

Avkastning av ferdigheter i arbeidslivet i
Norge

Arne Bjørdal Langan

Juni 2014

En empirisk analyse med PIAAC 2013 data

Institutt for samfunnsøkonomi
Norges teknisk-naturvitenskapelige universitet

Forord

Denne masteroppgaven er utarbeidet ved Institutt for samfunnsøkonomi, NTNU våren 2014, som en avsluttende del av masterprogrammet i samfunnsøkonomi. Arbeidet utgjør en arbeidsbelastning på 30 studiepoeng.

Jeg vil gjerne takke veileder Bjarne Strøm for gode råd og tilbakemeldinger. Jeg vil også rette en takk til medstudentene på lesesalen for gode råd og samarbeid. Spesielt takk til Kristoffer F. Hanssen som tok seg tid til korrekturlesing. Eventuelle feil og mangler i oppgaven kan ikke klandres andre enn meg selv.

Trondheim, våren 2014

Arne Bjørdal Langan

Innhold

Forord	i
1 Innledning og disposisjon	1
2 Teoriramme	3
2.1 Humankapitalmodell	3
2.1.1 Utdanningsvalg	5
2.2 Signalisering	7
2.3 Ferdigheter og formelt utdanningsnivå	8
2.4 Kohorteffekter	11
3 Tidligere forskning	13
3.1 Tradisjonelle studier av avkastning på utdanning	13
3.2 Kohorteffekter	15
3.3 Nyere studier av avkastning av ferdigheter	16
4 Empirisk tilrettelegging og strategi	19
4.1 Generell modell	19
4.1.1 Utelatt variabel	20
4.1.2 Målefeil	21
4.2 Alternative spesifikasjoner	22
4.2.1 Ferdighetsmål	22
4.2.2 Kohorteffekter	23
4.3 Instrumentvariabel	23
5 Databeskrivelse	25
5.1 PIAAC	25
5.1.1 Ferdighetsmål	25
5.1.2 Vekting	27
5.2 Tilrettelegging av datamaterialet	27
5.3 Variabler	28
5.3.1 Avhengig variabel	28
5.3.2 Forklaringsvariabler	28
5.3.3 Kontrollvariabler	28
5.3.4 Instrumentvariabel	29
5.3.5 Kohorteffekter	29
5.4 Deskriptiv statistikk	30
6 Empiriske Resultater	31

6.1	Grunnmodell	31
6.2	Kohortspesifikk effekt av ferdigheter	35
6.3	Alderseffekter	37
6.4	Ikke-linearitet i ferdighetseffekter	38
6.5	Instrumentvariabel-estimering	41
6.5.1	Hvorfor er IV-estimatene høyere enn ved MKM?	43
6.6	Analyse på delutvalg	44
6.6.1	Avkastning av ferdigheter etter sektor	44
6.6.2	Avkastning av ferdigheter etter aldersgrupper	47
7	Oppsummering og konklusjon	50
	Vedlegg	54
A	Førstestegslikninger	54
B	Grunnmodell med foreldres utdanning inkludert som kontrollvariabel	55
C	Forskjeller i avkastning mellom privat og offentlig sektor	56

Tabeller

1	Leseferdigheter	26
2	Tallforståelse	26
3	Problemløsning i teknologiske miljøer	27
4	Deskriptiv statistikk	30
5	Generell grunnmodell, avkastning av ferdigheter estimert ved MKM.	32
6	Avkastning av ferdigheter med gjennomsnittresultat som ferdighetsmål	34
7	Avkastning av ferdigheter inkludert kohorteffekter	36
8	Avkastning av ferdigheter inkludert alder som kontrollvariabel	37
9	Avkastning av ferdigheter, ikke-lineære effekter	38
10	Avkastning av ferdigheter til forskjellige ferdighetsnivå	40
11	Instrumentvariabel-regresjon	42
12	Avkastning av ferdigheter i offentlig og privat sektor	44
13	Avkastning av ferdigheter i offentlig og privat sektor	46
14	Avkastning av ferdigheter til forskjellige aldersgrupper	48
15	Avkastning av ferdigheter til forskjellige aldersgrupper	49

1 Innledning og disposisjon

I den moderne, kunnskapsbaserte økonomien i dag er populasjonens ferdigheter generelt sett på som en nøkkelingrediens, men det finnes overraskende få beviser på avkastningen av ferdigheter (Hanushek et al., 2013).

De fleste studier som er gjort på humankapital benytter utdanning og arbeidserfaring som forklaringsvariabler. Det antas at jo mer utdanning og arbeidserfaring et individ har, jo høyere er individets humankapital. Ved å studere sammenhengen mellom utdanning, arbeidserfaring og inntekt kan man finne hvor høyt humankapital verdsettes i arbeidsmarkedet og dermed hvor høy avkastning man får av å investere i utdanning.

En stor svakhet ved disse målene på humankapital er at de ikke differensierer mellom ulike typer utdanning og arbeidserfaring. Et gitt antall år med skolegang vil gi forskjellige resultater etter forskjellige typer skole og utdanningsretning. Dermed bør heller disse variablene tolkes som innsatsfaktorer og et mer riktig mål på individets humankapital vil være hva man sitter igjen med etter skolegang og arbeid, altså de faktiske ferdighetene man besitter og benytter i arbeidslivet. Skolesystemet i Norge har vært gjennom flere endringer de siste tiårene, og i en tid da stadig flere nordmenn tar høyere utdanning, ønsker jeg derfor å se nærmere på hva dagens arbeidere får igjen for sine ferdigheter i arbeidslivet.

Det eksisterer mange studier på humankapital og avkastning av utdanning, men det er få som har benyttet direkte mål på ferdigheter som humankapital. Mange har undersøkt studenters avkastning av kunnskaper når de går inn i arbeidslivet, men nyutdannedes avkastning av ferdigheter vil generelt underestimere den totale avkastningen gjennom hele livet, da særlig de flinkeste studentene raskt vil øke i lønn (Haider og Solon, 2006)

Ved å benytte data fra PIAAC¹-undersøkelsen har jeg data på testresultater til et stort utvalg individer mellom 16-65 år i Norge. Disse testresultatene kan jeg benytte som direkte mål på individuelle ferdigheter og humankapital. Dette gir meg muligheten til å studere avkastningen av ferdigheter sammenliknet med utdanning og arbeidserfaring.

Ved hjelp av minste kvadraters metode ønsker jeg å finne den kausale effekten av økte ferdigheter på inntekt. Jeg vil ta utgangspunkt i en klassisk Mincer-modell (Mincer, 1974) for humankapital, og utvide denne for å inkludere ferdigheter. Jeg vil så se nærmere på spesifiseringen av denne modellen og skille mellom undergrupper for å fordype meg i forskjeller i avkastning av ferdigheter i arbeidslivet i Norge. Jeg vil også benytte en instrumentvariabel-metode for å omgå potensielle skjevheter i modellen.

¹Programme for the International Assessment of Adult Competencies

Jeg vil se nærmere på tidligere studier av humankapital for å undersøke hva som ligger bak valget av modelleringer og hvilke spesifiseringer som bør gjøres for å best mulig tilpasse en modell for mitt formål. En interessant utvidelse er å se på forskjeller i avkastning av ferdigheter i privat og offentlig sektor. Siden lønssystemet i offentlig sektor er stivere og mindre fleksibel enn i privat sektor vil det være interessant å undersøke om avkastningen av ferdigheter for gitt utdanning er lavere i offentlig enn i privat sektor.

En annen utvidelse er å inkludere kohort-effekter for å se på forskjeller i avkastning mellom kohorter. Dersom disse kohort-effektene er til stede vil en persons fødsels- og utdanningsår ha påvirkning på ferdighetsnivået og kunne gi forskjeller i avkastningen av ferdigheter for de forskjellige årskullene. Dette kan forklare om individers valg av utdanning og arbeid er knyttet til eksogene variabler som konjunktursvingninger i økonomien. Det kan også forklare eventuelle endringer i utdanningskvalitet. Ble folk flinkere av å gå på skolen før?

En siste interessant utvidelse i modellen er å se på signaleffekten av utdanning. Dersom det viser seg at utdanning har en effekt på individets inntekt som ikke forklares gjennom de kognitive ferdighetene vil det tyde på at utdanning gir andre kvaliteter som verdsettes i arbeidslivet, som sosiale ferdigheter og evnen til å lære. Det kan også tyde på at utdanning har en signaleffekt. Ved å ha fullført høyere utdanning vil et individ signalisere uobserverbare kvaliteter for arbeidsgiver som at man er strukturert og tålmodig.

Avkastningen av ferdigheter i arbeidslivet vil også kunne gi et bilde av incentiver til å tilegne seg kunnskap gjennom utdanning. Ettersom jeg vil se på voksne individer i arbeidslivet vil jeg også kunne undersøke hvilke variabler som påvirker ferdighetene. Avhenger ferdigheter i dag av om man har gått lenge på skole, eller får man like god avkastning av arbeidserfaring?

Videre er oppgaven bygget opp som følger:

I kapittel 2 vil jeg forklare humankapitalmodellen og signaliseringsteorien som forklarer avkastning av individers utdanning, videre vil jeg utvide disse teoriene for å inkludere direkte mål på kognitive ferdigheter. I kapittel 3 vil jeg presentere resultater fra tidligere forskning på avkastning av utdanning og ferdigheter. I kapittel 4 vil jeg presentere modellen jeg vil estimere og strategier for det empiriske arbeidet. I kapittel 5 vil jeg beskrive datamaterialet og variablene som benyttes i estimeringen. I kapittel 6 vil jeg presentere resultatene fra det empiriske arbeidet jeg har gjennomført.

2 Teoriramme

Jeg vil i dette kapittelet se nærmere på modeller og teorier som forklarer hvordan utdanning og ferdigheter påvirker individers inntekt. Jeg vil først ta utgangspunkt i en enkel humankapitalmodell, først formulert av Mincer (1974), som forklarer sammenhengen mellom utdanning og inntekt. Jeg vil så se på en alternativ teori på avkastning av utdanning gjennom signalisering. I de to siste delene av kapittelet vil jeg se nærmere på hvordan disse teoriene kan og bør utvides for å studere avkastningen av ferdigheter i arbeidslivet i Norge.

