi ser på *next_year_fertility* sammenlignet med *current_year_fertility*.

6 Diskusjon

Diskusjonsdelen er delt inn i tre seksjoner. Vi begynner med en vurdering og diskusjon av hvorvidt analysen vi har gjort bryter med forutsetningene for regresjonsanalysen som er presentert i delkapittel 3.2. Deretter diskuterer vi resultatene vi har fått i regresjonsanalysen for OECD-land, både når det kommer til effekt av totale familieytelser og ytelser delt inn i *inkind* vs *incash*. Til slutt diskuterer vi resultatene for de nordiske landene og sammenligner dette med resultatene vi har fått i OECD-landene.

6.1 Evaluering av forutsetningene

Den første forutsetningen, MLR.1, er at parameterne i modellen er lineære. Dette kan problematiseres. En utfordring er at en økning med ett prosentpoeng i familieytelser kan ha større effekt hvis støtten er på et lavt nivå kontra et høyt nivå. For eksempel at en økning i familieytelser har større effekt hvis man går fra 2% til 3%, sammenlignet med å gå fra 49% til 50%. Dette kunne man løst ved å skrive ligningen på et annet format, for eksempel ved å bruke en logaritmisk form på ligningen som i publikasjonen til Gauthier og Hatzius. Siden vi ikke har gjort dette vil modellen mest sannsynlig gi dårlige predikasjoner på hva som vil skje hvis et land velger å innføre en stor økning i familieytelser.

Forutsetning 2 er at vi har et tilfeldig utvalg. Her er det to mulige utfordringer med modellen. For det første har vi kun brukt data fra OECD-land. Ved å kun inkludere OECD-land, velger vi typisk bort land med svak økonomi og lite velferd. I mange av disse landene er det ofte andre utfordringer knyttet til fruktbarhet enn det vi legger til grunn for problemstillingen vår. Modellen vår kan derfor kun brukes til å estimere sammenhengen mellom fertilitet og støtte til familier i OECD-land. Den gir ikke et representativt bilde av sammenhengen i alle verdens land.

Den andre utfordringen ved forutsetning 2 i regresjonsanalysen er at enkelte verdier mangler for noen land. I ligning 5 har antall observasjoner gått fra 720 til 344. Manglende data bryter i seg selv ikke med forutsetningen om tilfeldig utvalg, men vi er nødt til å reflektere over hvorfor enkelte land mangler data. Variabelen vi mangler mest data på er *gender wage gap*. Det kan være sannsynlig at land som fører politikk for å hindre lønnsgap mellom kjønn samler inn data på dette området, mens land som ikke har politikk på dette heller ikke har samlet inn dataen. Dermed kan estimeringen av lønnsgapet bli i overkant positiv og vi bryter MLR.2. Dette er hvorfor vi ikke har valgt å ta med enda flere kontrollvariabler i analysen. Jo flere variabler vi legger til, jo mer data vil også mangle og jo større er sannsynligheten for at resultatene våre blir uriktige eller unøyaktige.

MLR 3 antar at det ikke er en perfekt lineær relasjon mellom variablene i modellen. Dette kunne for eksempel vært et problem når det kommer til *incash* og *inkind*

benefits. Det er sannsynlig at disse to variablene kan være positivt lineært avhengig av hverandre. For eksempel ved at land enten gir mye eller lite støtte av begge type-
ne. En annen mulighet er at disse er negative lineært avhengige av hverandre, land
velger en av måtene å gi støtte på. I korrelasjonsmatrisen ser vi at forholdet mellom
de to variablene ikke er perfekt lineære. Det som derimot er nesten perfekt lineært,
er arbeidsledighetsrate for kvinner og for alle. Vi har derfor gjort et valg om å ikke
ta med begge disse variablene i regresjonsanalysen, og unngår denne utfordringen.
På bakgrunn av dette ser ikke forutsetningen om perfekt lineær korrelasjon ut til å
være brutt.

