



Danmark og Polen: Forskjeller i avkastning av utdanning

Bacheloroppgave vår 2019

Innholdsfortegnelse

Kapittel 1: Innledning	3
1.1. <i>Motivering</i>	3
1.2. <i>Våre problemstillinger</i>	3
Kapittel 2: Teoretisk rammeverk og tidligere litteratur	4
2.1. <i>Innledning</i>	4
2.2. <i>Teoretisk rammeverk</i>	4
2.3. <i>Tidligere litteratur</i>	4
2.4. <i>Oppsummering</i>	5
Kapittel 3: Datamaterialet	6
3.1. <i>Innledning</i>	6
3.2. <i>Om datamaterialet</i>	6
3.3. <i>Deskriptiv statistikk for avhengig variabel og interessevariabel</i>	7
3.4. <i>Deskriptiv statistikk for kontrollvariabler</i>	9
3.5. <i>Oppsummering</i>	10
Kapittel 4: Økonometrisk modell	11
4.1. <i>Innledning</i>	11
4.2. <i>Valg av funksjonsform</i>	11
4.3. <i>Empirisk strategi</i>	12
4.4. <i>Oppsummering</i>	13
Kapittel 5: Empiriske resultater	14
5.1. <i>Innledning</i>	14
5.2. <i>Empiriske hovedresultater</i>	14
Øvrige resultater basert på tabell A1 (appendiks 1)	16
5.3. <i>Analyse av tilleggs spørsmål</i>	21
5.4. <i>Empiriske resultater av tilleggs spørsmål</i>	21
5.5. <i>Oppsummering</i>	23
Kapittel 6: Oppsummering og konklusjon	24
6.1. <i>Oppsummering og konklusjon</i>	24
6.2. <i>Svakheter ved modellen</i>	25
Referanser	26
Appendiks 1	27
Appendiks 2	28

Kapittel 1: Innledning

1.1. Motivering

Arbeidsinnvandring er et aktuelt tema, og i den forbindelse kan det vært interessant å se på forskjeller og likheter i lønnsavkastning av formell utdanning mellom et tidligere kommunistisk land i Øst-Europa, og et sosialdemokratisk land i Vest-Europa. Dersom man kan forvente store forskjeller i lønnen med samme utdanningsnivå, er det ikke rart om mange ønsker å arbeide i et annet land enn sitt hjemland. EU og EØS er samarbeidsavtaler som blant annet innebærer fri flyt av arbeidskraft innad i unionen, som betyr at store deler av Europa er del av samme arbeidsmarked. Dette gjør arbeidsinnvandring mulig (European Union 2019).

Vi har valgt kun to land for at analysen skal bli mer ryddig, og fordi det også er store forskjeller innad i Øst- og Vest-Europa. En fullstendig analyse av alle tidligere kommunistiske land i Øst-Europa mot alle land i Vest-Europa hadde nok gitt et mer helhetlig bilde, men datamaterialet vårt begrenser seg til et maksimum på 4 kommunistiske land. Vi har valgt Polen og Danmark fordi begge er flate, landbruksdominerte land. Begge er nå med i EU og arbeidsledigheten er per dags dato relativt lav i begge land. Vi har valgt to land med såpass sentrale likhetstrekk for at forskjellene mellom dem skal komme tydeligere frem (Lundbo 2019, Pihl 2019).

1.2. Våre problemstillinger

1. Gir formell utdanning samme lønnsavkastning i et vesteuropeisk land som Danmark og i et tidligere kommunistisk land som Polen?
2. I hvilken grad kan forskjellen i lønnsavkastning på utdanning mellom landene forklares av forskjeller i avkastning på arbeidserfaring, kjønn, kognitive ferdigheter, sektor, fulltidsarbeid, foreldres utdanningsnivå og utdanning gjennomført før 1989?

Tilleggsspørsmål:

1. Er avkastningen ulik for de som har tatt utdanning før og etter kommunismens fall?
2. Kan noe av forskjellen i avkastning på utdanning forklares av kjønnsforskjeller i arbeidslivet?
3. Kan noe av forskjellen i utbytte av utdanning forklares av ulik avkastning på deltidsarbeid?

Kapittel 2: Teoretisk rammeverk og tidligere litteratur

2.1. Innledning

I dette kapitlet skal vi introdusere rammeverket for analysen vår. Her introduserer vi variabler som vi mener kan ha påvirkning på lønnen, annet enn formell utdanning. På den måten kan vi senere isolere effekten utdanning har på lønn i henholdsvis Danmark og Polen. Vi har også inkludert tidligere litteratur om effekten av formell utdanning på lønn.

2.2. Teoretisk rammeverk

Vi skal se om formell utdanning gir forskjellig lønnsavkastning i Polen og Danmark. For å finne ut av dette må vi se på andre eventuelle faktorer som kan påvirke resultatet.

1. Erfaring. Dersom Polen og Danmark verdsetter erfaring ulikt, kan dette påvirke avkastningen av utdanning i de to landene.
2. Kognitive ferdigheter. Dersom kognitive ferdigheter lønnes ulikt i Polen og Danmark, kan dette gi forskjeller i utbytte av utdanningen.
3. Sektor. Det kan være at lønnsforskjellen mellom privat og offentlig sektor er ulik i Danmark og Polen. Dette kan påvirke avkastning av utdanning hvis lønn avhenger av hvilken sektor man ender opp i.
4. Stillingsandel. Dersom fulltidsarbeidende tjener relativt bedre eller dårligere enn deltidsarbeidende i Danmark og Polen, kan dette påvirke avkastning på utdanning.
5. Utdanningsnivå til foreldre. Dersom barn av høyere utdannede foreldre tjener relativt bedre i et av landene, kan dette påvirke avkastning på utdanning.
6. Kjønnforskjeller. Det kan være at kvinner tjener relativt dårligere enn menn i et av landene slik at dette påvirker avkastning på utdanning.
7. Alder. Det kan være store forskjeller i utdanningsnivå mellom ulike aldersgrupper. Det kan være at generasjonen som tok utdanning under kommunismen har en annen avkastning på utdanning enn de som har tatt det senere.

2.3. Tidligere litteratur

Klassisk økonomisk teori tilsier at høyere tilbud fører til lavere pris. Dersom mange tar høyere utdanning vil dette føre til økt tilbud av høye ferdigheter. Dersom man antar etterspørselen etter slike ferdigheter konstant, vil lønnen reduseres for høye utdanningsnivåer, for å klarere arbeidsmarkedet.

Samtidig er det verdt å nevne at industrier og sektorer som krever høye ferdigheter, ofte har høy verdiskaping. Siden hvert individ i teorien skal tjene en lønn lik sin marginale produktivitet, bør mer verdiskapende industrier gi høyere lønn. Dersom man antar at høyere ferdigheter i utgangspunktet gir høyere lønn enn lavere ferdighetsnivå, vil dette gapet reduseres ved at de “på toppen” får sin lønn redusert.

Dette betyr at et land med generelt høyere utdanningsnivå sannsynligvis har lavere avkastning på den utdanningen. (Riis og Moen 2016).

I en rapport skriver OECD om noen eksterne faktorer som påvirker avkastning på utdanning. Undersøkelsene deres viser at høy skattesats, trygge vilkår for arbeidsledige og høye skolepenger alle bidrar til å redusere avkastning på utdanning. Høy skattesats og høye skolepenger vil direkte redusere avkastningen ved å påføre arbeidere mer kostnader. Trygge vilkår for arbeidsledige vil redusere avkastningen ved at alternativkostnaden ved å ikke arbeide blir lavere. (Boarini og Strauss 2007)

Polen var et kommunistisk regime fra 1948 til 1989, og gjennomførte omfattende reformer for å nasjonalisere industrien og kollektivisere landbruket. (Encyclopædia Britannica 2019 (2)). Under det kommunistiske regimet var idealet likere arbeidere (Encyclopædia Britannica 2019 (1)). Dette tilsier at Polen hadde en jevnere lønnsfordeling, og en lavere avkastning på utdanning enn Danmark, i hvert fall før 1989. Vi er interesserte i å se om dette historiske likhetsprinsippet veier tyngre enn annen klassisk økonomisk teori, som tilsier at økonomier med lavere gjennomsnittlig utdanningsnivå bør lønne de (færre) med utdanning relativt bedre. (Kent 2018; Stoltz 2014).

2.4. Oppsummering

I dette kapitlet har vi sett på andre faktorer som potensielt kan påvirke lønnen, og dermed avkastning på utdanning. Dette inkluderer kjønnsforskjeller i arbeidslivet, kognitive ferdigheter, arbeidserfaring, stillingsprosent, sektor, aldersgruppe og foreldres utdanningsnivå. Det finnes også andre faktorer som ikke er direkte inkludert i modellen. Noen er skattesats, skolepenger og trygdeordninger. I tillegg kommer vi tilbake til noen andre faktorer i kapittel 6. Økonomisk teori tilsier at et land der mange tar utdanning vil ha lavere avkastning på den utdanningen.

Kapittel 3: Datamaterialet

3.1. Innledning

I dette kapittelet skal vi ta for oss datamaterialet for analysen. Her definerer vi alle variablene brukt i analysen. Vi gir også en oppsummering av sentral informasjon om alle variabler. Den deskriptive statistikken blir introdusert i form av tabeller og histogrammer for å visualisere tallene. Dette kan gi noen mulige tolkninger før de empiriske resultatene i kapittel 5.