2.1 Humankapitalmodell

Utdanning spiller en viktig rolle i det moderne arbeidsmarkedet. Det eksisterer hundrevis av studier som viser at individer med høyere utdanning har høyere inntekt og opplever mindre arbeidsledighet. Men likevel har de fleste forskere vært forsiktige med å trekke for sterke slutninger av effekten utdanning har på inntekt (Lemieux, 2006). Det er vanskelig å vite om høyere lønn skyldes høyere utdanning eller om individer som tjener godt har valgt en lengre skolegang grunnet ambisjoner og ferdigheter. For å se nærmere på dette vil jeg presentere humankapitalmodellene som er benyttet i noen av disse studiene.

Hanushek et al. (2013) tok utgangspunkt i en enkel grunnmodell for humankapital:

$$y_i = \gamma H_i + \epsilon_i \quad (1)$$

Hvor: y : individets inntekt

H : individets humankapital.

Dette danner grunnlaget for humankapitalteorien. Et individs inntekt er en funksjon av individets humankapital. Jo flere og bedre kvaliteter individet innehar som verdsettes i arbeidsmarkedet, jo høyere er humankapitalen og jo høyere inntekt vil individet få. En sentral del av kritikken mot tidlige humankapital-modeller er nettopp målet på humankapital, H (Hanushek et al., 2013).

Det meste av nyere studier av humankapitals påvirkning på lønnssetting er forankret i ”Mincer’s Human Capital Earnings Function” (HCEF) utviklet av Mincer (1974). I følge denne modellen kan logaritmen av et individs inntekt i en gitt tidsperiode dekomponeres til en funksjon av utdanning og kvadratisk arbeidserfaring.

$$\log y_i = \alpha_0 + \alpha_1 S_i + \alpha_2 X_i + \alpha_3 X_i^2 + e_i \quad (2)$$

Hvor: S : Antall år fullført skolegang

X: Antall år arbeidserfaring etter utdanning

e: Stokastisk restledd

Implisitt antar Mincers formulering at skolegang er den eneste systematiske kilden til variasjoner i ferdigheter (Hanushek et al., 2013). Selv om Mincer utviklet denne likningen fra en teoretisk modell om valg av skolegang og opplæring etter skolen, så kan mønsteret i HCEF vises å forklare mye av avkastningen av utdanning også i dag. Men modellen har definitivt svakheter og forbedringspotensiale som jeg vil studere videre i dette kapitlet.

Jeg vil nå se nærmere på hvilke forenklinger som er gjort og forutsetninger som må være til stede for at denne modellen skal gi et godt bilde av avkastningen av utdanning.

Mål på utdanning.

Man antar ved Mincers HCEF at logaritmen av inntekt er en lineær funksjon av antall år fullførte utdannelse. Det er to viktige forutsetninger som ligger til grunn for i denne spesifikasjonen:

- Det beste målet på utdanning er antall år fullført utdanning
- Hvert ekstra år med utdanning har den samme proporsjonale effekten på inntekt

Under disse antakelsene vil koeffisienten α_1 i likning (2) gi den fulle effekten utdanning har på inntekt i arbeidsmarkedet. Antar man i tillegg at utdanning er gratis og at studenter ikke tjener noe under utdanningen kan α_1 tolkes som avkastningen av en investering i utdanning.

Flere har argumentert for at en grad (studiekompetanse, bachelor) betyr mer enn antall års skolegang. Denne hypotesen kalles "Sheepskin effect", altså at det er en lønnspremie for å fullføre en påbegynt utdannelse. Card og Krueger (1992) fant særlig en ikke-linearitet mellom utdanning og inntekt ved 15-16 års skolegang, altså ved fullført college i det amerikanske skolesystemet. Lemieux (2006) fant at den lineære funksjonen forklarer sammenhengen mellom utdanning og logaritmisk inntekt godt bortsett fra ved svært lavt utdanningsnivå. Han argumenterer for at den lineære tilnærmingen stemmer godt i stabile økonomier hvor vekst i relativ etterspørsel motsvares av vekst i relativt tilbud.

Kanskje har Mincer-funksjonen vist seg å være for vellykket, slik at forskere har ignorert viktige spørsmål som antakelsen om at utdanning er den eneste systematiske kilden til ferdighetsforskjeller? (Hanushek et al., 2013)

Mål på inntekt. I humankapital-teorien er inntektsfunksjonen analysert ved en rekke forskjellige mål: Årlig inntekt, månedslønn eller timelønn, men nesten alltid på logaritmisk form. Den logaritmiske formen har flere heldige egenskaper: Fordelingen er nært en normalfordeling, det nært lineære forholdet med utdanning, og det er praktisk for tolkning (Hanushek et al., 2013). Individuer med høyere utdanning har en tendens til å arbeide mer

enn individer med lavere utdanning og dette vil slå ut ved at avkastningen av utdanning synes å være større ved mål av ukes-, måneds-, eller årslønn. Når timelønn benyttes som avhengig variabel vil fokuset være på produktiviteten eller verdsettingen av arbeiderens egenskaper i arbeidsmarkedet og dermed gi et tydeligere mål på individuelle forskjeller i humankapital.

2.1.1 Utdanningsvalg

For å se nærmere på avkastningen av utdanning og ferdigheter er det også nødvendig med en teori som danner grunnlaget for individers valg av utdanning. I det meste empiriske arbeidet basert på humankapitalmodeller er utdanning behandlet som en eksogen variabel, selv om utdanning helt klart er en endogen, valgbar variabel. Utdanningsvalg er en investeringsbeslutning fra individenes side, basert på fremtidige inntekter og nåværende kostnader. Altså forskjellen i inntekt ved å ta eller å ikke ta utdanning, og den totale kostnaden ved å ta utdanning inkludert tapte inntekter (Harmon et al., 2003). For å se nærmere på hvilke variabler som påvirker utdanningsvalget, tar jeg utgangspunkt i en enkel nyttefunksjon av utdanning fra (Card, 1999):

$$U(S, y) = \log y(S) - h(S) \quad (3)$$

Hvor: S : utdanningsnivå

$y(S)$: Forventet inntekt til gitt utdanningsnivå. $y'(S) > 0$

$h(S)$: Kostnad av utdanningsnivå, inkludert tapt inntekt ved å ikke arbeide.

Individet vil finne den optimale utdanningsmengden som maksimerer nytten. Ved å investere i utdanning må man betale en kostnad i dag for å få en økt fremtidig inntekt. Investeringen i utdanning vil fortsette til forskjellen mellom marginal kostnad og marginal avkastning av utdanning er lik null. Den optimale mengden utdanning vil da være gitt ved:

$$h'(S) - \frac{y'(S)}{y(S)} = 0 \quad (4)$$

Det medfører noen implikasjoner å se på utdanning som en investering. For det første må man ta høyde for en individuell diskonteringsrente (Internal Rate of Return) som benyttes for å beregne nåverdien av fremtidig nytte og kostnader. Dersom individets IRR er høyere enn markedsrenten, som individet kan låne til, vil det lønne seg å investere i utdanning (Harmon et al., 2003). Med utgangspunkt i denne optimeringsbeslutningen kan det trekkes noen slutninger om utdanningsvalg:

- Individer som har en høyere verdsetting av nåværende inntekt enn fremtidig inntekt vil ha en høyere diskonteringsrente og lavere incentiver til å ta høyere utdanning.
- Direkte kostnader knyttet til skolegang reduserer netto-nytten av utdanning.

- Dersom sannsynligheten for å få jobb er høyere med mer utdanning, vil en økning i arbeidsledighetstrygd svekke belønningen ved å ta høyere utdanning.
- Det kan være andre nytteverdier knyttet til høyere utdanning og høyere betalt jobb, som status.
- En økning i individuelle incentiver til utdanning til gitte lønninger, vil motsvares av endringer i fordelingen av arbeidskraft ved et økt relativt tilbud av høyt utdannede arbeidere.

Den optimale mengden utdanning gitt ved likning (4) vil ikke være lik for alle individer. Det vil være store forskjeller i preferanser knyttet til både kostnader ved og nytten av utdanning. Individuelle forskjeller i kostnaden av utdanning er gitt ved heterogenitet i $h'(S)$. Denne kostnaden reflekterer også individuelle forskjeller i glede av utdanning. Forskjeller i økonomisk nytte av utdanning representeres ved heterogenitet i $\frac{y'(S)}{y(S)}$

En enkel spesifisering av disse heterogenitetene kan da være gitt ved (Hanushek et al., 2013):

$$\frac{y'(S)}{y(S)} = b_i - k_1 S \quad (5)$$

$$h'(S) = r_i + k_2 S \quad (6)$$

b_i tolkes som et individs beste estimat på sin økte inntekt per år med utdanning. Ettersom de fleste valg av utdanning skjer tidlig i livet er det naturlig å anta at de ikke selv kjenner sin avkastning av utdanning. Det er også naturlig å forvente at denne parameteren endres over tid med arbeidsmarkedsforhold og liknende men for enkelhets skyld ser vi her på den som konstant over tid.

r_i representerer individets marginale kostnad ved et ekstra år med utdanning.

Den optimale mengden utdanning for individ i er da gitt ved:

$$S_i^* = \frac{b_i - r_i}{k_1 + k_2} \quad (7)$$

Slik at jo høyere b_i og lavere r_i , jo flere år med utdanning vil være optimalt. Noen av disse individuelle forskjellene i optimal utdanning kan reflekteres gjennom kognitive ferdigheter. Høyere ferdigheter vil kunne redusere kostnaden ved å ta et ekstra år på skolen og redusere r_i . Endringer i b_i vil kunne skje gjennom eksogene sjokk på arbeidsmarkedet noe som vil gjøre at det optimale utdanningsvalget varierer over tid. Alt dette tyder på at antall års utdanning ikke er en eksogen variabel eller et optimalt mål på humankapital.