Den fjerde forutsetningen, MLR4, går ut på at det ikke er noen forhold i rest-
leddet u som har en påvirkning på både familieytelser og fertilitet. Dette gjør at
en enkel lineær regresjonsmodell som vi har brukt i ligning 1 blir unøyaktig. Vi
har funnet en modell som beskriver en korrelasjon, men fordi restleddet innehol-
der mange faktorer som påvirker størrelsen på disse variablene, blir nok den enkle
lineære regresjonsmodellen i vårt tilfelle unøyaktig. R-squared i den første lignin-
gen for regresjonsanalysene vi har gjort er også veldig lav, 0,026 for alle land og
0 for nordiske land. Selv om r-squared ikke er en fasit på hvorvidt modellen er en
god beskrivelse, kan det såpass lave tallet være en indikasjon på at vi har utelatt noe.

Det å legge til flere kontrollvariabler er et forsøk på å få en modell som gir en mer
nøyaktig beskrivelse av hvordan familieytelser påvirker fertilitet. Her blir det en ba-
lanse med å prøve å ikke bryte MLR2 og å ikke bryte MLR4. Med flere kontrollvari-
abler, gitt at vi mangler data, risikerer vi å bryte MLR2, men uten kontrollvariablene
bryter vi MLR4. Det at vi har lagt til kontrollvariablene betyr heller ikke at MLR4
er oppfylt. Det er lite realistisk å forutsette at vi har funnet alle mulige variabler
som ligger i restleddet og gjør modellen unøyaktig. Ligning 3 til 6 er likevel et forsøk
på å prøve å opprettholde forutsetningen. R-squared har også økt i disse ligningene,
men det trenger ikke nødvendigvis å skyldes at modellen er mer nøyaktig. Selv om
modellen ikke har beskrevet situasjonen bedre vil R-squared, på generell basis, øke
med flere variabler i modellen.

Den femte forutsetningen er homoskedastisitet. Dette kriteriet kan også være brutt.
I dataen er det mye spredning i observasjonene når andelen av BNP brukt på fami-
lieytelsen er liten. Når andelen blir større, er variasjonen mindre, og observasjonene
er nærmere det lineære estimatet. Utfordringen blir at tallene vi får for standard
error ikke blir riktige og at hypotesetestingen derfor også blir feil. Det at vi bruker en
estimering med faste effekter i regresjonsanalysen kan gjøre dette problemet mindre.

Den sjette forutsetningen, normalitet i restleddet, krever at MLR4 og MLR5 også
holder. Siden disse mest sannsynlig er brutt virker det rimelig å anta at denne
forutsetningen også er brutt.

6.2 OECD-land

6.2.1 Effekt av totale ytelser

Hovedfunnet i modellen for alle OECD-land er at totale ytelser har en positiv effekt på fertiliteten. Dette samsvarer med litteratur vi har sett på - både teorien fra Becker og studiene fra Rønsen og Skrede, Luci-Greulich og Thévenon, og Gauthier og Hatzius. Det er likevel mulig å problematisere denne konklusjonen. Selv om vi har funnet en korrelasjon mellom totale ytelser og fertilitet betyr ikke dette nødvendigvis at det er en kausalitet. Det sier heller ikke noe om hvorvidt det er riktig prioritering for myndighetene.

Et mulig problem med regresjonsmodellen vår er at sammenhengen mellom fertilitet og familieytelser også kan ligge andre veien. Altså at en økning i fertilitet fører til økte offentlige utgifter. Hvis det for eksempel er faste satser for barnetrygd per barn, så vil økt fertilitet føre til at utgiftene til myndighetene automatisk øker. I ligning 6 er problemet med en motsatt årsak-virking sammenheng i høyeste grad gjeldende fordi fertilitet og familieytelser er målt i samme år.