3.2. Om datamaterialet

Datamaterialet er basert på den internasjonale undersøkelsen om lese- og tallforståelse i voksenbefolkningen (PIAAC) som er gjennomført i regi av OECD. Undersøkelsen pågikk fra august 2011 til mars/april 2012 i til sammen 24 land. Vår oppgave er basert på datamaterialet fra landene Danmark og Polen.

Alle variablene er kun definert for personer eldre enn 25, slik at tallene inkluderer de som mest sannsynlig har fullført utdanning.

Se tabell 1 og 2 for definisjon av variablene.

Tabell 1. Definisjon av interessevariabler.

Variabel	Definisjon
wage	timelønn i PPP-justerte amerikanske dollar
poland	=1 hvis arbeider fra Polen =0 hvis arbeider fra Danmark
educ	antall år utdanning

Tabell 2. Definisjon av kontrollvariabler.

Variabel	Definisjon
exper	antall år arbeidserfaring
lit	Lesetestscore som avvik fra gjennomsnittet*
num	Regnetestscore som avvik fra gjennomsnittet*
psl	Problemløsingscore som avvik fra gjennomsnittet*
pub	=1 hvis arbeider jobber i offentlig sektor =0 hvis arbeider jobber i privat sektor
fulltime	=1 hvis fulltidsarbeidende =0 hvis ikke fulltidsarbeidende
parent	=1 hvis minst én forelder har fullført høyere utdanning =0 hvis ingen foreldre har fullført høyere utdanning

fem	=1 hvis kvinne =0 hvis mann
communist**	=1 hvis eldre enn 55 =0 hvis yngre enn 55

*Testene er gjennomført av OECD i 2011/2012. Test i lesing innebærer lesing av lange og fyldige tekster, dessuten retoriske virkemidler. Test i regning innebærer løsning av komplekse oppgaver som ofte krever flere metoder. Test av problemløsning er også en test av IKT-kunnskaper. Testresultatene er standardisert med gjennomsnitt 0 og standardavvik 1.

**Hvis personen er eldre enn 55 har personen mest sannsynlig fullført utdanning før 1989 da Polen var et kommunistisk styrt land.

3.3. Deskriptiv statistikk for avhengig variabel og interessevariabel

I tabell 3 finnes deskriptiv statistikk for avhengig variabel, lønn, og interessevariabelen, utdanning, i de to landene Polen og Danmark.

Tabell 3. Deskriptiv statistikk for interessevariabel og avhengig variabel.

	Polen		Danmark	
	Lønn	Utdanning	Lønn	Utdanning
Gjennomsnitt	8.915268	12.80099	24.88045	13.18846
Minimum	2.05834	6	5.913182	6
Maksimum	38.42235	21	66.13428	20
Standardavvik	5.141913	3.101045	8.326811	2.747325
Antall observasjoner	1,986	4,226	4,056	6,224

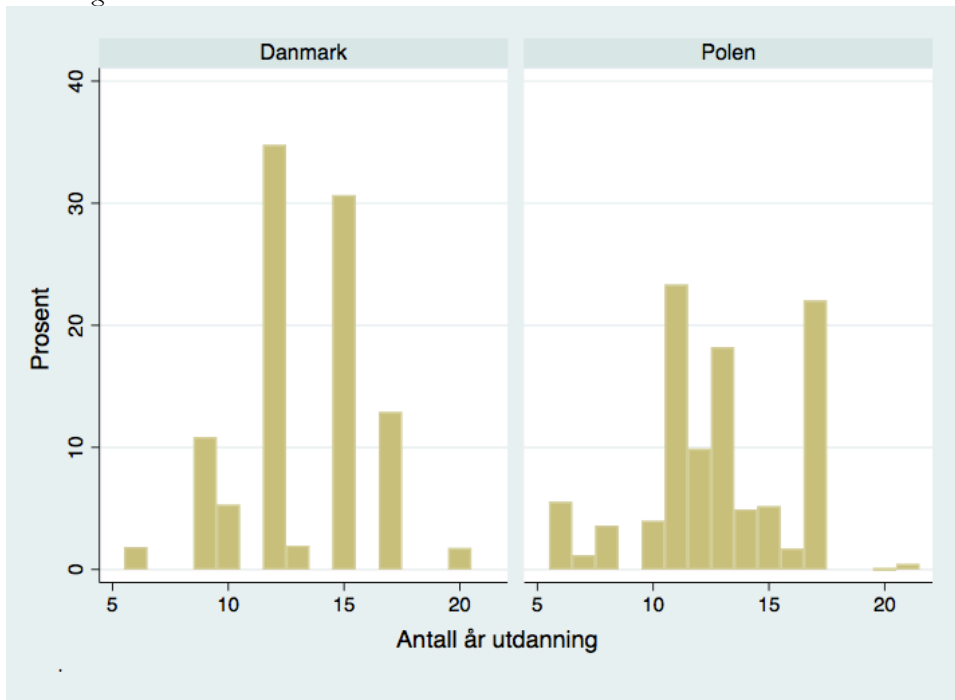
Kommentarer:

- Når vi sammenligner gjennomsnittlig timelønn i Polen og Danmark ser vi at det er over dobbelt så høy timelønn i Danmark som i Polen.
- Gjennomsnittet av antall utdanningsår er derimot ikke så forskjellig, men man kan likevel se at det generelle utdanningsnivået i Danmark er høyere enn i Polen.

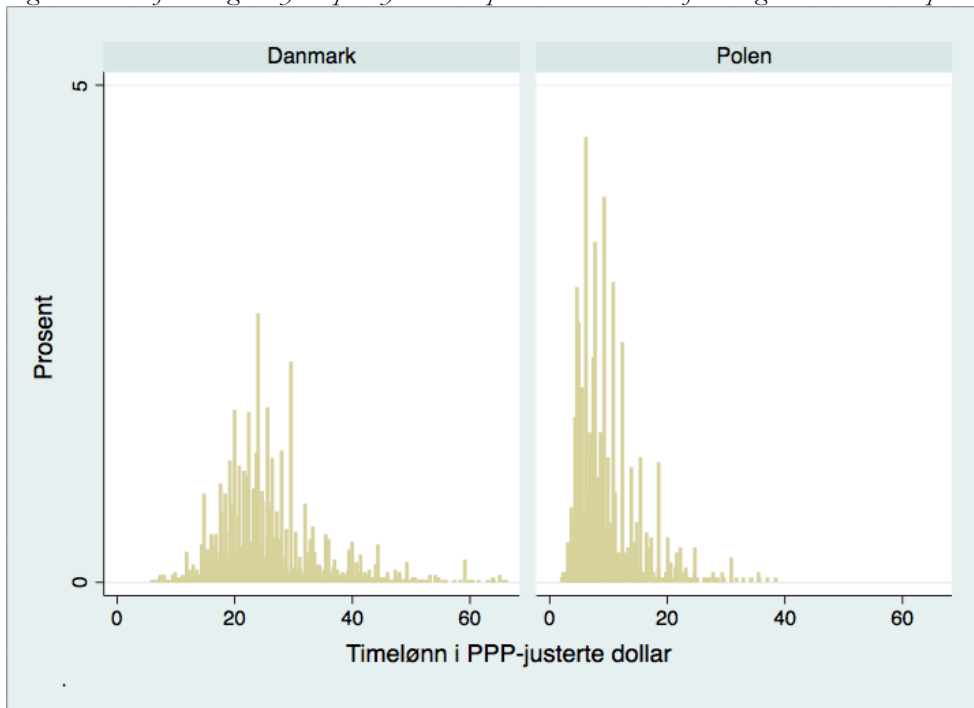
- Standardavviket i lønn er større i Danmark noe som kan tyde på at det er større lønnsbredning i Danmark enn i Polen. Dette stemmer overens med kommunistisk tankegang i Polen.

Se histogram nedenfor om utdanningsnivå og lønnsfordeling.

Figur 1. Utdanningsnivå. Høyden på søylene viser prosentandelen av befolkningen som har gjennomført hvert utdanningsnivå.



Figur 2. Lønnsfordeling. Høyden på søylene viser prosentandelen av befolkningen ved hver nivå på timelønn.



3.4. Deskriptiv statistikk for kontrollvariabler

Tabell 4. Deskriptiv statistikk for kontrollvariable.

	Polen				Danmark			
	Gj.snitt	Std.avvik	Min	Maks	Gj.snitt	Std.avvik	Min	Maks
Erfaring, antall år (exper)	17.873	12.336	0	47	26.572	12.722	0	47
Lese- forståelse (lit)*	-0.204	1.052	-4.654	3.741	-0.029	1.019	-4.633	2.774
Tall- forståelse (num)*	-0.155	1.048	-4.665	2.861	0.009	1.019	-4.069	3.264
Problem- løsning (psl)*	-0.225	1.056	-4.208	3.486	-0.072	1.00	-4.553	3.054
Offentlig sektor (%)**	27.09				38.7			
Privat sektor (%)**	72.91				61.23			
Foreldre høyere utdannet (%)	12.17				26.55			
Kvinner (%)	51.41				50.88			
Ferdig utdannet før 1989 (%)	22.95				34.97			
Fulltids- arbeider (%)	58.10				66.54			

*Variablene for hhv lese- og tallforståelse er standardisert med gjennomsnitt 0 og standardavvik 1 for hele datasettet. Siden vi har tatt vekk alle under 25 år viser tallene relativt avvik fra landsgjennomsnittet.