2.2 Signalisering

I markeder med skjev informasjon vil vi kunne få likevekter med "adverse selection". Jeg har til nå sett på teorier om hvordan individers humankapital bestemmes utdanning, under antakelsen at utdanning gir egenskaper som verdsettes i arbeidslivet. Arbeidsgivere kan ikke observere alle disse egenskapene og arbeidsmarkedet vil derfor individer med lavere produktivitet prøve å skjule dette. På samme måte vil derfor individer med høyere produktivitet ønske å signalisere dette til arbeidsgiver. Dette kan studeres nærmere gjennom et eksempel fra (Varian, 1992, s. 470). Vi tar utgangspunkt i to typer arbeidere med produktivitet v_1 og v_2 . Arbeiderne jobber like mye, og det er ikke mulig for arbeidsgiver å skille mellom de mindre produktive og mer produktive. Dermed mottar alle lønn:

$$\bar{s} = \pi_1 v_1 + \pi_2 v_2 \quad (8)$$

Mer produktive individer vil få betalt mindre enn sin marginale produktivitet og de mindre produktive vil få betalt mer enn sin marginale produktivitet. Dermed vil de mer produktive ønske å signalisere sin produktivitet.

En måte å signalisere sin produktivitet på er gjennom utdanning. Vi antar at utdanning ikke har noen innvirkning på produktiviteten til arbeiderne, men arbeidsgivere vil foretrekke høyere utdannede individer da disse har høyere produktivitet. Det er naturlig å anta at det er mindre kostbart for de mer produktive å ta utdanning. Kostnaden for å ta e år utdanning kan da antas å være henholdsvis $c_1 e$ og $c_2 e$ for de mer og mindre produktive arbeiderne, og at $c_1 > c_2$. Arbeiderne vil forvente at bedrifter gir en lønn $s(e)$ hvor s er en stigende funksjon av e . Ved å la e_1 og e_2 være utdanningsnivået som de to gruppene velger, må en likevekt da tilfredsstillende følgende betingelser:

$$\begin{aligned} s(e_1) &= v_1 \\ s(e_2) &= v_2 \end{aligned} \quad (9)$$

$$\begin{aligned} s(e_1) - c_1 e_1 &\geq s(e_2) - c_1 e_2 \\ s(e_2) - c_2 e_2 &\geq s(e_1) - c_2 e_1 \end{aligned} \quad (10)$$

Likningene (9) angir inntekten som en funksjon av utdanning til de forskjellige produktivitetsnivåene. Likningene (10) er "self-selection"-betingelser og sier at det vil lønne seg for en arbeider i en gitt produktivitets-gruppe å velge en utdanning tilpasset denne produktiviteten. Dersom det vil lønne seg for individer med lavere produktivitet å ta like høy utdanning som de med høy produktivitet vil ikke signaliseringen fungere. Generelt vil

det finnes mange funksjoner $s(e)$ som tilfredsstiller betingelsene i (9) og (10). I likevekt vil det lønne seg for de mer produktive arbeiderne å signalisere sin produktivitet ved å ta høyere utdanning.

I eksempelet ovenfor er det antatt at utdanning kun fungerer som en signaleffekt og ikke øker arbeideres ferdigheter, noe som selvfølgelig er en urealistisk antakelse, men det får tydelig frem teorien først formulert i (Spence, 1973). Spence argumenterer for at en ansettelsesprosess er en usikker investering fra arbeidsgivers side. Jo større sjans det er for at arbeideren er god, jo mer er de villige til å satse i form av lønn. Dermed vil arbeidere signalisere at de er dyktige gjennom høyere utdanning. Denne signaliseringsteorien peker i retning av spennende effekter som bør tas høyde for i humankapital-modellene. Antall år utdanning er ikke nødvendigvis det beste målet på humankapitalen et individ besitter.

2.3 Ferdigheter og formelt utdanningsnivå

Til nå har jeg kun sett på teorier hvor utdanning behandles som forklaringsvariabel og et mål på individets humankapital. Både teoriene om utdanningsvalg og signalisering indikerer at dette målet har flere svakheter. I andre deler av forskningen har det derfor blitt tatt høyde for at ferdigheter også påvirkes av en rekke andre variabler enn bare utdanning (Hanushek et al., 2013). Jeg vil nå se nærmere på hvordan ferdighetsmål kan benyttes for å gi et bedre bilde på individuell humankapital.

I denne oppgaven vil jeg benytte direkte mål på kognitive ferdigheter. Standardiserte testresultater brukes for å måle ferdighetene. Dersom disse ferdighetene fanger opp all variasjon i humankapitalen, H kan testresultatene benyttes direkte i den enkleste humankapitalmodellen i likning (1) som et mål på humankapital.

Dersom dette er tilfelle vil det å legge til antall års skolegang, i motsetning til i signaliseringsmodellen, ikke ha noen betydning ettersom alle relevante ferdigheter fanges opp i testresultatene. Det er lite sannsynlig at kognitive ferdigheter fanger opp all relevant variasjon i individers humankapital. Derfor kan heller testresultatet, C , sees på som et mål på humankapitalen, H , som kan inneholde en målefeil, μ .

$$H_i = C_i + \mu \tag{11}$$

Ved å estimere en likning med kun ferdighetene som et mål på humankapital vil det dermed oppstå en målefeil, ettersom ferdighetene ikke fanger opp all variasjon i humankapital.

$$\begin{aligned} y_i &= \gamma H_i + \epsilon_i \\ &= \gamma(C + \mu) + \epsilon_i \end{aligned}$$

Med denne modellen kan det dermed forventes at det oppstår en skjevhet i estimatet slik at γ vil være skjev mot null.²

Ved inkludering av ferdighetsmålet i Mincer-likningen fra likning (2) er det sannsynlig at ferdighetsscoren, C , er korrelert med antall års utdanning, S , siden bedre evner vil føre til økt skolegang gjennom en reduksjon i marginalkostnaden av utdanning.

$$\log y_i = \alpha_1 + \alpha_2 S_i + \alpha_3 X_i + \alpha_4 X_i^2 + \alpha_5 C_i + e_i \quad (12)$$

Dermed vil α_2 kunne ha en positiv skjevhet, selv om S ikke har noen annen effekt på inntekten enn den gjennom C . Disse modellene tyder altså på at det kan forventes en skjevhet i estimatet og at koeffisienten for C vil være en nedre grense for effekten av humankapitalen på inntekt (Hanushek et al., 2013).

Green og Riddel (2001) utvider Mincer-modellen noe mer for å benytte ferdigheter som forklaringsvariabel, og tar i tillegg høyde for at noen ferdigheter er observerbare og andre ikke. De argumenterer for at man må se på utdanning som en innsatsfaktor som øker ferdigheter og dermed humankapitalen, fremfor å se på det som et direkte mål på humankapitalen.

$$\log y_i = A_i p + Z_i \delta + \epsilon_i \quad (13)$$

Hvor A_i er en vektor av ferdigheter og kunnskaper som individ i er i besittelse av.

p er en vektor av verdier eller "implisitte markedspriser" på ferdighetene.

Z_i er en vektor av andre variabler som påvirker lønn

ϵ_i er et stokastisk restledd.

Denne modelleringen antar at økt utdanning ikke uavhengig fører til økt inntekt annet enn gjennom økte ferdigheter og at et individs inntekt avhenger av markedets verdsetting av ferdighetene de er i besittelse av. Dersom alle relevante ferdigheter er observerbare vil (13) gi estimater på priser og avkastning av alle ferdighetene på arbeidsmarkedet, slik at inntektene til individ i er lik summen av markedsverdien til alle ferdighetene de innehar. Men ettersom ikke alle ferdigheter er observerbare må vi ta utgangspunkt i innsatsfaktorene som påvirker individets ferdigheter og humankapital:

$$A_i = X_i B + v_i \quad (14)$$

Hvor:

X_i er en vektor av variabler som påvirker individets humankapital, som arbeidserfaring, utdanning, helse og liknende

B er en matrise som angir hvordan disse påvirkningene fører til ferdigheter.

v_i er et stokastisk restledd.

²Nærmere forklaring av dette i kapittel 4

Jeg kan da sette inn for (14) i (13):

$$\begin{aligned} \log y_i &= (X_i B + v_i)p + Z_i \delta + \epsilon_i \\ &= X_i \beta + Z_i \delta + u_i \end{aligned} \tag{15}$$

β er da en vektor som angir hvilken påvirkning hver enkelt innsatsfaktor i humankapitalen har på inntekten, altså både hvordan en innsatsfaktor påvirker ferdigheter og hvordan disse ferdighetene verdsettes i markedet. Dersom det ikke finnes direkte mål på ferdigheter vil det ikke være mulig å skille mellom disse to effektene. For å ta høyde for at visse ferdigheter er observerbare og andre ikke er det, må ferdighetsvektoren, A_i , deles opp:

$$A_i = (A_i^o A_i^u) = (X_i^o B^o X_i^u B^u) + v_i \tag{16}$$

Hvor A_i^o og A_i^u angir henholdsvis observerbare og uobserverbare ferdigheter som har tilhørende implisitte markedspriser P_i^o og P_i^u . Settes dette inn i (13) gir det da:

$$\log y_i = A_i^o P_i^o + X_i^u B^u P^u + Z_i d + u_i = A_i^o P_i^o + X_i^u \beta^u + Z_i d + u_i \tag{17}$$

Ettersom ferdighetene A_i^o er observerbare kan det også skilles mellom hvordan innsatsfaktorer påvirker ferdighetene og hvordan ferdighetene verdsettes i markedet, og dermed direkte estimere de implisitte prisene i arbeidsmarkedet. Fra estimatet av de uobserverbare ferdighetene A_i^u kan det ikke skilles mellom påvirkningen av innsatsfaktorer og avkastningen av ferdighetene. Derfor bør det inkluderes tradisjonelle humankapitalvariabler som utdanning, arbeidserfaring, kjønn og foreldres utdanning for å ta høyde for dette.