I ligning 1 til 5 ser vi på hvordan fertiliteten vil påvirkes året etter at familieytelsen er gitt. Fertilitetsraten i år $t+1$ kan ikke ha en innvirkning på familieytelsene i år t . Som vi ser i korrelasjonsmatrisen i figur 6 er likevel *next year fertility* og *current year fertility* nesten perfekt korrelert, og hvis dette skyldes at fertilitetsraten er ganske lik mellom år t og $t + 1$, kan det allikevel hende at vi i også i ligning 1-5 har funnet en motsatt kausalitet.

En annen svakhet er at dataen på fertilitet kun er aggregert på landsnivå i vår regresjonsanalyse. Dette gjør det vanskelig å se om barnetrygd har forskjellig effekt på fertilitet hos ulike sub-grupper i populasjonen. Det kunne blant annet vært interessant å se på i hvor stor grad barnetrygd har en effekt på fertilitet i ulike grupper, basert på inntekt eller utdanning. Man vil kunne tenke seg til at en økning i barnetrygden kanskje kan ha en høyere effekt på fertiliteten i lavinntektsfamilier enn familier med bedre råd.

Det er også verdt å merke seg at andre ordninger, slik som foreldrepermisjon, ikke er inkludert i datasettet. Likevel er en naturlig antakelse at dette er et gode som kan ha innvirkning på fertilitetsraten. I analysen har vi ikke gjort en vurdering av om andre ordninger som ikke er inkludert i OECD-dataene kan være mer effektive.

En utfordring når man sammenligner barnetrygdordninger mellom forskjellige land, er at ordningene er utformet forskjellig. Noen land kan for eksempel ha behovsprøvd trygd. Dette gjør at trygder kan ha ulik effekt på fertilitet, selv om det totale beløpet som blir betalt i trygd er det samme.

Vi har heller ikke foretatt noen vurdering av den samfunnsøkonomiske effektiviteten av en økning i familieytelser. Det kan være kostbart for et land å skulle bruke en høy prosentats av BNP på familieytelser. Dette kan også gjøre relevansen av resultatene våre lavere. Selv om vi finner at det er en positiv sammenheng mellom BNP og fertilitet, kan økningen i prosentandel av BNP man er nødt til å gjennomføre for å komme opp til ønsket fertilitetsnivå være så høy at det ikke er realistisk mulig for et land å gjennomføre. Siden modellen vår gir en dårlig predikasjon av hva som skjer ved en større endring i familieytelser, er det heller ikke sikkert det ville hatt en stor effekt i det hele tatt.

6.2.2 Effekt av *inkind* vs *incash*-ytelser

Når vi deler inn i *inkind* og *incash*, vil resultatet avhenge av hvor mange kontrollvariabler vi legger til. Ved en hypotesetest på et 5 prosent signifikansnivå, er det en forventet positiv effekt på fertilitet av *incash*-ytelser i alle ligningene, men kun en virkning av *inkind*-ytelser i ligning 5 og 6. Det er også forskjeller mellom ligningene i hvilken av ytelsene som har den største forventede virkningen på fertilitet. I ligning 2 og 4 har *incash*-ytelser en klart større innvirkning på fertiliteten enn *inkind*-ytelser. I ligning 5 og 6 hvor flere kontrollvariabler er tatt med, er det derimot *inkind*-ytelser som får den største effekten.

Den positive sammenhengen mellom familieytelser i *incash* og fertilitet vises også i flere av studiene vi har gjennomgått, blant annet Rønsen og Skrede (2008). Det er gode grunner til å tro at ytelser som trykker den økonomiske situasjonen til foreldre, kan ha en positiv virkning på fruktbarheten. På denne måten kan man gå økte kostnader i møte uten å risikere et større velferdstap, og det åpner også opp for at foreldre kan jobbe mindre uten at det innebærer en stor nedgang i inntekt.