**En liten prosentandel av arbeiderne i undersøkelsen jobbet i ideelle organisasjoner. Siden dette ikke kan kategoriseres som verken offentlig eller privat er det fjernet fra datasettet.

Kommentarer:

- Det er betydelig høyere gjennomsnittlig arbeidserfaring i Danmark enn i Polen.
- Polen skårer dårligere enn Danmark i alle ferdighetstestene.
- I Danmark er det en høyere prosentandel som jobber i offentlig sektor enn i Danmark, men i begge land jobber de fleste i privat sektor.
- Andelen foreldre med høyere utdanning er større i Danmark.
- Andelen eldre (over 55) med utdanning i Danmark er høyere enn i Polen.
- Det er flere personer med fulltidsjobb i Danmark.

Mulig tolkning:

Det er høyere lønn i Danmark og nesten like mye utdanning i begge land. Etter å ha sett på kontrollvariablene kan dette kanskje ha noe å gjøre med at flere dansker jobber lenger og får mer jobberfaring. Siden ferdighetsnivået i Danmark er høyere, kan det også være at det er høyere kvalitet på utdanningen i Danmark enn i Polen. Ser også at andelen foreldre i Danmark med høyere utdanning er større, som kan føre til høyere lønnsnivå for barna deres.

3.5. Oppsummering

I dette kapitlet har vi definert alle variablene vi skal inkludere i modellen vår. Disse er utdanning og land, som er interessevariablene våre. I tillegg kontrollerer vi for erfaring, ferdigheter, sektor, stillingsandel, foreldres utdanning, kjønn og alder. Alt dette kan ha innflytelse på lønnen og dermed lønnsavkastningen av utdanning.

Vi ser at gjennomsnittslønnen i Danmark er høyere enn i Polen, men at landene har relativt likt utdanningsnivå, om noe høyere i Danmark.

Kapittel 4: Økonometrisk modell

4.1. Innledning

I kapittel fire introduserer vi den økonometriske modellen for analysen. Her forklarer vi valg av funksjonsform og empirisk strategi. Empirisk strategi viser alle versjoner av modellen vi har brukt. Den er bygd opp ledd for ledd for å vise den isolerte effekten av hver av kontrollvariablene.

4.2. Valg av funksjonsform

Diskusjonen i teorikapitlet (kapittel 2) tilsier at: $wage = f(educ, poland, x)$

der x representerer et sett andre faktorer som kan tenkes å påvirke lønnen.

Siden vi skal se på forskjellen i avkastning på utdanning i Danmark og Polen, inkluderer vi et samvariasjonsledd mellom utdanning og land. Vi får:

$$(i) \text{ wage}_i = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{educ}_i * \text{poland}_i + \beta_4 \text{poland}_i + \beta_i x_i + \varepsilon_i$$

Vi velger en log-lineær funksjonsform hvor $lwage$ representerer logaritmen til

timelønnen ($wage$). Denne funksjonen inneholder den avhengige variabelen lønn ($lwage$), variabelen utdanning ($educ$), samt et samvariasjonsledd ($educpoland$) som representerer effekten av å være utdannet i Polen.

x_i representerer alle kontrollvariablene som kommer i tillegg med tilhørende koeffisienter β_i . Det siste leddet, ε_i , er et stokastisk restledd som fanger opp alle virkninger fra variabler som ikke er eksplisitt inkludert i modellen.

Siden vi lurer på hvilket land som har høyest lønnsavkastning på utdanning blir $educpoland$ funksjonens interessevariabel. Den tilhørende koeffisienten, β_3 , viser hvor innflytelsesrik variabelen er på lønnen. Tolkning av β_3 : $\beta_3 * 100$ viser i prosent hvor mye et ekstra år med utdanning i Polen øker timelønnen mer enn i Danmark.

Har valgt en logaritme-lineær funksjon siden variablene blir estimert prosentvis mot lønnen. Måleenheten er dermed likegyldig. Dette gjør at vi om ønskelig kan estimere variablene på tvers av studier og land.

4.3. Empirisk strategi

Vi ønsker å estimere om lønnsavkastningen på formell utdanning er forskjellig i Polen og Danmark. For å kunne estimere den isolerte effekten av de ulike kontrollvariablene har vi delt opp modellen i 8 deler. Restleddet, ε_i , er ikke korrelert med de andre koeffisientene $\beta_1 - \beta_{14}$. Vi estimerer med en regresjonsanalyse og ved bruk av minste kvadraters metode, også kalt OLS-estimering. En regresjonsanalyse er en kvantitativ analyse innen statistikk som analyserer sammenhenger mellom en eller flere avhengige variabler, og en eller flere uavhengige variabler. OLS er en estimeringsmetode som minimerer summen av kvadrerte residualer, hvor residualene er de observerte verdiene minus de predikerte verdiene. Jo mindre summen er, jo nærmere sannheten er estimeringen.

Egenskaper ved OLS:

- Introduserer en klassisk, enkel regresjonsmodell: $y_i = \alpha + \beta_i x_i + \varepsilon_i$
- Presiserer forutsetningene om restleddet ε_i .
 - Klassisk situasjon: x-ene er ikke stokastiske.
 - $E(\varepsilon_i) = 0$, forventningen til residualleddet er lik null.
 - $Var(\varepsilon_i) = E(\varepsilon_i - E(\varepsilon_i))^2 = E(\varepsilon_i^2) = \sigma^2$, variansen til restleddet er konstant.
 - $Cov(\varepsilon_i, \varepsilon_j) = E((\varepsilon_i - E(\varepsilon_i))(\varepsilon_j - E(\varepsilon_j))) = 0$, ingen systematisk samvariasjon mellom restledd ved ulike observasjoner.
 - ε_i er normalfordelt.
- OLS-estimatoren for stigningstallet, parameteren β , er gitt ved: $b = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$, der \bar{x} er gjennomsnittet til observasjonene for x, og \bar{y} er gjennomsnittet for observasjonene for y.
- OLS-estimatoren for konstantleddet, α , er gitt ved: $a = \bar{y} - b\bar{x}$.

Modell:

Variant (i): $lwage_i = \beta_1 + \beta_2 educ_i + \beta_3 educ_i poland_i + \beta_4 poland_i + \varepsilon_i$

For Danmark: (i') $lwage_i = \beta_1 + \beta_2 educ_i + \varepsilon_i$

For Polen: (i'') $lwage_i = \beta_1 + \beta_4 + (\beta_2 + \beta_3) educ_i + \varepsilon_i$

Variant (ii): $lwage_i = \beta_1 + \beta_2 educ_i + \beta_3 educ_i poland_i + \beta_4 poland_i + \beta_5 exper_i + \beta_6 exper_i^2 + \varepsilon_i$

Variant (iii): $lwage_i = \beta_1 + \beta_2 educ_i + \beta_3 educ_i poland_i + \beta_4 poland_i + \beta_5 exper_i + \beta_6 exper_i^2 + \beta_7 lit_i + \beta_8 num_i + \beta_9 psl_i + \varepsilon_i$

Variant (iv): $lwage_i = \beta_1 + \beta_2 educ_i + \beta_3 educ_i poland_i + \beta_4 poland_i + \beta_5 exper_i + \beta_6 exper_i^2 + \beta_7 lit_i + \beta_8 num_i + \beta_9 psl_i + \beta_{10} pub_i + \varepsilon_i$

Variant (v): $lwage_i = \beta_1 + \beta_2 educ_i + \beta_3 educ_i poland_i + \beta_4 poland_i + \beta_5 exper_i + \beta_6 exper_i^2 + \beta_7 lit_i + \beta_8 num_i + \beta_9 psl_i + \beta_{10} pub_i + \beta_{11} fulltime_i + \varepsilon_i$

Variant (vi): $lwage_i = \beta_1 + \beta_2 educ_i + \beta_3 educ_i poland_i + \beta_4 poland_i + \beta_5 exper_i + \beta_6 exper_i^2 + \beta_7 lit_i + \beta_8 num_i + \beta_9 psl_i + \beta_{10} pub_i + \beta_{11} fulltime_i + \beta_{12} parent_i + \varepsilon_i$

Variant (vii): $lwage_i = \beta_1 + \beta_2 educ_i + \beta_3 educ_i poland_i + \beta_4 poland_i + \beta_5 exper_i + \beta_6 exper_i^2 + \beta_7 lit_i + \beta_8 num_i + \beta_9 psl_i + \beta_{10} pub_i + \beta_{11} fulltime_i + \beta_{12} parent_i + \beta_{13} fem_i + \varepsilon_i$

Variant (viii): $lwage_i = \beta_1 + \beta_2 educ_i + \beta_3 educ_i poland_i + \beta_4 poland_i + \beta_5 exper_i + \beta_6 exper_i^2 + \beta_7 lit_i + \beta_8 num_i + \beta_9 psl_i + \beta_{10} pub_i + \beta_{11} fulltime_i + \beta_{12} parent_i + \beta_{13} fem_i + \beta_{14} communist_i + \varepsilon_i$

(i) viser kun en grov estimering av lønnsavkastningen på utdanning i Danmark og Polen. Første ledd er et konstantledd, som kan tolkes som lønnen man får uansett utdanning. Ledd to (*educ*) viser hvor stor avkastningen av formell utdanning er i Danmark. For å få avkastning på utdanning i Polen må man i tillegg ta hensyn til ledd tre (*educpoland*), som adderes med *educ* (se i’). Ledd fire kontrollerer for effekten av å være polsk på lønnen. For Danmark blir denne null. Man kan gjennomføre t-tester der man kontrollerer om parameterne er statistisk signifikante. Dette kan også gjøres for variablene kontrollert for senere. Se kapittel 5.