Inkluderingen av ferdigheter i modellen gir også innvirkninger på utdanningsvalget som ble forklart tidligere i kapittelet. Høyere ferdigheter vil gi individer større muligheter til å effektivt "konvertere" utdanning til humankapital. Dermed vil individer med høyere ferdigheter ha høyere diskonteringsfaktor. Individer med høyere ferdigheter vil da, til gitt utdanningsnivå, få høyere humankapital. Individer med høyere ferdigheter vil også ha en høyere alternativkostnad ved å ta høyere utdanning, som reduserer deres diskonteringsrente (Harmon et al., 2003). Denne utvidelsen er skissert i Card (1999). Card lar optimal utdanning variere mellom individer både fordi forskjellig avkastning av utdanning oppstår ved at individer med høyere ferdigheter har en høyere marginal avkastning av utdanning, og fordi individer også kan ha forskjellige substitusjonsrater mellom nåværende og fremtidig inntekt.

Dersom ferdighetene er like for alle individer vil individer med lavere diskonteringsrente velge mer skole. Men dersom ferdigheter varierer mellom individer vil det oppstå en negativ korrelasjon mellom disse elementene. Dette fører til en skjevhet ved at individer med høyere marginalnytte eller lavere marginalkostnad av utdanning velger høyere utdanning

(Harmon et al., 2003).

Dersom det ikke er noen varians i diskonteringsraten mellom individer vil skjevheten oppstå kun på grunn av korrelasjonen mellom ferdigheter og utdanning, og siden denne mest sannsynlig er positiv vil effekten overestimeres.³ Dersom det ikke er noen variasjon i ferdigheter vil endogeniteten kun oppstå på grunn av den negative korrelasjonen mellom diskonteringsrente og optimal utdanning slik at effekten underestimeres.⁴ Dermed vil totaleffekten og retningen på skjevheten ved MKM-estimering⁵ på avkastningen av utdanning være uklar.

Selv med denne tilnærmingen med kan det fremdeles være problematisk å estimere kausale tolkninger av ferdigheter på inntekt. Det er flere potensielle kilder til skjevhet. Det kan også være uobserverbare variabler eller innsatsfaktorer som påvirker både observerbare og uobserverbare ferdigheter, som talent e.l. Disse og flere økonometriske utfordringer skal jeg se nærmere på i kapittel 4.

2.4 Kohorteffekter

En annen effekt som ikke inkluderes i den klassiske Mincer-modellen, men har blitt foreslått å inkluderes av Lemieux (2006), er kohorteffekter. Kohorteffekter angir variasjon i modellen mellom kohorter eller årskull. De fanger dermed opp variasjoner over tid som kan ha innvirkning på avkastningen av ferdigheter, slik som konjunktursvingninger, arbeidsledighet og størrelse på årskull. Teoretisk sett vil humankapital-modellen være tilpasset å studere individers avkastning til forskjellige tidspunkt i livet, men i de fleste empiriske arbeid, inkludert mitt eget, ser vi på et utvalg av forskjellige individer på forskjellige tidspunkt i livet. Mincer var klar over dette og forventet derfor at modellen ville underestimere livstidsinntekten (Mincer, 1974). Dersom utdanning og økning i humankapital skal studeres som en investering, er det også viktig at all avkastning blir fanget opp. Flere nyere studier, deriblant Green og Riddell (2001) finner at Mincer-modellen vil overestimere livstidsinntekten til kanadiske menn, siden nye kohorter går inn i arbeidsmarkedet med stadig lavere lønn. Det er altså flere tidseffekter som kan peke i hver sin retning, og dette viser at det kan være problematisk å skille mellom alders-, års- og kohort-effekter.

Hanushek et al. (2013) tok høyde for deler av denne effekten ved å kun se på individer i sin "prime age". "Prime age" defineres som alderen 35-54 år, da det antas at individets inntekt er på et nivå som best gjenspeiler livstidsinntekten. Men selv ved å kun se på et delutvalg kan det fremdeles eksistere kohortforskjeller innad i utvalget. Spørsmålet er om man trenger å inkludere kohorteffektene for å se på forholdet mellom ferdigheter, utdanning,

³Gitt at høyere ferdigheter øker lønninger senere i livet mer enn tidlig i livet

⁴Gitt at diskonteringsrente og lønn er positivt korrelert

⁵MKM: Minste kvadraters metode. Forklares nærmere i kapittel 4

arbeidserfaring og inntekt. Mincer (1974) konkluderte med at dette ikke var nødvendig, men bruk av mincer-modellen på nyere data har vist seg å ikke stemme like godt (Lemieux, 2006). Lemieux konkluderer med at i en stabil økonomi hvor utdanningsnivået vokser jevnt mellom kohortene, vil det være unødvendig å kontrollere for kohorteffekter. Hvis det derimot har vært store svingninger i det relative tilbudet av forskjellige yrkesgrupper, vil dette kunne føre til endringer i lønnsstrukturen og dermed gi grunnlag for å inkludere kohorteffektene i Mincer-funksjonen.

Boockmann og Steiner (2006) argumenterte for at kohorteffekter har innvirkning på individers utdanning og humankapital på flere områder:

- Det kan representere eksogene skift i utdanningsbeslutning over tid. En periode med økt utdanningsetterspørsel gir flere arbeidere med universitetsutdanning og dermed større tilbud på arbeidsmarkedet som vil presse ned lønnspremien for utdanning.
- Endringer i kohortstørrelser, altså det totale arbeidstilbudet, vil påvirke det relative lønnsforholdet i markedet. Studier fra Tyskland tyder på at elastisiteten er lavere for høyt utdannede og dermed at en liten kohort-størrelse vil gi en større relativ økning i lønninger for disse.
- Rigiditet i lønninger til de etablerte ansatte gjennom sosiale normer. Dersom mange unge arbeidere kommer inn i arbeidsmarkedet kan man forvente at dette vil presse ned lønningene i markedet. Men de eldre ansatte vil være vant til et visst lønnsnivå og lønnsreduksjonen vil derfor slå ut hos de yngre i større grad.⁶

Oppsummering: Den klassiske Mincer-modellen for humankapitalens effekt på inntekt er både vellykket og populær, men som nevnt har det vært mye diskusjon rundt hva som er et godt mål på humankapital. Antall års skolegang er det vanligste målet, men det er flere faktorer og teorier som peker i retning av at dette gir et skjevt bilde av individers humankapital. Ved å ha data på individuelle testresultater på kognitive ferdigheter gir det meg et mye mer nyansert bilde av humankapital og en unik mulighet til å både utvide og teste relevansen til den klassiske Mincer-modellen og klassiske studier på humankapital.

⁶Dette vil innebære at gamle arbeidere blir betalt over sin produktivitet og de yngre blir betalt under, noe som vil gi arbeidsgivere incentiver til å ansette unge arbeidere fremfor gamle.

3 Tidligere forskning

Jeg vil her oppsummere de viktigste tidligere resultatene som er funnet på avkastning av utdanning og ferdigheter. Spesielt avkastning av utdanning er det gjort utallige studier på, med ulike versjoner av Mincer-modellen og ulike resultater. Jeg vil først ta for meg disse som et grunnlag for mine studier. I delkapittel to vil jeg se nærmere på studier som inkluderer kohort-effekter og variasjon i avkastningen mellom årskull, og hvordan dette bør inkluderes for å få et best mulig resultat i estimeringene. Til slutt vil jeg presentere studier som er gjort på direkte avkastning av ferdigheter. Da vil jeg spesielt fordype meg i Hanushek et al. (2013) sin studie da de benytter det samme datamaterialet som meg, og dermed danner et godt grunnlag for sammenlikning.

3.1 Tradisjonelle studier av avkastning på utdanning

Card (1999) undersøker og oppsummerer den eksisterende litteraturen på den kausale sammenhengen mellom utdanning og inntekt. For å undersøke og tolke nyere resultater på feltet fokuseres det på fire områder:

- Teoretiske og økonometriske fordeler ved å modellere den kausale effekten av utdanning ved heterogene avkastninger av skolegang.
- Nyere studier som benytter institusjonelle aspekter ved utdanningssystemet som eksogene kilder til variasjon i utdanningsresultater.
- Nyere studier på inntekt og utdanningsresultater hos tvillinger.
- Nyere studier som eksplisitt modellerer heterogenitet i avkastning av utdanning mellom grupper og individer.

Et resultat som går igjen i mange av studiene er at avkastning av utdanning ikke er en enkelt parameter i populasjonen, men en tilfeldig variabel som påvirkes av blant annet familiebakgrunn, ferdigheter og utdanningsnivå.

Han finner at regresjoner gjort ved instrumentvariabel (IV) gir høyere estimater enn ved MKM, ofte med 20% eller mer. Denne høyere IV-estimatoren kan skyldes flere ulike grunner og har skapt noe forvirring da mange vil anta at MKM vil overestimere modellen. Card mener en mulig grunn kan være en underliggende heterogenitet i avkastningen av utdanning, slik at instrumenteringen påvirker individer med høyere marginal avkastning og dermed øker estimatoren. I modeller hvor familiebakgrunn benyttes som instrument vil uobserverte ferdigheter føre til en høyere skjevhet enn ved MKM. Det store spørsmålet som forblir ubesvart er om den individuelle avkastningen av utdanning er mindre eller større enn den samfunnsøkonomiske avkastningen.

Lemieux (2006) benytter eksisterende litteratur og nye estimater for å teste hvor godt Mincer-modellen forklarer avkastning av utdanning i dag sammenliknet med for 30 år

siden. Han mener det er oppsiktsvekkende at en modell utviklet for så lenge siden fremdeles står like sterkt, selv med nye estimeringsmetoder og avanserte data. Det er nyttig å bruke denne modellen dersom det er en god tilnærming for å kunne sammenlikne med tidligere studier, men dersom det viser seg at Mincer-modellen ikke lenger er en god tilnærming vil vi få unøyaktige estimater fra mange studier som alle gjør samme feil.