Luci-Greulich og Thevenon har i tillegg til direkte kontantoverføringer, undersøkt sammenhengen mellom fruktbarhet og barnepass – som går inn under familiestøtte gjennom tjenester (*inkind*). De har funnet at denne typen *inkind*-ytelser gir en positiv effekt på fertiliteten. Våre resultater av effekten for *inkind*-ytelser er mer usikre, og samsvarer derfor ikke helt med deres funn. Vår definisjon av *inkind*-ytelser er likevel videre enn kun barnepass og det er derfor ikke helt sammenlignbart.

Som drøftet under evalueringen av forutsetningene er det fordeler og ulemper med de forskjellige ligningene. Isolert sett kan man betrakte ligning 5 og 6 som mer korrekte fordi de inneholder flere kontrollvariabler. Hvis vi derimot antar at vi bryter forutsetningen om tilfeldig utvalg, kan det heller være at ligning 3 og 4 gir et bedre bilde av den faktiske situasjonen.

Det virker logisk at *incash* kan ha en større effekt på fertiliteten enn *inkind* slik vi fant i ligning 3 og 4. Begge formene for støtte hjelper foreldre økonomisk, men *incash* gir større økonomisk frihet da midlene kan disponeres etter egne preferanser

og behov. Noen familier/foreldre vil antakelig få større nytte av *incash*-ytelser fordi midlene kan brukes på en måte som er optimalt tilpasset deres situasjon. *Inkind* støtte er derimot begrenset til mer spesifikke tjenester, som for eksempel barnehage. En familie som ikke ønsker å sende barnet i barnehage vil ikke få noe økt nytte hvis barnehage blir gratis. Det er også mulig å tenke at en mer direkte støtte gjennom *incash* kan ha en større effekt gitt at tjenester som barnehage allerede holder et tilfredsstillende nivå. Det kan hende at en ytterligere styrking av disse ordningene ikke er det som er mest avgjørende for å få barn.

Spørsmålet blir hvorfor *inkind*-ytelser får en større effekt enn *incash* i ligning 5 og 6. Det kan som sagt handle om at vi bryter forutsetningen om tilfeldig utvalg, men en annen mulig forklaring er at vi legger til en tidstrend. Ser man på utviklingen av *inkind* støtte over tid (se figur 2), så kan man observere at den jevnt over har steget, mens *incash* i større grad er flat. Det har derfor, i hvert fall i noen land, vært en trend hvor nivået på *inkind*-ytelser har steget. Når det kommer til fertilitet, har den holdt seg ganske jevn, men fra 2009 har det vært en liten nedgang i fertilitet samtidig som det har vært en økning i *inkind*-ytelser. Det er en mulighet for at den begrensede effekten vi finner av *inkind*-ytelser i ligning 1-4 skyldes at vi ikke har tatt hensyn til en avtagende tidstrend når det kommer til fertilitet. Denne får vi skilt ut i ligning 5 og 6 og vi får dermed et bilde av den faktiske effekten *inkind*-ytelser har.

Det er derfor vanskelig å gi noen fasit på hvilke av ligningene som er mest korrekte. Dette viser et generelt problem ved å bruke en regresjonsanalyse til å studere en sammenheng mellom to variabler. Det er vanskelig å ikke bryte noen av forutsetningene for modellen, og man kan bli nødt til å vurdere hvilke av forutsetningene som er minst viktige. Det er også enkelt å trekke konklusjoner basert på egen bias. Vi ønsker derfor ikke å gi noe konklusjon på hvilken ligning som er mest korrekt.

6.3 Nordiske land

Hovedfunnet i modellen for nordiske land er at totale ytelser kan ha en effekt på fertiliteten, men resultatene avhenger av om vi har med kontrollvariablene eller ikke. Uten kontrollvariabler fant vi ikke en effekt av totale ytelser på fertilitet. Når vi deler det inn *incash* og *inkind*, er det kun *incash* som har en signifikant effekt på fertiliteten.