(ii) kontrollerer i tillegg for effekten av arbeidserfaring på lønnen. Vi har også kontrollert med arbeidserfaring kvadrert for å fange opp at det ikke nødvendigvis er en lineær effekt av arbeidserfaring på lønn. F.eks. hjelper det ikke å ha 50 års erfaring som gymlærer hvis du er 100 år gammel, de vil nok ansette noen yngre.

(iii) kontrollerer i tillegg for kognitive ferdigheter. Vi har inkludert både lese- og regneferdigheter, samt IKT-problemløsning.

(iv) kontrollerer for sektor. Det vil si effekten av å jobbe i privat eller offentlig sektor.

(v) kontrollerer for effekten av å være fulltidsarbeidende.

(vi) kontrollerer for effekten av å ha høyere utdannede foreldre.

(vii) kontrollerer for kjønn, det vil si effekten av å være kvinne på lønn.

(viii) kontrollerer for hvor mange som tok utdanning under kommunisttiden.

4.4. Oppsummering

I dette kapitlet har vi gått igjennom den økonometriske modellen. Vi har valgt å bruke en log-lineær funksjon for å kunne analysere på tvers av land. Vi estimerte modellen i 8 deler med minste kvadraters metode. Hver kontrollvariabel har sin isolerende effekt som påvirker avkastningen av utdanning. Modellen kan nå gi oss de empiriske resultatene som kommer i neste kapittel.

Kapittel 5: Empiriske resultater

5.1. Innledning

I dette kapitlet skal vi ta for oss resultatene av regresjonsanalysen vi har gjennomført og hva slags implikasjoner disse har. Vi tar først for oss hovedproblemstillingen om forskjell i lønnsavkastning på utdanning i Danmark og Polen, der vi, én etter én kontrollerer for andre faktorer. Vi tar så for oss tilleggsspørsmålene om kommunisme, kjønn og fulltidsarbeid.

5.2. Empiriske hovedresultater

I tabell 5 finnes hovedresultatene fra den økonometriske analysen.

Tabell 5. Estimert lønnsforskjell basert på likninger (i)-(viii). Fullstendige estimeringsresultater i Appendixs tabell A1. Estimerte standardavvik i parentes. Estimeringsmetode: OLS.

	i	ii	iii	iv	v	vi	vii	viii
Variable	lwage	lwage	lwage	lwage	lwage	lwage	lwage	lwage
educ	0.0463 (0.00216)	0.0519 (0.00213)	0.0402 (0.00234)	0.0441 (0.00238)	0.0448 (0.00238)	0.0438 (0.00241)	0.0448 (0.00239)	0.0448 (0.00241)
educpoland	0.0293 (0.00348)	0.0338 (0.00341)	0.0342 (0.00437)	0.0358 (0.00435)	0.0341 (0.00435)	0.0341 (0.00432)	0.0363 (0.00432)	0.0364 (0.00434)
poland	-1.517 (0.0484)	-1.498 (0.0474)	-1.452 (0.0647)	-1.482 (0.0644)	-1.456 (0.0644)	-1.487 (0.0642)	-1.487 (0.0642)	-1.488 (0.0645)
Kontroll for erfaring	nei	ja	ja	ja	ja	ja	ja	ja
Kontroll for ferdigheter	nei	nei	ja	ja	ja	ja	ja	ja
Kontroll for sektor	nei	nei	nei	ja	ja	ja	ja	ja
Kontroll for fulltid	nei	nei	nei	nei	ja	ja	ja	ja
Kontroll for høyere utdannede foreldre	nei	nei	nei	nei	nei	ja	ja	ja
Kontroll for kjønn	nei	nei	nei	nei	nei	nei	ja	ja
Kontroll for alder >55	nei	nei	nei	nei	nei	nei	nei	ja
Konstant	2.538 (0.0299)	2.152 (0.0341)	2.275 (0.0377)	2.246 (0.0377)	2.310 (0.0394)	2.306 (0.0393)	2.353 (0.0392)	2.353 (0.0392)
Antall obs.	5,901	5,879	4,587	4,584	4,584	4,539	4,539	4,539
R ²	0.706	0.727	0.708	0.712	0.714	0.715	0.721	0.721

Kommentarer:

Estimert forskjell i lønnsavkastning på utdanning mellom Danmark og Polen varierer mellom 2,9 (i) og 3,6 (viii) prosent. Det vil si at et års ekstra utdanning i Polen gir rundt 3,4¹ prosent høyere timelønn enn i Danmark. Dette kommer i tillegg til en estimert 4,5² prosent høyere timelønn for hvert ekstra år utdanning i begge land.

Dersom resultatene er signifikante vil et ekstra år utdanning i Polen tilføre timelønnen nesten 8 prosent, mens det samme året i Danmark øker timelønnen kun med 4,5 prosent.

For å se om resultatene er statistisk signifikante gjennomfører vi en t-test der vi tester hypotesene $H_0: \beta_3 = 0$ mot $H_1: \beta_3 \neq 0$. T-testen kan finnes i sin helhet i appendiks 2. Her tester vi sannsynligheten for at lønnsavkastningen av utdanning er lik i Danmark og Polen. Vi må forkaste nullhypotesen i alle modeller, og finner en overveiende sannsynlighet for at lønnsavkastning av utdanning er *ulike* i de to landene.

Vi må også teste om utdanning har betydning for timelønn i seg selv. Vi gjennomfører derfor enda en t-test. I modell (i) er et års ekstra utdanning estimert til å gi 4,6 prosent høyere timelønn. Vi tester $H_0: \beta_2 = 0$ mot $H_1: \beta_2 \neq 0$ i alle modeller. T-verdien er langt over kritisk verdi ved 5 prosent signifikansnivå, slik at vi må forkaste nullhypotesen i alle modeller om at formell utdanning er uten betydning for timelønn.

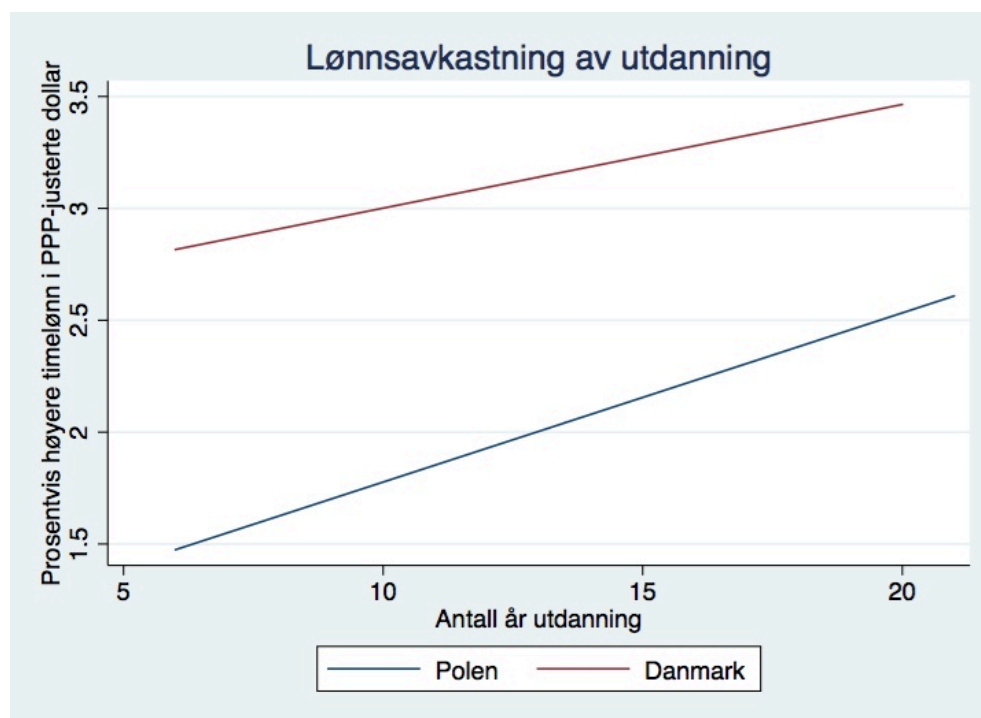
Vi kan dermed klart si at utdanning har betydning for timelønn og at lønnsavkastningen på utdanning i Polen er *høyere* enn i Danmark.

Samtidig ser vi at timelønnen i Polen er estimert til å være mellom 145,2 (iii) og 151,7 (i) prosent *lavere* enn i Danmark. Vi tester om disse resultatene er statistisk signifikante. Også her brukes en t-test, der vi tester $H_0: \beta_4 = 0$ mot $H_1: \beta_4 \neq 0$. I alle modeller får vi t-verdier langt under kritisk verdi på -1,96. Det vil si at det er en overveiende sannsynlighet for at lønnsnivået er lavere i Polen. Det vil si at selv med høy utdanning, får du lavere lønn i Polen enn i Danmark. Dette tilsier at lønnsforskjellene mellom Polen og Danmark er større for lave utdanningsnivåer, se figur 3.