Ved å sammenlikne flere studier konkluderer han med at modellen stemmer overraskende godt med datamaterialet også i dag. Det viser seg å være en god tilnærming å benytte logaritmisk inntekt som avhengig variabel, lineær utdanning og kvadratisk arbeidserfaring, men det trengs noe finjustering. Mincer-modellen forklarer ikke like godt data fra 1980-1990, men ved å inkludere kohort-dummys bedres dette. En av grunnene til at modellen ikke passer like godt til datamateriale fra 80- og 90-tallet er en økende konveksethet i lønnen med hensyn på utdanning. En annen utfordring er at erfaring-lønn profilene ikke lenger er parallelle for forskjellige utdanningsgrupper. Spesielt blant college- og high school-studenter er lønnen nå mye høyere for individer med mindre erfaring enn da Mincer gjennomførte sine studier.

Lemieux konkluderer med at Mincer-modellen fortsatt fungerer som et godt utgangspunkt i stabile økonomier hvor utdanning og lønn vokser jevnt over kohorter slik det gjorde i perioden da Mincer utarbeidet sine studier. I studier gjennomført på datamateriale fra nyere tid er det viktig å bekrefte robustheten til modellen ved å inkludere kvadratiske ledd og kohort-effekter.

Harmon et al. (2003) fokuserer på utdanning som en privat investering i humankapital og hvordan dette påvirker utdanningsvalg og avkastningen av utdanning. De benytter seg av regresjoner utført på flere forskjellige datasett for å sammenlikne spesifikasjoner og resultatene de gir. Ved å sammenlikne forskjellige studier ser de på hvordan ulike modeller påvirker estimatene av avkastningen av utdanning. Gjennom en meta-analyse kombinerer de resultater fra en rekke studier over flere land og konkluderer bl.a. med at avkastningen av utdanning i Norge er blant de laveste i utvalget på like over 5%. De studerer også alternative spesifiseringer som signaliseringsmodeller og IV-estimeringer. Ved en meta-analyse av de forskjellige estimeringsmetodene viser det seg at IV-estimatene gir en mye høyere avkastning enn MKM-estimatene. 7%-9% med MKM og 11%-15% med IV. De finner lite direkte bevis på signalisering.

De konkluderer med at de enkleste estimeringene av gjennomsnittsinntekt til forskjellige nivå av utdanning skjuler flere utfordringer. De finner at MKM-spesifikasjonene er relativt lite sensitive for endringer i spesifikasjonene, som inkludering av flere kontrollvariabler. Estimaten tyder på en høyere avkastning av utdanning for individene i toppdesilen av inntektsfordelingen. Dette kan forklares ved sammenhengen mellom høyere ferdigheter og høyere utdanning. En utfordring med IV-estimatorene er at disse kan reflektere avkast-

ningen for undergruppen som påvirkes av behandlingen eller eksperimentet som benyttes som instrument. De finner flere eksempler på at familiebakgrunn er et svakt instrument. På tross av mange spissfindigheter i estimatene, er det utvilsomt en positiv effekt på individuell inntekt av utdanning. Størrelsen på avkastningen virker også stor relativ til andre investeringer. Dette kan implisere en samfunnsøkonomisk gevinst, men kan også tolkes som en marginalavkastning.

3.2 Kohorteffekter

Boockmann og Steiner (2006) ser på hvordan avkastningen av utdanning varierer mellom kohorter i Tyskland. De benytter en utvidet Mincer-modell hvor de inkluderer dummyvariabler for kohorter på data fra "the German Socio-Economic Panel" (GSOEP) fra 1984-1997. Ved interaksjonsledd mellom kohort-dummyene og antall års utdanning finner de beviser for en redusert avkastning av utdanning i vest-tyske kohorter etter Andre Verdenskrig. Reduksjonen er mye sterkere hos kvinner enn for menn (henholdsvis 7,8% og 3,6% lavere avkastning av utdanning for individer født i 1970-1974 enn i 1925-1929). De inkluderer også et interaksjonsledd mellom utdanning og arbeidserfaring for å ta høyde for at avkastningen av utdanning kan øke med høyere arbeidserfaring. Dette reduserer kohorteffektene noe.

En utfordring med spesifiseringen er å skille de rene kohorteffektene fra erfaring- og alderseffekter. Særlig i offentlig sektor er dette et problem da lønningene i større grad øker med alder. Dette i tillegg til mer rigide lønninger i offentlig sektor gjør at de rene kohorteffektene er lettere å tolke fra privat sektor.

Hægeland et al. (1999) ser på avkastning av utdanning i Norge over tid. Det er interessant å studere sammenhengene mellom inntekt og avkastning av utdanning over tid på tvers av kohorter i Norge, særlig siden utdanningssystemet har vært gjennom store endringer de siste tiårene. Det er også interessant for å sammenlikne med andre land som England og USA som har opplevd økt avkastning av utdanning og dermed større inntektsforskjeller i dette tidsrommet.

De benytter data fra folke- og boligtellingerne 1960, 1970 og 1980 for å studere sammenhengen mellom inntekt og utdanning mellom forskjellige kohorter. Her benyttes også dummyvariabler for regionen hvor individet vokste opp som instrument for utdanning. Det bygger da på forutsetningen om at dette har direkte innvirkning på utdanningsvalget, men ikke på inntekten. De finner en stabil avkastning av utdanning fra 1980 til 1990, og ikke noe som tyder på at den økte andelen som tar høyere utdannelse har ført til en reduksjon i avkastningen av utdanning.

3.3 Nyere studier av avkastning av ferdigheter

Murnane et al. (1995) undersøker i hvor stor grad viktigheten av kognitive ferdigheter har endret seg de siste tiårene ved å sammenlikne hvordan matteferdighetene til "high school seniors" påvirker deres lønninger seks år etter endt utdanning. De ønsker å undersøke om økte krav i arbeidslivet grunnet utvikling i industri og tjenester har ført til endringer i ferdighetene til studentene, og om kognitive ferdigheter har blitt viktigere i lønnsdetermineringen nå enn før. De benytter data fra "The National Longitudinal Study of the High School Class of 1972" (NLS72) og "High School and Beyond" (HS & B) som inneholder data for individer som har fullført formell utdanning i henholdsvis 1972 og 1980 og befinner seg i betalt arbeid 6 år etter. Ved å undersøke sammenhengen mellom logaritmisk timelønn og testresultater i matematikk finner de en mye større avkastning av matteferdigheter i arbeidslivet fra den seneste kohorten. En økning i matteferdigheter med ett standardavvik, 6,25 poeng gir henholdsvis \$0,24 og \$0,57 høyere timelønn for menn som var ferdig utdannet i 1972 og 1980 når det kontrolleres for utdanning, arbeidserfaring, familiebakgrunn, rase og geografi.

De finner klare bevis på at grunnleggende kognitive ferdigheter var en viktigere indikator på lønn seks år etter fullført high school i midten av 1980-tallet enn på slutten av 1970-tallet. Dette tyder på et skift i etterspørselen etter arbeidere med høyere ferdigheter. De finner også at de kognitive ferdighetene hadde mye mindre effekt på lønninger to år etter ferdig utdanning enn ved seks år etter ferdig utdanning, noe som tyder på at det tar tid før arbeidsgivere kan observere ferdighetene.

Murnane et al. (2000) undersøker sammenhengen mellom tenåringers kognitive ferdigheter og deres lønninger omtrent ti år senere for å finne hvor stor del av lønnsforskjeller i 1985 og 1991 som kan forklares ved skolekarakterer. De benytter de samme datasettene som Murnane et al. (1995), NLS72 og HS&B for å studere sammenhengen mellom matte-resultatene når de gikk ut fra high school med lønningene når de var henholdsvis 31 og 27 år gamle. De finner at et standardavviks økning i matteferdigheter vil gi en økt årsinntekt med 3,7% for 31 år gamle menn i 1985.

Ved å sammenlikne med modeller hvor de inkluderer dummyvariabler for høyeste fullførte utdanningsgrad, konkluderer de med at en tredjedel av avkastningen av kognitive ferdigheter er en indirekte effekt som virker gjennom at individer med høyere ferdigheter har større sannsynlighet for å fullføre college. De konkluderer med at investeringer i skole for å tilegne seg kognitive ferdigheter vil gi avkastning i arbeidslivet, men at en av de viktigste grunnene til at høyere ferdigheter gir høyere inntekt er på grunn av høyere sannsynlighet for å ta høyere utdanning. Dette betyr at for at en tenåring skal få fullt utbytte av sine ferdigheter, er han eller hun avhengig av muligheter for høyere utdanning.

Green og Riddell (2001) benytter data fra "The International Adult Literacy Survey" (IALS) fra 1994 hvor leseferdigheter ble målt i tre forskjellige kategorier, og benytter dette for å studere avkastning av ferdigheter i Canada. Ved å ta utgangspunkt i Mincer-modellen finner de en høy, positiv korrelasjon mellom leseferdigheter og inntekt. En økning med ti prosentiler i fordelingen av ferdighetsscoren vil øke timelønnen med 4,6%.

De benytter også IV-regresjoner for å håndtere den eventuelle endogeniteten knyttet til uobserverbare variabler som påvirker både utdanning og inntekt. Som instrument benyttes foreldres utdanning og begrunnelse for å avslutte utdanning. Dette gir noe høyere estimatorer på avkastningen enn ved MKM-estimeringen, særlig ved årlig inntekt som avhengig variabel og kun begrunnelse for å avslutte utdanning som instrument, (0,0178 mot 0,0055). Når foreldres utdanning inkluderes som instrument er koeffisienten høyere enn MKM-estimatoren, men ikke signifikant forskjellig. De benytter også en "three-stage least squares"-modell for å fange opp den felles bestemmelsen av utdanning, leseferdigheter og inntekt. Dette gir også et noe høyere estimat enn ved MKM (0,0074 mot 0,0055). De konkluderer med at leseferdigheter har en stor, positiv effekt på inntekt ved flere estimeringsmetoder, men klare tall på den kausale effekten krever et større antall observasjoner og instrumenter.