Analysen av Norden viser at gjennomsnittlig fertilitet er betraktelig høyere sammenlignet med analysen gjennomført for OECD-landene. Tidligere forskning har vist at fertilitetstrender observert i de nordiske landene er relativt like. Mye kan tyde på at en felles nordisk modell, hvor fokus på likestilling på arbeidsmarkedet har blitt tillagt større vekt, har vært avgjørende for utarbeidelsen av generøse familieytelser. Dette kan ha gitt et betydelig positivt bidrag til fertilitet sammenlignet med andre OECD-land. Fertiliteten i nordiske land er i gjennomsnitt høyere enn fertiliteten i de andre OECD-landene i datasettet vårt.

Vi vil likevel påpeke at resultatet vi har fått i regresjonsanalysen for nordiske land virker urealistisk. I ligning 5 får vi at en økning i *incash* støtte som andel av BNP på ett prosentpoeng, vil gi en forventet økning i neste års fertilitet med 0,838 barn og i ligning 6 hvor vi har current year fertilitet, har vi en forventet økning på 1,002 barn. Et slik resultat vil gjøre det svært enkelt å drive familiepolitikk i de nordiske landene. Sannheten er nok heller at det er flere problemer med modellen vår som gjør at den passer dårlig for predikasjon av hva som vil skje med fertiliteten ved en økning i familieytelser, i hvert fall når det kommer til nordiske land.

Her vil nok både det at vi bryter mange av forutsetningene for modellen og at fertilitet kan spille inn på størrelsen på familieytelsene spille inn. Vi har funnet en korrelasjon, men ikke en kausalitet. I hvert fall ikke en kausalitet som viser at økte familieytelser er årsaken til økt fertilitet. I tillegg har vi mange færre observasjoner i de nordiske landene - noe som gjør at tilfeldigheter spiller en mye større rolle enn det gjorde når vi hadde alle OECD-landene.

7 Konklusjon

I konklusjonen ønsker vi å besvare de tre spørsmålene som ble stilt i innledningen.

Påvirker familieytelser fertiliteten i OECD-land?

I regresjonsanalysen har vi funnet en positiv sammenheng mellom totale familieytelser og fertilitet i OECD-land. Det ser derfor ut til at en økning i familieytelser i et OECD-land vil gi økt fertilitet. Dette resultatet styrkes av at vi har funnet denne sammenhengen både når vi ikke har tatt med kontrollvariabler, og når vi har tatt med flere kontrollvariabler. Resultatet svekkes av at flere av forutsetningene vi har gjort for regresjonsanalysen mest sannsynlig er brutt og at det finnes en mulighet for en motsatt kausalitet. Vi har heller ikke sett på forskjeller mellom sub-grupper i populasjonen - noe som kunne gitt et mer nøyaktig bilde.

Påvirkes fertiliteten av hvilke typer familieytelser som gis?

Vi har ikke funnet en klar forskjell mellom *inkind* vs *incash* ytelser. Resultatet avhenger av om man ønsker å legge størst vekt på hensynet om tilfeldig utvalg eller om at variabler i restleddet ikke påvirker resultatet. Det vi derimot har funnet er at resultatet for *incash*-ytelser er mer konsistent. Uavhengig av ligning og om vi ser på nordiske land eller OECD-land, har *incash*-ytelser en statistisk signifikant effekt på fertiliteten. Mens effekten av *inkind*-ytelser kun har en statistisk signifikant effekt (med et signifikansnivå på 5 prosent) i ligning 5 og 6 i regresjonsanalysen for alle OECD-land. Selv om vi ikke kan være sikre på at *incash*-ytelser har en større effekt enn *inkind*-ytelser på fertiliteten, har vi i hvert fall flere resultater som tyder på at det faktisk er en effekt.

Er det noen forskjeller i de nordiske landene sammenlignet med alle OECD-landene?

I de nordiske landene finner vi samme korrelasjon som i OECD-landene, ved at økte totale familieytelser gir økt fertilitet. Effekten av familieytelser er betydelig større i nordiske land enn når man ser på alle OECD-landene. Resultatet vi har funnet for de nordiske landene virker såpass urealistisk at vi antar at resultatet enten skyldes at forutsetningene for modellen brytes eller en motsatt kausalitet.