¹ Gjennomsnitt av resultatene for variabelen *educpoland* i likning (i)-(viii).

² Gjennomsnitt av resultatene for variabelen *educ* i likning (i)-(viii).

Figur 3. Lønnsavkastning av utdanning i Danmark og Polen uten kontrollvariabler.



Øvrige resultater basert på tabell A1 (appendiks 1)

Kan forskjellen i lønnsavkastning forklares av systematisk ulik lengde på arbeidserfaring i Danmark og Polen?

Først skal vi undersøke om arbeidserfaring har betydning for timelønn. Igjen må vi teste om effekten av arbeidserfaring på timelønn er statistisk signifikant. I modell (ii)-(viii) har vi inkludert to ledd med arbeidserfaring (se kapittel 4.3. Empirisk strategi om hvorfor). Vi må derfor teste $H_0: \beta_5 = \beta_6 = 0$ mot $H_1: \text{minst én av } \beta_5 \text{ og } \beta_6 \neq 0$ i modell (ii).

Til dette må vi bruke en F-test fordi vi tester hypoteser om flere parametere samtidig.

I H_0 har vi to restriksjoner ($h=2$) på den underliggende modellen, altså kreves det at minst 2 av parameterne er ubetydelige for timelønnen.

Vi lar $SSR(ur)$ være residualkvadratsummen for modellen uten restriksjoner, og $SSR(r)$ være residualkvadratsummen for modellen pålagt restriksjoner under H_0 . Dersom vi får en stor økning i residualkvadratsummen (SSR) når restriksjonen pålegges, tyder det på at restriksjonen er ugyldig.

Vi velger et signifikansnivå på 5 prosent.

Vi får følgende testobservator:

$$TS = \frac{(SSR(r) - SSR(ur))/h}{SSR(ur)/(n - k)} \sim F(h, n - k)$$

Vi har 6 parametere (k), 2 restriksjoner under H_0 (h) og 5879 observasjoner (n). $SSR(ur)$ er 688,61 og $SSR(r)$ er 744,21. Vi får da:

$$F(2, 5873) = 227,86$$

Dette er langt over kritisk verdi på 2,99 slik at vi med stor sikkerhet kan si at arbeidserfaring har betydning for lønn.

Vi ser oss også nødt til å teste om effekten av arbeidserfaring er lineær eller ikke. I kapittel 4.3 introduserte vi to ledd med erfaring fordi antakelsen er at et ekstra års erfaring på et tidspunkt ikke gir meravkastning. Ett års ekstra arbeidserfaring gir i modell (ii) tilnærmet lik $100 * (0.0213 + 2 * (-0.000296) * exper)$. Vi tester en hypotese om at arbeidserfaring er lineær ved å teste $H_0: \beta_6 = 0$ mot $H_1: \beta_6 \neq 0$. Får en t-verdi på -9.72, som er langt under noe kritisk verdi. Må forkaste nullhypotesen om at effekten av arbeidserfaring er lineær.

Eksempel: Antar initial arbeidserfaring på 10 år ($exper=10$). Et års ekstra arbeidserfaring etter 10 år vil med våre estimerte koeffisienter i modell (ii) innebære

$$100 * (0.0213 + 2 * (-0.000296) * 10) = 1.5\% \text{ høyere lønn.}$$

Estimatene innebærer også at marginalavkastningen av utdanning er null ved 36 års arbeidserfaring, og negativ etter det.

Vi kan dermed si med stor sikkerhet at effekten av arbeidserfaring ikke er lineær.

Fra tabell 3 ser vi at modell (ii) estimerer en høyere lønnsavkastning på utdanning i Polen enn det estimert i modell (i). Det vil si at, kontrollert for arbeidserfaring, blir forskjellen i avkastning på utdanning i Polen og Danmark større. Vi så i kapittel 3 at det er flere med mer arbeidserfaring i Danmark. Hadde det ikke vært for at Polen har færre med arbeidserfaring som trekker lønnsavkastningen ned, ville forskjellene mellom Danmark og Polen vært større.

Vi kan dermed si at det er sannsynlig at man får høyere avkastning av utdanningen sin i Polen med samme erfaring som i Danmark. Likevel kan ikke forskjellen i lønnsavkastning på utdanning i de to landene forklares av erfaring fordi dette trekker i en annen retning.

Kan forskjellen i lønnsavkastning forklares av systematisk ulike kognitive ferdigheter i Danmark og Polen?

Vi skal nå teste om timelønnen påvirkes av kognitive ferdigheter, gitt formell utdanning, erfaring og land. De estimerte verdiene på β_6 , β_7 og β_9 er henholdsvis 0,0103; 0,0561 og 0,0254 i modell (iii). De estimerer virkningen av ferdigheter i henholdsvis lesing, regning og problemløsning på timelønn. Siden testresultatene er standardisert med gjennomsnitt 0 og standardavvik 1, vil verdiene på *lit* (lesing), *num* (regning) og *psl* (IKT-problemløsning) gi relativt avvik fra

gjennomsnittet. Om man scorer lavere enn gjennomsnittet vil ferdighetene dermed slå negativt ut i lønn siden de estimerte parameterne foran er positive. Motsatt om man scorer bedre enn gjennomsnittet.

Eksempel: Dersom en person er 1 enhet (standardavvik) bedre enn gjennomsnittet på alle tester vil dette gi en estimert $0.0103 * 1 + 0.0561 * 1 + 0.0254 * 1 = 0.0918 = 9,18$ prosent høyere timelønn.

Dersom en person er 1 enhet (standardavvik) dårligere enn gjennomsnittet på alle tester vil dette gi $0.0103 * (-1) + 0.0561 * (-1) + 0.0254 * (-1) = -0.0918$, altså en estimert 9,18 prosent *lavere* timelønn.

Nå skal vi teste om resultatene våre er statistisk signifikante. Siden vi tester hypoteser om flere parametere samtidig, må vi igjen bruke en F-test.

Tester $H_0: \beta_7 = \beta_8 = \beta_9 = 0$ mot $H_1: \text{minst én av } \beta_7, \beta_8 \text{ og } \beta_9 \neq 0$ i modell (iii).

Får at: $F(3, 4578) = 71,62$; som er langt høyere enn kritisk verdi på 2,6. Vi kan dermed klart forkaste nullhypotesen om at ferdigheter er uten betydning for timelønn.

Samtidig ser vi at β_3 , som estimerer forskjellen i lønnsavkastning på utdanning, øker marginalt i modell (iii). Dette tyder på at høyere ferdigheter i Danmark bidrar til å dempe forskjellene mellom Danmark og Polen i lønnsavkastning. Likevel er dette en så liten endring at ferdigheter sannsynligvis har lite å si for forskjellen i avkastning.

Kan forskjellen i lønnsavkastning forklares av systematisk større forskjell mellom offentlig og privat sektor i Danmark og Polen?

Vi skal nå teste om timelønnen påvirkes av hvilken sektor et individ jobber i, gitt formell

utdanning, erfaring, land og kognitive ferdigheter. β_{10} måler effekten av å være ansatt i offentlig sektor. Fra tabell A1 får vi en estimert verdi på -0,0778 for denne parameteren i modell (iv).

Dette innebærer at personer i offentlig sektor tjener en estimert 7,78 prosent *mindre* enn de med lik lengde på utdanning, samme ferdigheter og samme erfaring i privat sektor.

Må nå teste om resultatet er statistisk signifikant. I dette tilfellet må vi bruke en t-test, og teste

$H_0: \beta_{10} = 0$ mot $H_1: \beta_{10} \neq 0$. Får en t-verdi på -7.91. Dette er igjen lavere enn kritisk verdi på -1,96, og vi må forkaste nullhypotesen om at timelønn er uavhengig av sektor.

Samtidig ser vi at forskjellen i avkastning på utdanning øker betydelig i modell (iv) til 3,58

prosent. Dette er høyere enn i modell (i)-(iii). Dette antyder at sektorforskjeller mellom

Danmark og Polen er med på å trekke forskjellene mellom dem ned. Dette antyder at man får relativt mindre igjen på utdanningen sin dersom man jobber i offentlig sektor i Polen.

Sektorforskjeller kan dermed være med på å redusere avkastningen på utdanning i Polen, men kan ikke forklare den underliggende forskjellen mellom Danmark og Polen.

Kan forskjellen i lønnsavkastning forklares av systematisk forskjell mellom fulltids- og deltidsarbeidende i Danmark og Polen?

Vi skal først teste om timelønn blir påvirket av stillingsprosent, gitt formell utdanning, erfaring, ferdigheter og sektor. I modell (v) er effekten av å være fulltidsarbeider på timelønn estimert til å være -0,0866. Det vil si at en person med fulltidsjobb tjener en estimert 8,66 prosent mindre i timen enn en med deltids- eller annen ikke-fulltidsjobb. Tester om resultatet er statistisk signifikant. Tester $H_0: \beta_{11} = 0$ mot $H_1: \beta_{11} \neq 0$. Får en t-verdi på -5.48, som igjen er lavere enn kritisk verdi ved 5 prosent signifikansnivå. Må forkaste nullhypotesen om at stillingsprosent er uten betydning for timelønn.