Hanushek et al. (2013) har benyttet datamaterialet fra PIAAC for å sammenlikne avkastning av ferdigheter 22 i forskjellige land. De tar utgangspunkt i en klassisk Mincer-modell, som utvides med mål på ferdigheter. I hovedsak benyttes tallforståelse som forklaringsvariabel da det lettest lar seg sammenlikne på tvers av landegrenser. Her studeres et utvalg individer i sin "prime age" i full jobb, ettersom dette gir det beste bildet av inntekten et individ vil ha i løpet av sin levetid, og dermed den totale avkastningen av ferdighetene.

De finner at bedre kognitive ferdigheter i stor grad er knyttet til høyere lønninger. Resultatet viser at en økning i et standardavvik i ferdighetene gjennomsnittlig gir en økning i timelønn på 17,8%. De finner store forskjeller mellom land. Avkastningen i landene med høyest avkastning (Tyskland, USA og Irland) er omtrent dobbelt så stor som i de med lavest (Norge, Sverige og Tsjekkia). De finner også en tydelig høyere avkastning hos individer i sin "prime-age" enn yngre og eldre arbeidere, høyere avkastning ved arbeid i privat enn offentlig sektor, og høyere avkastning ved foreldre med høyere utdanning. Avkastningen er systematisk lavere i land med sterkere fagforeninger og større offentlig sektor. De tester robustheten i modellen ved å inkludere flere kontrollvariabler og ta høyde for heterogene effekter, og konkluderer med at det i alle land i utvalget er en positiv sammenheng mellom ferdigheter og inntekt.

Oppsummering: Det er gjort veldig mye forskning på avkastning av humankapital. De fleste av humankapitalmodellene benytter utdanning som forklaringsvariabel og har ikke direkte mål på individenes ferdigheter. Mesteparten av studiene på avkastning av ferdigheter er basert på data fra nyutdannede og gir dermed et skjevt bilde av den totale avkastningen over hele livstiden. Jeg vil legge meg tett opptil hva Hanushek et al. (2013) har gjort, men med et mer kritisk blikk på Mincer-modellen slik som foreslått av Lemieux (2006). På denne måten vil jeg kunne sammenlikne mine resultater med resultatene fra Mincer-modelleringen i Hanushek et al. (2013) og sette et klarere søkelys på funksjonsformen som fungerer best for å forklare avkastningen av ferdigheter i det norske arbeidsmarkedet i stedet for hva som best lar seg sammenlikne mellom land.

4 Empirisk tilrettelegging og strategi

Jeg vil i dette kapittelet gå nærmere inn på strategiene jeg har benyttet i det empiriske arbeidet i oppgaven. Jeg vil først beskrive grunnmodellen som benyttes og mulige problemer dette kan gi meg i estimeringen. Jeg vil så gå nærmere inn på alternative spesifikasjoner og strategier for å omgå disse eventuelle problemene.

4.1 Generell modell

Med bakgrunn i Mincer-modellen (Mincer, 1974) og utvidelsene gjort av Lemieux (2006) vil jeg ta utgangspunkt i en utvidet modell sammenliknet med Hanushek et al. (2013) for så å kunne sammenlikne med nevnte artikkels resultater og gi et tydeligere bilde av avkastningen av ferdigheter i arbeidslivet i Norge. Jeg vil dermed ta utgangspunkt i modellen:

$$\begin{aligned} \ln y_i = & \beta_0 + \beta_1 \textit{lese}_i + \beta_2 \textit{tall}_i + \beta_3 \textit{data}_i + \beta_4 \textit{erfaring}_i + \beta_5 \textit{erfaring}_i^2 + \delta_1 \textit{kvinne}_i \\ & + \delta_2 \textit{kohort2}_i + \delta_3 \textit{kohort3}_i + \delta_4 \textit{kohort4}_i + \delta_5 \textit{kohort5}_i + u_i \end{aligned} \quad (18)$$

Hvor:

- y_i : Inntekt gitt ved timelønn

- \textit{lese}_i : Testresultat for leseferdigheter

- \textit{tall}_i : Testresultat for tallforståelse

- \textit{data}_i : Testresultat for problemløsning i teknologirike miljøer

- $\textit{erfaring}_i$: Antall års arbeidserfaring.

- \textit{kvinne}_i : Binær variabel, lik 1 hvis individ i er kvinne.

Kohort-variablene er dummyvariabler som er lik én dersom individet er i aldersgruppen:

- $\textit{kohort1}_i$: 35-38 år

- $\textit{kohort2}_i$: 39-42 år

- $\textit{kohort3}_i$: 43-46 år

- $\textit{kohort4}_i$: 47-50 år

- $\textit{kohort5}_i$: 51-54 år

- u_i : stokastisk restledd

$\textit{kohort1}$ benyttes som referansegruppe slik at koeffisientene for de andre kohortene angir resultater som må sees i forhold til denne gruppen. Inntekten er oppgitt i logaritmiske verdier basert på diskusjon i kapittel 2, det samme gjelder kvadratisk erfaringsledd. Jeg vil estimere modellen ved hjelp av MKM, minste kvadraters metode, hvor β_1 , β_2 og β_3 da er interesseparameterne som angir individets ferdigheters påvirkning på inntekten. Jeg vil gi en nærmere forklaring av disse variablene i kapittel 5

For å få konsistente og forventningsrette estimatorer ved MKM-estimering av denne likningen, må den oppfylle Gauss-Markov-forutsetningene (Wooldridge, 2009, s. 104).

-Modellen er lineær i parameterne.

-Det er et tilfeldig, representativt utvalg.

-Det er ingen perfekt kollinearitet. Det er ingen eksakt lineær sammenheng mellom noen av variablene.

-Restleddet u_i har en forventningsverdi lik null.

$$E(u_i|X_1, X_2, \dots, X_k) = 0$$

-Homoskedastisitet; restleddet u_i har den samme variansen gitt verdiene til alle forklaringsvariablene.

$$Var(u_i|X_1, X_2, \dots, X_k) = \sigma^2$$

Et av hovedproblemene ved estimeringen er å finne den kausale effekten av fedigheter, siden det vil være veldig mange effekter som spiller inn og påvirker et individs inntekt. Det kan altså oppstå et utelatt variabel-problem

4.1.1 Utelatt variabel

For å se nærmere på hvordan en utelatt variabel vil kunne påvirke estimatet vil jeg ta utgangspunkt i en enkel regresjon for å eksemplifisere:

$$y_i = \beta_0 + \beta_1 \text{ferdighet}_i + u_i \quad (19)$$

Hvor: ferdighet_i Er testresultatet for en av ferdighetene.

Dersom jeg nå antar at den sanne modellen for likning (19) er gitt ved:

$$y_i = \beta_0 + \beta_1 \text{ferdighet}_i + \beta_2 X_i + v_i \quad (20)$$

Slik at variabelen X_i påvirker venstresidevariabelen, y_i men ikke er medregnet i modellen. Det eksisterer altså en utelatt variabel. I likningen (19) vil da denne variabelen inngå i restleddet, slik at:

$$u_i = \beta_2 X_i + v_i \quad (21)$$

Ved å estimere likning (19) ved MKM vil estimatoren være gitt ved:

$$\hat{\beta}_1 = \frac{\sum(y_i - \bar{y})(\text{ferdighet}_i - \overline{\text{ferdighet}})}{\sum(\text{ferdighet}_i - \overline{\text{ferdighet}})^2} \quad (22)$$

Sannsynlighetsgrensen av estimatoren vil da være gitt ved:

$$plim(\hat{\beta}_1) = \beta_1 + \frac{\beta_2 cov(ferdighet_i, X_i)}{var(ferdighet_i)} \quad (23)$$

Dersom $cov(ferdighet_i, X_i) = 0$ vil da fortsatt estimatoren være konsistent. Men dersom det er en korrelasjon mellom forklaringsvariabelen og den uobserverte/utelatte variabelen vil $cov(X_i, u_i) \neq 0$ og vi får et brudd på en av restleddsforutsetningene. Dermed vil det resultere i en skjev estimator. Retningen på denne skjevheten avhenger av retningen på korrelasjonen.

I modellen fra likning (18) vil dette skape problemer dersom det eksisterer variabler som ikke er inkludert i likningen, men er korrelert med individenes ferdigheter. Slike variabler kan for eksempel være talent og medfødte evner, sosiale ferdigheter eller familiebakgrunn som gir individuelle forskjeller i læring og ferdigheter men ikke kan observeres.

4.1.2 Målefeil

Et annet potensielt problem og kilde til feilestimering i modellen er målefeil. Målefeil oppstår dersom vi ikke observerer den sanne verdien til forklaringsvariablene. I mitt tilfelle med ferdighetsmål som forklaringsvariabel er dette et tydelig potensielt problem ettersom ferdighetsmålene ikke kan observeres direkte og vi må stole på at testene gir et riktig mål på individenes kognitive ferdigheter. Dersom testscoren ikke gir et riktig mål vil MKM-estimatoren bli skjeve. Målefeilen kan også oppstå ved at de kognitive ferdighetene ikke fanger opp all variasjon i humankapital, da den også kan avhenge av uobserverbar variasjon.

Dette kan enklere forklares ved å igjen ta utgangspunkt i den forenklede likningen:

$$y_i = \beta_0 + \beta_1 ferdighet_i + u_i \quad (24)$$

Dersom det eksisterer en målefeil, vil forklaringsvariabelen egentlig være gitt ved:

$$ferdighet_i^* = ferdighet_i + \mu_i \quad (25)$$

Variabelen som benyttes i estimeringene er dermed ikke den sanne verdien på ferdighetene. Målefeilen, gitt ved μ_i , vil føre til en skjevhet i estimatet av likning (24). Målefeil i forklaringsvariabelen gjør at estimatoren får en skjevhet mot null. (Wooldridge, 2009, s. 319) Den potensielle målefeilen gir derfor grunn til å oppfatte MKM-estimatet som en nedre grense for avkastning av ferdigheter (Hanushek et al., 2013).