Referanser

- Becker, Gary S (1991). *A treatise on the family: Enlarged edition*. Harvard university press.
- Gauthier, Anne Helene og Jan Hatzius (1997). «Family benefits and fertility: An econometric analysis». I: *Population studies* 51.3, s. 295–306.
- Hart, Rannveig K og Øystein Kravdal (2020). «Fallende fruktbarhet i Norge». I: *Report*. Oslo: Folkehelseinstituttet.
- Krokedal, Linn (2023). *Rekordlav fruktbarhet i 2022*. <https://www.ssb.no/befolkning/fodte-og-dode/statistikk/fodte/artikler/rekordlav-fruktbarhet-i-2022>. Accessed on: May 3, 2023.
- Laland, Christian (2019). *Statsministerens nyttårstale 2019*. <https://www.example.com>. Accessed on: May 3, 2023.
- Luci-Greulich, Angela og Olivier Thévenon (2013). «The Impact of Family Policies on Fertility Trends in Developed Countries: L'influence des politiques familiales sur les tendances de la fécondité des pays développés». I: *European Journal of Population/Revue européenne de Démographie* 29, s. 387–416.
- Malthus, Thomas Robert (1872). *An Essay on the Principle of Population*.
- OECD (2023a). *Fertility rates*. <https://data.oecd.org/pop/fertility-rates.htm>. Accessed on: May 5, 2023.
- (2023b). *Find, compare and share the latest OECD data: charts, maps, tables and related publications . . .* <https://data.oecd.org/>. Accessed on: May 3, 2023.
- (2023c). *Foreign-born population*. Available from: <https://data.oecd.org/migration/foreign-born-population.htm>. Accessed on 13th May 2023.
- (2023d). *Gender wage gap*. Available from: <https://data.oecd.org/earnwage/gender-wage-gap.htm>. Accessed on 13th May 2023.
- (2023e). *Gross domestic product (GDP)*. Available from: <https://data.oecd.org/gdp/gross-domestic-product-gdp.htm>. Accessed on 13th May 2023.
- (2023f). *Infant mortality rates*. Available from: <https://data.oecd.org/healthstat/infant-mortality-rates.htm>. Accessed on 13th May 2023.
- (2023g). *Our global reach*. <https://www.oecd.org/about/members-and-partners/>. Accessed on: May 5, 2023.
- (2023h). *Population with tertiary education*. Available from: <https://data.oecd.org/eduatt/population-with-tertiary-education.htm>. Accessed on 13th May 2023.
- (2023i). *Together, we create better policies for better lives*. <https://www.oecd.org/about/>. Accessed on: May 3, 2023.
- (2023j). *Unemployment rate*. Available from: <https://data.oecd.org/unemp/unemployment-rate.htm>. Accessed on 13th May 2023.
- (2023k). *Working age population*. Available from: <https://data.oecd.org/pop/working-age-population.htm>. Accessed on 13th May 2023.
- Rønsen, Marit (2005). «Fruktbarhetsutviklingen i Norge». I.
- Rønsen, Marit og Kari Skrede (2008). «Fertility trends and differentials in the Nordic countries—Footprints of welfare policies and challenges on the road ahead». I: *Vienna yearbook of population research*, s. 103–123.
- (2010). «Can public policies sustain fertility in the Nordic countries? Lessons from the past and questions for the future». I: *Demographic research* 22, s. 321–346.

-
- Sobotka, Tomáš, Anna Matysiak og Zuzanna Brzozowska (2019). «Policy responses to low fertility: How effective are they». I: *United Nations Population Fund*.
- Stortinget (2022). *Statsbudsjettet 2023*. https://www.stortinget.no/globalassets/pdf/budsjettsiden/blaabok_2023.pdf. Accessed on: May 5, 2023.
- Ward, Andrew (2013). «Spurious correlations and causal inferences». I: *Erkenntnis* 78, s. 699–712.
- Wooldridge, Jeffrey M (2015). *Introductory econometrics: A modern approach*. Cengage learning.