Ser også at avkastning på utdanning i Polen nå er estimert til 3,41% høyere enn i Danmark. Dette er en lavere avkastning enn estimert i modell (iii) og (iv). Dette antyder at fulltidsarbeid i Polen fører til en relativt høyere avkastning på utdanning enn i Danmark. Altså ser det ut til at noe av forskjellen i avkastning på utdanning i Danmark og Polen skyldes forskjell i relativ timelønn for fulltidsarbeidende og deltidsarbeidende. Mer om dette i kapittel 5.4.

Kan forskjellen i lønnsavkastning forklares av systematisk ulik andel høyere utdannede foreldre i Danmark og Polen?

Vi skal nå teste om en arbeiders timelønn blir påvirket av om foreldrene har tatt høyere utdanning eller ikke. I modell (vi) er effekten av foreldres høyere utdanning estimert til å være 0,0263. Det vil si at det å ha minst én høyere utdannet forelder gir en estimert 2,63 prosent høyere timelønn. Vi må teste om resultatet er statistisk signifikant. Tester

$H_0: \beta_{12} = 0$ mot $H_1: \beta_{12} \neq 0$. Får en t-verdi på 2,38. Med signifikansnivå på 5% og tosidig test, må t-verdien være høyere enn kritisk verdi på 1,96 for å kunne forkaste nullhypotesen. Vi kan forkaste nullhypotesen om at høyere utdannede foreldre er uten betydning for timelønn.

Ser nå på hvordan dette påvirker forskjellen i avkastning på utdanning. I modell (vi) er avkastning på utdanning i Polen estimert til å være 3,41 prosent høyere enn i Danmark, altså helt likt som i modell (v). Dette tilsier at foreldres utdanningsnivå *ikke* kan forklare forskjellen i avkastning på utdanning i Danmark og Polen.

Kan forskjellen i lønnsavkastning forklares av systematisk ulik likestilling i arbeidslivet i Danmark og Polen?

Skal nå teste om timelønn blir påvirket av kjønn, gitt utdanning, nasjonalitet, erfaring, ferdigheter, sektor, stillingsprosent og foreldres utdanning. I modell (vii) er effekten av å være kvinne estimert til å være -0,100. Det vil si at det å være kvinne gir en estimert 10% lavere timelønn. Tester om resultatet er statistisk signifikant. Tester $H_0: \beta_{13} = 0$ mot $H_1: \beta_{13} \neq 0$. Får en t-verdi på -10,16, som er langt lavere enn kritisk verdi på -1,96. Vi må forkaste nullhypotesen om at kjønn er uten betydning for timelønn.

Ser nå på hvordan avkastningen på utdanning i Polen versus Danmark påvirkes av at vi har kontrollert for kjønn. Avkastning på utdanning i Polen er i modell (vii) estimert til å være 3,63 prosent høyere enn i Danmark. Dette er et høyere estimat enn i modell (i)-(vi). Gitt samme utdanningsnivå, erfaring og ferdigheter, ser det ut til at kvinner trekker avkastningen på utdanning ned i Polen. Samtidig ser vi at effekten av å være kvinne generelt er veldig negativ. Mer om dette i kapittel 5.4. Vi kan dermed si at forskjellen i avkastning på utdanning ikke skyldes kjønnsforskjeller. Siden polske kvinner får lavere relativ lønnsavkastning enn danske kvinner, bidrar dette til å dempe forskjellene i avkastning mellom landene.

Kan forskjellen i lønnsavkastning forklares av systematisk forskjell i antall eldre arbeidere i Danmark og Polen?

Skal først teste om timelønn blir påvirket av om man er ferdig utdannet før 1989 eller ikke. I Polen tilsier dette at man er utdannet under kommunisttiden. Mer om dette i kapittel 5.4. I modell (viii) er effekten av å være over 55 år estimert til å være 0,00263, altså vil en alder over 55 gi en estimert 0,26% høyere timelønn. Tester om dette er statistisk signifikant. Tester $H_0: \beta_{14} = 0$ mot $H_1: \beta_{14} \neq 0$. Får en t-verdi på 0.15. Kan forkaste nullhypotesen dersom t-verdien er høyere enn kritisk verdi på 1,96 (ved 5% signifikansnivå). Ser at vi ikke kan forkaste nullhypotesen. Det vil si at alder over 55 sannsynligvis er uten betydning for timelønn. Ser også at resultatene for avkastning på utdanning i Polen er lite endret fra modell (vii), så vi kan med relativt stor sikkerhet si at alder over 55 har lite å si for forskjellen i avkastning på utdanning i Danmark og Polen.

5.3. Analyse av tilleggsspørsmål

Vi ønsker å besvare de tre tilleggsspørsmålene om kommunisme under utdanning, kjønn og stillingsandel har betydning for forskjellen i avkastning på utdanning i Polen og Danmark. Vi behandler dem hver for seg.

(1): Er avkastningen ulik for de som har tatt utdanning før og etter kommunismens fall (1989)?

Estimerer (viii) og legger til et interaksjonsledd mellom *poland* og *communist*:

$$\begin{aligned}(ix) \text{ lwage}_i = & \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{educ}_i \text{poland}_i + \beta_4 \text{poland}_i + \beta_5 \text{exper}_i + \beta_6 \text{exper}_i^2 + \beta_7 \text{lit}_i \\ & + \beta_8 \text{num}_i + \beta_9 \text{psl}_i + \beta_{10} \text{pub}_i + \beta_{11} \text{fulltime}_i + \beta_{12} \text{parent}_i + \beta_{13} \text{fem}_i \\ & + \beta_{14} \text{communist}_i + \beta_{15} \text{communist}_i \text{poland}_i + \varepsilon_i\end{aligned}$$

(2): Kan noe av forskjellen i avkastning på utdanning forklares av kjønnsforskjeller i arbeidslivet?

Estimerer (viii) og legger til et interaksjonsledd mellom *poland* og *fem*:

$$\begin{aligned}(ix) \text{ lwage}_i = & \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{educ}_i \text{poland}_i + \beta_4 \text{poland}_i + \beta_5 \text{exper}_i + \beta_6 \text{exper}_i^2 + \beta_7 \text{lit}_i \\ & + \beta_8 \text{num}_i + \beta_9 \text{psl}_i + \beta_{10} \text{pub}_i + \beta_{11} \text{fulltime}_i + \beta_{12} \text{parent}_i + \beta_{13} \text{fem}_i \\ & + \beta_{14} \text{communist}_i + \beta_{15} \text{fem}_i \text{poland}_i + \varepsilon_i\end{aligned}$$

(3): Kan noe av forskjellen i utbytte av utdanning forklares av ulik avkastning på deltidsarbeid?

Estimerer (viii) og legger til et interaksjonsledd mellom *fulltime* og *poland*:

$$\begin{aligned}(ix) \text{ lwage}_i = & \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{educ}_i \text{poland}_i + \beta_4 \text{poland}_i + \beta_5 \text{exper}_i + \beta_6 \text{exper}_i^2 + \beta_7 \text{lit}_i \\ & + \beta_8 \text{num}_i + \beta_9 \text{psl}_i + \beta_{10} \text{pub}_i + \beta_{11} \text{fulltime}_i + \beta_{12} \text{parent}_i + \beta_{13} \text{fem}_i \\ & + \beta_{14} \text{communist}_i + \beta_{15} \text{fulltime}_i \text{poland}_i + \varepsilon_i\end{aligned}$$

Kommentarer:

Når alle dummyer er lik null vil vi se referansegruppen som er:

- Dansk, mann, jobber i privat sektor, jobber ikke fulltid, har foreldre uten høy utdanning og er mellom 25 og 55 år.
- I referansegruppa blir β_2 tolket som avkastningen av utdanning på lønnen for en danske som ikke har tatt utdanning før 1989. Hvis det er snakk om en polsk person har vi to interaksjonsledd. β_3 viser avkastningen av utdanning hvis du er polsk.

(ix) β_{15} tolker effekten av å ha tatt utdanning i polen under kommunisttiden.

(x) β_{16} tolker effekten på lønnen av å være en kvinne fra Polen.

(xi) β_{17} tolker effekten på lønnen av å jobbe fulltid i Polen.

5.4. Empiriske resultater av tilleggsspørsmål

I tabell 6 finnes resultatene av regresjonsanalysen av tilleggsspørsmålene.

Tabell 6. Resultater av regresjonsanalysen for likninger (ix)-(xi).