For å ta høyde for en eventuell heteroskedastisitet, altså at restleddsvariansen varierer

mellom forklaringsvariablene, vil jeg også benytte robuste standardavvik i estimeringene. Et brudd på forutsetningen om homoskedastisitet vil ikke føre til skjeve estimatorer, men vil gi feil ved blant annet t- og F-fordelinger og dermed skape problemer ved tester som tar i bruk disse. (Wooldridge, 2009, s. 265)

4.2 Alternative spesifikasjoner

Grunnmodellen gitt ved likning (18) gir utgangspunktet for estimeringen, men det er ikke gitt at dette er den beste fremstillingen. Mye av utfordringen i estimeringen ligger i å finne den riktige funksjonsformen for best mulig å forklare avkastningen av ferdigheter.

4.2.1 Ferdighetsmål

Jeg har resultater fra tester på tre forskjellige ferdigheter; leseferdigheter, tallforståelse og problemløsning i teknologirike miljøer. Jeg kommer tilbake til disse ferdighetsmålene i kapittel 5. En utfordring med disse målene er at et individs resultater på de forskjellige testene er sterkt korrelert. Er man flink til å lese er sannsynligheten høy for at man også har god tallforståelse. Når alle ferdighetene inkluderes i modellen vil det altså kunne være en kollinearitet mellom variablene som vanskeliggjør det å finne de separate effektene av den enkelte ferdighet på inntekten. Hanushek et al. (2013) løser dette ved å kun benytte tallforståelse som forklaringsvariabel og argumenterer også for at dette lettere lar seg sammenlikne på tvers av landegrenser enn leseferdigheter. Jeg vil inkludere alle tre ferdighetsmålene for så å teste om effektene av hvert ferdighetsmål er like. Dersom dette er tilfelle vil jeg benytte et gjennomsnittresultat på de tre testene som forklaringsvariabel.

En annen mulig utfordring er ikke-linearitet i avkastning av ferdigheter. Det er ikke nødvendigvis en god tilnærming at marginalavkastningen av økte ferdigheter er lik uansett hvilket nivå av ferdighetene individet befinner seg på. Det kan tenkes at forskjeller i ferdigheter på et grunnleggende nivå gir større utslag på lønn enn forskjeller på et høyere nivå, noe resultatene til Murnane et al. (1995) tydet på.

Denne eventuelle ikke-lineariteten kan formuleres på flere måter. En strategi er å inkludere et kvadratisk ledd for å ta høyde for en parametrisert ikke-linearitet hvor det kan forventes et negativt kvadratisk ledd dersom avkastningen er avtakende. En annen strategi kan være å inkludere dummyvariabler for forskjellige nivå for ferdighetene. I undersøkelsen deles resultatene opp i fem forskjellige nivå, noe som kan utnyttes ved å inkludere en dummyvariabel for hvert av disse nivåene for å studere hvordan avkastningen varierer mellom de forskjellige ferdighetsnivåene.

4.2.2 Kohorteffekter

Jeg har i grunnmodellen, likning (18), inkludert dummyvariabler for kohorter for å fange opp eventuelle forskjeller i avkastning knyttet til forskjeller i aldersgrupper som kan skyldes variasjoner i kullstørrelse, konjunkturer eller liknende, som diskutert i kapittel 2.4. Som Lemieux (2006) konkluderte med, vil det være nødvendig å kontrollere for kohort-effekter dersom det har vært svingninger i arbeidstilbudet som har kunnet påvirke lønnsstrukturen i arbeidsmarkedet. Jeg vil derfor i min modell ta utgangspunkt i å inkludere kohort-effektene, for så å teste om modellen benyttet i (Hanushek et al., 2013) uten disse, er en gyldig forenkling for arbeidsmarkedet i Norge. Dersom kohorteffektene er til stede og det er klare forskjeller i inntekt mellom kohorter, vil det også være interessant å undersøke forskjeller i avkastning av ferdigheter mellom kohortene ved å inkludere interaksjonsledd mellom kohort-dummyene og ferdighetsmålet.

4.3 Instrumentvariabel

Som nevnt tidligere i kapittelet er det flere potensielle problemer ved MKM-estimering av modellen gitt ved likning (18). Det er årsaker til at forklaringsvariabelen, $ferdighet_i$, kan være korrelert med restleddet, u_i . Enten via utelatte variabler eller målefeil. For å adressere de potensielle skjevhetene ved korrelasjon mellom forklaringsvariabelen og restleddet vil jeg benytte en instrumentvariabelmetode. En instrumentvariabel, Z_i , må oppfylle følgende kriterier:(Wooldridge, 2009, s. 508)

1. $cov(Z_i, ferdighet_i) \neq 0$, relevanskriteriet
2. $cov(Z_i, u_i) = 0$, eksklusjonskriteriet

Relevanskriteriet sier at instrumentvariabelen må være korrelert med forklaringsvariabelen. Denne korrelasjonen bør også være høy for at det skal være et godt instrument, slik at den kan forklare mye av variasjonen i ferdighetene. Dette kriteriet lar seg enkelt teste i førstestegslikningen.

Eksklusjonskriteriet sier at instrumentvariabelen må være ukorrelert med restleddet, og dermed ikke ha noen direkte effekt på den avhengige variabelen når vi har kontrollert for forklaringsvariablene. Variabelen må være eksogen. Dette kriteriet lar seg generelt ikke teste på samme måte som relevanskriteriet, så jeg må ta utgangspunkt i resultatene gitt forutsetningen om at kriteriet holder.

Et ofte benyttet instrument ved studier på avkastning av utdanning er foreldres utdanning. Dette viser seg å ha stor sammenheng med individers utdanningsvalg og ferdigheter, men sannsynligvis påvirker det ikke inntekten direkte gjennom noen annen effekt enn kontrollvariablene (Card, 1999). I lys av dette vil jeg benytte foreldres utdanning som

instrument for ferdigheter og undersøke om en slik empirisk strategi gir andre resultater enn MKM.

I praksis gjøres dette ved å benytte en to-steps-MKM-metode (2SLS). Jeg gjennomfører en regresjon på en førstestegslikning som estimerer instrumentenes, samt kontrollvariablenes, effekt på forklaringsvariabelen, *ferdighet*. Denne estimeringen angir i hvor stor grad instrumentet forklarer variasjonen i ferdighetene og dermed også om relevanskriteriet er oppfylt. Denne førstestegslikningen vil da estimeres ved MKM:

$$\begin{aligned} ferdighet_i = & \alpha_0 + \alpha_1paredmed_i + \alpha_2paredhi_i + \beta_4erfaring_i + \beta_5erfaring_i^2 \\ & + \delta_1kvinne_i + \alpha_3alder_i + \epsilon_i \end{aligned} \quad (26)$$

Hvor *paredmed* og *paredhi* er dummyvariabler for hvorvidt individets foreldre har høyere utdanning.⁷ Interesseparameterne er dermed α_1 og α_2 som angir foreldres utdannings påvirkning på individuelle ferdigheter. Ved å estimere denne likningen ved MKM finner jeg et estimat på individuelle ferdigheter til verdier på forklarings- og kontrollvariablene, $\hat{ferdighet}_i$. Jeg benytter så dette predikerte ferdighetsmålet $\hat{ferdighet}_i$ fra førstesteget, istedet for de faktiske ferdighetsmålene, i grunnmodellen for å estimere andrestegslikningen.

Oppsummering: Jeg har nå sett på problemer og utfordringer som kan oppstå ved å estimere modellen ved MKM, og forskjellige strategier som kan benyttes for å omgå disse. Modellformuleringen er ikke gitt, så selv om Hanushek et al. (2013) kun benyttet en enkel Mincer-modell, ønsker jeg å teste om dette er gyldige forenklinger og om det eventuelt finnes bedre formuleringer for å forklare avkastningen av ferdigheter i Norge.

⁷En nærmere forklaring av disse variablene fremkommer i kapittel 5

5 Databeskrivelse

Jeg vil i dette kapittelet beskrive datamaterialet som ligger til grunne for mitt empiriske arbeid. Jeg vil først beskrive PIAAC-undersøkelsen som har samlet inn data og hvordan denne undersøkelsen har lagt opp målene for kognitive ferdigheter, før jeg går nærmere inn på mitt delutvalg og hvilke variabler jeg vil benytte i mine estimeringer

5.1 PIAAC

Datamaterialet er et resultat av PIAAC-undersøkelsen (Programme for the International Assessment of Adult Competencies) fra 2013 gjennomført i regi av OECD⁸.

Undersøkelsen vil måle voksne individers kognitive ferdigheter som er nødvendige for å fungere i samfunnet. Resultatene fra undersøkelsen skal hjelpe land å bedre forstå hvordan utdanning og jobbtrening kan forbedre disse ferdighetene. Undervisere, myndigheter og økonomer kan benytte informasjonen til å utvikle politikk og retningslinjer til å forbedre voksnes ferdigheter (Bjørkeng, 2013).

Undersøkelsen er utviklet av OECD og gjennomført i 24 land. Jeg benytter kun resultatene fra Norge. I alt 4947 fullførte bakgrunnsintervjuer på voksne i alderen 16-65 danner bakgrunnen for datamaterialet. Individene har blitt intervjuet i sine hjem og løst oppgaver på datamaskiner, eventuelt på papir ved mangel på datakunnskaper. Undersøkelsen ble gjennomført i to faser; en hovedfase i desember 2011 og en oppfølgingsfase i april 2013. Spørsmålene tester individenes leseferdigheter, tallforståelse og problemløsning i teknologirike miljøer, i tillegg til bakgrunnsinformasjon om deltakerne og hvordan ferdighetene brukes i arbeids- og hverdagslivet. Undersøkelsen er designet slik at resultatene er sammenliknbare på tvers av landegrenser og målet er å repetere undersøkelsen senere slik at man kan studere utviklingen over tid.