Vedlegg

A Vedlegg 1 - Oversikt over variabler

Variabel	Beskrivelse
<i>CURRENT_YEAR_FERTILITY</i>	Fertiliteten målt i barn per kvinne, i inneværende år.
<i>NEXT_YEAR_FERTILITY</i>	Fertiliteten i landet det neste året.
<i>INCASH_CB</i>	Direkte pengestøtte til familier som andel av BNP.
<i>INKIND_CB</i>	Familiestøtte gjennom tjenester som andel av BNP.
<i>TOTAL_CB</i>	Total familiestøtte som andel av BNP.
<i>UNEMPLOYMENT_RATE_WOMAN</i>	Arbeidsledigheten for kvinner som prosent av arbeidsstyrken.
<i>UNEMPLOYMENT_RATE_ALL</i>	Arbeidsledigheten for begge kjønn som prosent av arbeidsstyrken.
<i>INFANT_MORTALITY</i>	Barnedødeligheten som død per 1000 vellykkede fødsler.
<i>GDP_PER_CAPITA</i>	BNP per innbygger målt i dollar delt på 1000.
<i>WORKING_AGE_POPULATION</i>	Andel av befolkningen som er i arbeidsstyrken (15 - 64 år).
<i>FOREIGN_BORN_POPULATION</i>	Andel av befolkningen født i et annet land.
<i>GENDER_WAGE_GAP</i>	Forskjellen mellom median lønn for menn og kvinner delt på median lønn for menn.
<i>PERCENT_WITH_TERTARY_EDUCATION</i>	Andelen folk mellom 25 og 34 år som har fullført den høyeste graden mulig for deres alder, forkortet til.
<i>TIME_TREND</i>	Tidstrend variabel som starter på 0 i år 1999, og øker med en for hvert år.

Tabell 1: Variabler med beskrivelse

B Vedlegg 2 - Oversikt over land

Land	Verdensdel
Australia	Oseania
Belgia	Europa
Canada	Amerika
Chile	Amerika
Danmark	Europa
Estland	Europa
Finland	Europa
Frankrike	Europa
Hellas	Europa
Irland	Europa
Island	Europa
Israel	Asia
Italia	Europa
Japan	Asia
Latvia	Europa
Litauen	Europa
Luxemburg	Europa
Mexico	Amerika
Nederland	Europa
New Zealand	Oseania
Norge	Europa
Østerrike	Europa
Polen	Europa
Portugal	Europa
Slovakia	Europa
Slovenia	Europa
Spania	Europa
Storbritannia	Europa
Sverige	Europa
Sveits	Europa
Sør-Korea	Asia
Tsjekkia	Europa
Tyskland	Europa
Tyrkia	Europa
Ungarn	Europa
USA	Amerika

Tabell 2: Land med verdensdel

C Vedlegg 3 - Oversikt over forkortelser

Variabelnavn	Forkortelse
<i>CURRENT_YEAR_FERTILTY</i>	fertility c y
<i>NEXT_YEAR_FERTILITY</i>	fertility n y
<i>INCASH_CB</i>	incash cb
<i>INKIND_CB</i>	inkind cb
<i>UNEMPLOYMENT_RATE_ALL</i>	unemployment rate
<i>INFANT_MORTALITY</i>	infant mortality
<i>GDP_PER_CAPITA</i>	gdp per capita
<i>WORKING_AGE_POPULATION</i>	working age pop
<i>FOREIGN_BORN_POPULATION</i>	foreign age pop
<i>GENDER_WAGE_GAP</i>	gender wage gap
<i>PERCENT_WITH_TERTARY_EDUCATION</i>	p with ter educ
<i>TIME_TREND</i>	time trend

Tabell 3: Forkortelser av variabler