VARIABLES	(ix) lwage	(x) lwage	(xi) lwage
educ	0.0448 (0.00241)	0.0448 (0.00241)	0.0442 (0.00240)
educpoland	0.0364 (0.00435)	0.0394 (0.00440)	0.0345 (0.00433)
poland	-1.490 (0.0648)	-1.487 (0.0644)	-1.241 (0.0751)
exper	0.0245 (0.00167)	0.0245 (0.00165)	0.0244 (0.00165)
exper_sq	-0.000335 (3.70e-05)	-0.000332 (3.67e-05)	-0.000336 (3.66e-05)
lit	0.0235 (0.0114)	0.0244 (0.0114)	0.0235 (0.0113)
num	0.0300 (0.0107)	0.0298 (0.0107)	0.0321 (0.0107)
psl	0.0248 (0.00832)	0.0241 (0.00831)	0.0229 (0.00829)
pub	-0.0588 (0.0100)	-0.0619 (0.0100)	-0.0608 (0.01000)
fulltime	-0.0997 (0.0158)	-0.0989 (0.0157)	-0.0502 (0.0175)
parent	0.0258 (0.0110)	0.0254 (0.0109)	0.0257 (0.0109)
fem	-0.100 (0.00989)	-0.0788 (0.0113)	-0.0993 (0.00985)
communist	0.00115 (0.0175)	-0.000389 (0.0171)	0.00479 (0.0170)
communistpoland	0.0163 (0.0412)		
fempoland		-0.0871 (0.0217)	
fulltimepoland			-0.241 (0.0382)
Constant	2.355 (0.0394)	2.341 (0.0393)	2.319 (0.0394)
Observations	4,539	4,539	4,539
R-squared	0.721	0.722	0.724

Standard errors in parentheses

(ix) Estimeringen viser at det å ha tatt utdanning i Polen under kommunisttiden øker timelønnen med 1,6 prosent. MEN t-testen viser at dette ikke er i statistisk signifikant. β_{15} har t- verdi på 0,4 som er lavere enn kritisk verdi i t-fordelingen ved 5 prosent signifikansnivå. Det har dermed sannsynligvis ingen effekt på lønnen å ha vært utdannet i Polen under kommunisttiden. Som nevnt i kapittel 5.2., har det heller ingen isolert effekt på lønnen å være over 55 år. Ser også at det ikke er noen endring i forskjellen på avkastning av utdanning, slik at denne variabelen ikke er betydelig for forskjellene mellom Danmark og Polen.

Likevel bør det nevnes at denne variabelen ikke kontrollerer for effekten av å ha vært et kommunistisk land. Variabelen vil kun gi et anslag på om det at utdanningen er gjennomført under et kommunistisk regime har påvirkning på avkastningen.

(x) Her viser estimeringen at timelønnen reduseres med 8 prosent dersom du er kvinne fra Polen. Dette er statistisk signifikant. Vi får en t-verdi på -4,01 for β_{16} , som er lavere enn kritisk verdi i t-fordelingen ved 5 prosent signifikansnivå.

Når vi kontrollerer for kjønnsforskjeller mellom Danmark og Polen ser vi også at forskjellen i avkastning på utdanning øker ytterligere. Dette er fordi lønnsnivået generelt er lavere for polske kvinner, og dermed blir avkastningen av utdanning trukket ned av dem. Det vil igjen si at kjønnsforskjeller ikke kan forklare den underliggende forskjellen i avkastning på utdanning i Danmark og Polen, men at den er med på å trekke forskjellen ned.

(xi) Det viser seg at det å være fulltidsarbeider i Polen reduserer timelønnen med en estimert 24 prosent. Dette er også statistisk signifikant, med en t-verdi på -3,11, som er lavere enn kritisk verdi på -1,96. Det kan også nevnes at en stor del av det råde lønnsgapet mellom Danmark og Polen kan forklares av forskjeller i fulltidsarbeid. Når vi kontrollerer for at fulltidsarbeidende i Polen tjener dårligere, reduseres det uforklarte gapet fra 148% til 124%.

Dette påvirker også avkastningen av polsk utdanning. Forskjellen i avkastning på utdanning er nå redusert fra 3,64 prosent i modell **(viii)**, til 3,45 prosent i modell **(xi)**. Dette kan tyde på at fulltidsarbeidende i Polen får bedre avkastning på utdanning enn sine likemenn i Danmark. Noe av forskjellen i avkastning på utdanning i Danmark og Polen kan dermed forklares av høyere avkastning i Polen for fulltidsarbeidende. Dette kan kanskje virke paradoksalt når de tjener så mye dårligere i utgangspunktet.

5.5. Oppsummering

I dette kapitlet har vi beskrevet resultatene av regresjonsanalysene våre. Vi har kvantifisert forskjellen i timelønnsavkastning på utdanning i Danmark og Polen til å være rundt 3,6 prosent, kontrollert for 6 ulike faktorer. Vi har også sett på mulighetene for å forklare forskjellen i avkastning på utdanning ved å se på forskjeller i likestilling i arbeidslivet, avkastning på fulltidsarbeid og aktiv kommunisme under utdanning. Det ser ut til at deler av forskjellen i avkastning kan forklares av at fulltidsarbeidende i Polen har relativt høyere avkastning på utdanningen sin enn i Danmark. Aktiv kommunisme under utdanning ser ut fra våre estimater til å ha lite å si for avkastning på utdanning. Kvinner tjener relativt dårligere i Polen, noe som bidrar til å dempe forskjellene i avkastning på utdanning.

Kapittel 6: Oppsummering og konklusjon

6.1. Oppsummering og konklusjon

Vi stilte spørsmålene om formell utdanning gir samme lønnsavkastning i Danmark og Polen; og om en eventuell forskjell kunne skyldes ulikheter mellom polakker og dansker i arbeidserfaring, likestilling i arbeidslivet, kognitive ferdigheter, sektor, stillingsprosent, alder og foreldres utdanningsnivå. Tilleggsspørsmålene våre gikk ut på om forskjellen i avkastning på utdanning kunne forklares av forskjeller mellom Danmark og Polen i lønning av fulltidsarbeidende, kvinner og personer eldre enn 55.

Økonomisk teori tilsier at avkastning på utdanning er høyere i land der færre personer tar utdanning. Dette tilsier at landet med lavest gjennomsnittlig utdanningsnivå burde ha høyere avkastning på utdanning. Ved kommunisme er idealet at alle skal tjene mer eller mindre likt uavhengig av arbeid, og dermed utdanning. Dette tilsier isolert sett at et land som Polen burde hatt lav avkastning på utdanning før 1989.

Dataen vår viser at gjennomsnittlig utdanningsnivå i Polen er litt lavere enn i Danmark, og at gjennomsnittslønn er betraktelig lavere i Polen.

For å undersøke problemstillingene våre nærmere har vi konstruert 8 modeller. Modellene er konstruert logaritmisk for å enklere kunne sammenligne på tvers av land og studier. Den første modellen inneholder kun variablene utdanning og land for å alene kunne vise forskjellen i avkastning på utdanning i Danmark og Polen. De andre modellene kontrollerer, én etter én, for andre faktorer som kan sies å påvirke timelønnen. Vi har også inkludert tre ekstra interaksjonsledd for å besvare tilleggsspørsmålene.

Hovedfunn i vår analyse:

- Det generelle lønnsgapet mellom Danmark og Polen er betydelig. Våre estimater tyder på at en polakk i gjennomsnitt tjener 150 prosent mindre i timen enn det en danske gjør.
- I alle modeller fant vi en statistisk signifikant forskjell i avkastning på utdanning i Danmark og Polen. Mens en danske kan forvente seg rundt 4,5 prosent høyere timelønn for hvert ekstra år utdanning, kan en polakk forvente nesten 8 prosent høyere timelønn for samme år.
- Det ser ut til at lengre arbeidserfaring, høyere grad av likelønn og mindre forskjell mellom sektorer i Danmark, gjør forskjellene i avkastning på utdanning mindre.
- Fulltidsarbeid viser seg å være betraktelig dårligere betalt i Polen enn i Danmark. Samtidig ser det ut til å være en relativt høyere avkastning på utdanning i Polen for

fulltidsarbeidende, som delvis kan forklare den underliggende forskjellen i avkastning på utdanning.

- Kognitive ferdigheter, alder over 55 og foreldre med høyere utdanning ser ut til å ha lite å si for forskjellen i avkastning på utdanning.

6.2. Svakheter ved modellen

Vi har konstruert en relativt enkel analyse av forskjellen i avkastning på utdanning i Polen og Danmark. Vi har kontrollert for en del faktorer som erfaring, utdanningsnivå, kognitive ferdigheter, kjønn, sektor, stillingsprosent, alder og høyere utdannede foreldre, men det er samtidig en hel del vi ikke har kontrollert for. I denne seksjonen oppsummerer vi det vi mener er de mest betydelige svakhetene ved vår modell, samt andre faktorer som bør kommenteres.

- a. Store forskjeller i arbeidsledighet i landene kan påvirke lønnsnivået betraktelig. Vår data begrenset seg til personer med inntekt. Klassisk økonomisk teori tilsier at en høyere arbeidsledighet fører til lavere lønn (Holden 2008).
- b. Datasettet vårt er begrenset, og dermed utgjorde respondenter med inntekt over dobbelt så mange i Danmark som i Polen. Dette kan potensielt gjøre dataene våre mer nøyaktige for Danmark ved at utvalget er større.
- c. Funksjonen vår er basert på timelønn og ikke årslønn. Dette kan påvirke dataene våre skjevt ved at personer som tjener mye i timen, men jobber få timer (frivillig eller ufrivillig), ser ut til å få bedre avkastning på utdanningen sin enn de som jobber mange timer og tjener mindre per time. Dette kan særlig slå ut på de som jobber deltid.
- d. Datasettet vårt begrenser seg til undersøkelser gjort ett år (2012). Av den grunn har vi ikke detaljerte data for alle som faktisk tok utdanning og arbeidet under kommunisttiden i Polen. Variabelen *communist*, som innbefatter alle over 55 år er dermed meget begrenset, og kan være årsak til at denne heller ikke ga megetsigende svar.
- e. Vi hadde lite informasjon om stilling og spesifikt yrke. Dersom respondentene i det ene landet hadde mer klassisk lavtlønte stillinger og/eller jobbet i lavtlønte yrker kan dette påvirke resultatene våre.

Referanser

1. Boarini, Romina og Strauss, Hubert. 2007. *The Private Internal Rates of Return to Tertiary Education: New Estimates for 21 OECD Countries*. OECD Economics Department Working Papers no. 591. OECD. Hentet 08.05.2019 fra https://read.oecd-ilibrary.org/economics/the-private-internal-rates-of-return-to-tertiary-education_031008650733#page1
2. Encyclopædia Britannica. 2019. "Communism". Hentet dato: 09.05.2019 fra <https://www.britannica.com/topic/communism>
3. Encyclopædia Britannica. 2019. "Communist Poland". Hentet dato: 09.05.2019 fra <https://www.britannica.com/place/Poland/Communist-Poland>
4. European Union. 2019. "Single Market". Hentet dato: 10.05.2019 fra: https://europa.eu/european-union/topics/single-market_en
5. Holden, Steinar. 2008. "Arbeidsmarked og likevektsledighet". Hentet dato: 10.05.2019 fra: <http://folk.uio.no/sholden/E1310/fnotat-Likevektsledighet.pdf>
6. Kenton, Will. 2018. "Demand for Labor". Hentet dato: 09.05.2019 fra: https://www.investopedia.com/terms/d/demand_for_labor.asp
7. Lundbo, Sten. 2019. "Polen". Hentet dato: 10.05.2019 fra: <https://snl.no/Polen>
8. Phil, Roger. 2019. "Danmark". Hentet dato: 10.05.2019 fra https://snl.no/Danmark#-%C3%98konomi_og_n%C3%A6ringsliv
9. Riis, Christian og Moen, Espen R. 2016. "Kapittel 1" i *Moderne Mikroøkonomi*, s.25. 3. utgave. Oslo: Gyldendal Akademisk.
10. Stoltz, Gerhard. 2014. "Tilbud". Hentet dato: 09.05.2019
11. Thomas, R. Leighton. 2004. *Using Statistics in Economics*. 1. Utgave. McGraw Hill Higher Education.

Appendiks 1

Tabell A1. Fullstendige estimeringsresultater for tabell 3 i hovedteksten. Standardavvik i parentes.

VARIABLES	(i) lwage	(ii) lwage	(iii) lwage	(iv) lwage	(v) lwage	(vi) lwage	(vii) lwage	(viii) lwage
educ	0.0463 (0.00216)	0.0519 (0.00213)	0.0402 (0.00234)	0.0441 (0.00238)	0.0448 (0.00238)	0.0438 (0.00241)	0.0448 (0.00239)	0.0448 (0.00241)
educpoland	0.0293 (0.00348)	0.0338 (0.00341)	0.0342 (0.00437)	0.0358 (0.00435)	0.0341 (0.00435)	0.0341 (0.00437)	0.0363 (0.00432)	0.0364 (0.00434)
poland	-1.517 (0.0484)	-1.498 (0.0474)	-1.452 (0.0647)	-1.482 (0.0644)	-1.456 (0.0644)	-1.450 (0.0648)	-1.487 (0.0642)	-1.488 (0.0645)
exper		0.0213 (0.00148)	0.0221 (0.00160)	0.0228 (0.00160)	0.0234 (0.00160)	0.0238 (0.00160)	0.0245 (0.00158)	0.0245 (0.00166)
exper_sq		-0.000296 (3.04e-05)	-0.000287 (3.24e-05)	-0.000295 (3.22e-05)	-0.000307 (3.22e-05)	-0.000312 (3.22e-05)	-0.000334 (3.19e-05)	-0.000337 (3.68e-05)
lit			0.0103 (0.0115)	0.0181 (0.0114)	0.0160 (0.0114)	0.0127 (0.0115)	0.0235 (0.0114)	0.0235 (0.0114)
num			0.0561 (0.0107)	0.0462 (0.0107)	0.0493 (0.0106)	0.0491 (0.0107)	0.0300 (0.0107)	0.0300 (0.0107)
psl			0.0254 (0.00842)	0.0218 (0.00838)	0.0227 (0.00835)	0.0249 (0.00838)	0.0247 (0.00829)	0.0248 (0.00832)
pub				-0.0778 (0.00984)	-0.0827 (0.00984)	-0.0826 (0.00984)	-0.0588 (0.0100)	-0.0589 (0.0100)
fulltime					-0.0866 (0.0158)	-0.0819 (0.0157)	-0.100 (0.0157)	-0.0999 (0.0157)
parent						0.0263 (0.0111)	0.0260 (0.0110)	0.0260 (0.0110)
fem							-0.100 (0.00989)	-0.101 (0.00989)
communist								0.00263 (0.0171)
Constant	2.538 (0.0299)	2.152 (0.0341)	2.275 (0.0377)	2.246 (0.0377)	2.310 (0.0394)	2.306 (0.0393)	2.353 (0.0392)	2.353 (0.0392)
Observations	5,901	5,879	4,587	4,584	4,584	4,539	4,539	4,539
R-squared	0.706	0.727	0.708	0.712	0.714	0.715	0.721	0.721

Appendiks 2

Vi har en enkel modell $Y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_n x_{ni} + \varepsilon_i$.

I oppgaven har vi brukt OLS (Ordinary Least Squares)-metoden til å estimere modellen til

$$\hat{Y}_i = a + b_1 x_{1i} + b_2 x_{2i} + \dots + b_n x_{ni}.$$

I en enkel regresjonsmodell blir OLS-estimatoren for β lik b . Vi vet at den er forventningsrett

$$(E(b) = \beta) \text{ og har varians } \sigma_b^2 = \frac{\sigma^2}{\sum(x_i - \bar{x})^2}.$$

Gitt at det stokastiske restleddet ε_i er normalfordelt, må også den virkelige Y være normalfordelt. Dette medfører at den estimerte b er normalfordelt.

Dette kan skrives som: $b \sim N(\beta, \sigma_b^2)$.

Standardisering gir: $\frac{b - \beta}{\sigma_b} \sim N(0, 1)$

T-test

Med denne informasjonen kan vi teste hypoteser om de virkelige variablene.

I seksjon 5.2. i hoveddelen stilles spørsmålet om forskjellen i lønnsavkastning på utdanning mellom Danmark og Polen er statistisk signifikant. Vi tester dermed hypotesen om variabelen β_3 , som angir prosentvis høyere (eller lavere) lønnsavkastning av utdanning i Polen enn i Danmark:

$$H_0: \beta_3 = 0$$

$$H_1: \beta_3 \neq 0$$

Velger et signifikansnivå på 5% ($\alpha = 0.05$). Vi har en tosidig test som innebærer signifikansnivå på 2.5% på hver side av fordelingen.

Vet at $\frac{b - \beta}{\sigma_b} \sim N(0, 1)$. Dersom nullhypotesen skal holde, må dette gjelde for $\beta_3 = 0$.

Forkaster H_0 dersom $\frac{b_3 - 0}{\sigma_{b_3}} > Z_{0.025}$ eller $\frac{b_3 - 0}{\sigma_{b_3}} < -Z_{0.025}$. Men: Vi kjenner ikke den sanne

verdien på variansen til b , og må bruke den forventningsrette estimatoren s ($E(s) = \sigma$), som vi kan lese av tabellen. Da kan vi ikke lenger bruke normalfordeling, men må bruke t-fordelingen.

Modell (i):

I modell (i) har vi 5899 frihetsgrader ($n=5901$) vi får da:

$$\frac{b_3 - \beta}{s_{b_3}} \sim t_{5899}$$

$$\rightarrow t_{0.05, 5899} < \frac{b_3 - 0}{s_{b_3}} < -t_{0.05, 5899}$$

Leser av verdiene fra tabell 3, modell (i), og setter inn:

$\frac{0.0293}{0.00348} = 8.4195 > 1.96$ (Dette kan også leses rett av regresjonstabellen i Stata, som er det vi kommer til å gjøre heretter).

8.42 er mye høyere enn kritisk verdi, slik at vi kan forkaste nullhypotesen om at lønnsavkastningen på utdanning i de to landene er lik. Estimert vårt på β_3 er dermed statistisk signifikant.

Modell (ii):

$$\frac{b_3}{s_{b_3}} = 9.92 > 1.96$$

Modell (iii):

$$\frac{b_3}{s_{b_3}} = 7.82 > 1.96$$

Modell (iv):

$$\frac{b_3}{s_{b_3}} = 8.23 > 1.96$$

Modell (v):

$$\frac{b_3}{s_{b_3}} = 7.85 > 1.96$$

Modell (vi):

$$\frac{b_3}{s_{b_3}} = 7.80 > 1.96$$

Modell (vii):

$$\frac{b_3}{s_{b_3}} = 8.40 > 1.96$$

Modell (viii):

$$\frac{b_3}{s_{b_3}} = 8.38 > 1.96$$

I alle modellene (i-viii) er resultatene statistisk signifikante, med t-verdier langt over noen kritisk verdi (for $n > 4539$). Fra Stata får vi også at p-verdien for alle testene er lik 0.000. Det vil si at vi kunne forkastet nullhypotesen ned til 0.0% signifikansnivå. Det er dermed en overveiende sannsynlighet for at lønnsavkastningen av utdanning i Polen er høyere enn i Danmark.