Det er gjennomført liknende undersøkelser på lese- og tallforståelse tidligere, IALS (The International Adult Literacy Survey) fra 1998 og ALL (The Adult Literacy and Life Skills Survey) fra 2003. Resultatene fra disse er ikke direkte sammenliknbare med PIAAC-resultatene da poengskalaen er endret.

5.1.1 Ferdighetsmål

Spørreundersøkelsen testet individene på tre forskjellige områder. Leseferdigheter, tallforståelse og problemløsning i teknologirike miljøer. På hvert av de tre områdene er ferdighetsnivå estimert på en skala fra 0 til 500 poeng. Poengscoren er så delt inn i fem eller tre

⁸Organisation for Economic Co-operation and Development. <http://www.oecd.org/site/piaac/>

ferdighetsmål, på en slik måte at om en person får en total poengscore på et av målene vil han eller hun med 67% sannsynlighet klare å løse en oppgave på dette vanskelighetsnivået. Inndelingen av ferdighetsmålene i PIAAC-undersøkelsen er annerledes enn hva som ble benyttet i IALS og ALL, noe som vanskeliggjør en direkte sammenlikning med resultatene fra disse.

Leseferdigheter, er definert som:

”Evnen til å forstå, bruke og gripe an skriftlig test for å delta i samfunnet, oppnå egne mål og utvikle sine ferdigheter og sitt potensiale”(Bjørkeng, 2013)

Dette innebærer altså ikke skriving. Testene gir tilsammen en totalscore 0-500. Og gjennomsnittsscoren for hele det norske utvalget ble 278. Testresultatene sorteres etter fem ferdighetsmål.

Tabell 1: Leseferdigheter

Nivå	0	1	2	3	4	5
testresultat	<176	176-225	226-275	276-325	326-375	376-500
prosentandel Norge	3	9	31	43	13	1
prosentandel OECD	3	12	34	39	11	1

Tallforståelse, er definert som:

”Evnen til å uthente, bruke tolke og formidle matematisk informasjon og ideer for å kunne håndtere en rekke situasjoner i voksenlivet som stiller krav til matematisk kunnskap. Tallforståelse stiller krav til at leseren kan håndtere et problem i en realistisk kontekst ved å respondere på matematisk innhold/informasjon/ideer som er presentert på et mangfold av måter”(Bjørkeng, 2013)

Gjennomsnittsscoren for hele landet ble 278 poeng.

Tabell 2: Tallforståelse

Nivå	0	1	2	3	4	5
testresultat	<176	176-225	226-275	276-325	326-375	376-500
prosentandel Norge	4	10	29	38	16	2
prosentandel OECD	5	12	34	39	11	1

Problemløsningsferdigheter i teknologirike miljøer er definert som:

”Evnen til å bruke digital teknologi, kommunikasjonsverktøy og nettverk til å tilegne seg og evaluere informasjon, kommunisere med andre og utføre praktiske oppgaver. Testen fokuserer på evnen til å løse problemer med private,

arbeidsmessige og samfunnsmessige formål, ved å sette opp egnede mål og planer, samt oppnå tilgang til og bruke informasjon gjennom datamaskiner og datanettverk.”(Bjørkeng, 2013)

Nivåene er inndelt på samme måte som ved de andre ferdighetene, men det er her kun tre nivå ettersom flere individer ikke var i stand til å gjennomføre testen på grunn av manglende datakunnskaper.

Tabell 3: Problemløsning i teknologiske miljøer

Nivå	ikke klassifisert	1	2	3
testresultat		<291	291-340	340<
prosentandel Norge	27,3	31,8	34,9	6,1
prosentandel OECD	36,5	29,4	28,2	5,8

5.1.2 Vekting

Det legges inn en vekting på utvalget basert på trekksannsynlighet og frafall slik at utvalget skal gi et best mulig bilde av populasjonen. Siden trekksannsynligheten var ulik for ulike alders- og utdanningsgrupper, legges det til et utvalgsvekt lik den inverse trekksannsynligheten. I tillegg justeres det for ferdighetsrelaterte eller ukjente frafall fra bruttoutvalget. PIAAC har egne krav til denne vektingen og kalibreringen (Gravem og Lagerstrøm, 2013).

5.2 Tilrettelegging av datamaterialet

Jeg vil ta utgangspunkt i resultatene fra estimeringene gjort i Hanushek et al. (2013) ettersom disse er gjort med det samme datamaterialet, for å enklere kunne undersøke og sammenlikne modellformuleringer. På bakgrunn av dette vil jeg benytte det samme utvalget fra datasettet. Dette gjøres ved å trimme datasettet for topp og bunn 1% av inntekt for å få et mer balansert utvalg og fjerne resultater av feiltasting. Jeg vil så ta utgangspunkt i individer i full jobb i sin beste alder, ”prime age”. Som nevnt i kapittel 2 ”Prime age” defineres som alderen mellom 35-54, da inntekten i denne alderen gir et best mulig bilde på livstidsinntekt. Individer som er yngre enn dette vil være i starten av sin karriere og derfor sannsynligvis øke i lønn senere i karrieren. Det er vist at de flinkeste individene vil øke raskere og mer i lønn, og dermed vil inntekten tidlig i karrieren gi et skjevt bilde av avkastningen av ferdigheter på livstidsinntekt.(Haider og Solon, 2006) En ulempe med å kun benytte dette delutvalget er at det vil underestimere effekten av at høyere ferdigheter gir høyere sannsynlighet for å få jobb. Full jobb defineres som som 30 arbeidstimer eller mer i uken.

5.3 Variabler

5.3.1 Avhengig variabel

Den avhengige variabelen som benyttes i de økonometriske modellene er logaritmisk verdi av inntekt, i form av timelønn uten bonuser. Bakgrunnen for dette ble begrunnet i kapittel 2. Ved å kun se på timelønn dannes et klarere bilde av avkastningen av ferdigheter ettersom det luker ut effekten av at økt inntekt også avhenger av antall arbeidstimer i uken. Den logaritmiske funksjonen gir et godt bilde av sammenhengen mellom utdanning og inntekt, og i tillegg resultater som er enkle å tolke. Som vist i den deskriptive statistikken i tabell 4 er gjennomsnittstidelønnen i utvalget gitt ved 235.975 med standardavvik 75.304.

5.3.2 Forklaringsvariabler

Forklaringsvariablene i modellen er resultatene fra ferdighetstestene i PIAAC-undersøkelsen; leseferdigheter, tallforståelse og problemløsning i teknologiske miljøer, gitt ved variablene henholdsvis *lese*, *tall* og *data*. Som i Hanushek et al. (2013) standardiseres variablene til en fordeling med gjennomsnitt 0 og standardavvik 1 for en enklere tolkning. Da vil koeffisientene i estimeringen kunne tolkes som prosentvis endring i inntekt ved ett standardavviks endring i ferdigheten. Ved å benytte denne variabel-formen forenkler jeg også muligheten for sammenlikning med modellformuleringen benyttet i Hanushek et al. (2013). I utvalget har *lese* en gjennomsnittsscore på 290,228 med standardavvik 41,712. *tall* har gjennomsnittsscore på 294,248 og standardavvik 48,992. *data* har gjennomsnittsscore på 290,057 og standardavvik 38,128.

5.3.3 Kontrollvariabler

Som kontrollvariabler tar jeg utgangspunkt i Mincer-modellen (Mincer, 1974) hvor det benyttes:

utdanning_i - Antall år individet har med utdanning. Variabelen har et gjennomsnitt i utvalget på 14,88 år og standardavvik på 2,32.

erfaring_i - Antall år individet har med arbeidserfaring. Variabelen har et gjennomsnitt i utvalget på 22,185 år og standardavvik på 7,385.

erfaring_i² - Kvadratisk ledd med arbeidserfaring. Som diskutert i kapittel 2 kan det tenkes at erfaring ikke har en lineær effekt på inntekt, men at den er avtakende. I regresjonene deler jeg verdiene av denne variabelen på 100 for å tydeligere få frem effekten.

kvinne_i - En dummyvariabel lik 1 dersom individet er kvinne. Variabelen har et gjennomsnitt i utvalget på 0,44 og standardavvik på 0,497.

5.3.4 Instrumentvariabel

Som instrument for å adressere den mulige endogeniteten i forklaringsvariabelen av interesse vil jeg benytte: *ParEd* - Foreldres utdanning. Denne variabelen kan ta tre verdier etter hvilket utdanningsnivå foreldrene til individet har:

- 1 - Ingen av foreldrene har videregående utdanning.
- 2 - Minst en av foreldrene har videregående utdanning som høyeste fullførte utdanning.
- 3 - Minst en av foreldrene har fullført høyere utdanning.

For å benytte dette som et instrument genererer jeg tre dummyvariabler:

ParEdLo_i Er lik 1 dersom individet har foreldre med lav utdanning, kategori 1.

ParEdMed_i Er lik 1 dersom individet har foreldre med middels utdanning, kategori 2.

ParEdHi_i Er lik 1 dersom individet har foreldre med høy utdanning, kategori 3.

Ved å inkludere to av disse dummyvariablene i en førstestegslikning kan jeg da estimere effekten av disse på individets ferdigheter, for så å benytte den predikerte ferdigheten fra 1. steget i stedet for faktiske ferdigheter i estimeringen av strukturlikningen.

5.3.5 Kohorteffekter

Som forklart i kapittel 4 vil jeg utvide modellen i forhold til formuleringen i (Hanushek et al., 2013) ved å kontrollere for kohorteffekter. Jeg inkluderer disse effektene ved å dele utvalget inn i fem kohorter.

kohort1: 35-38 år

kohort2: 39-42 år

kohort3: 43-46 år

kohort4: 47-50 år

kohort5: 51-54 år

Jeg inkluderer dummyvariabler for hver enkelt av disse, bortsett fra referansegruppen, som er lik én dersom individet er del av kohorten. På denne måten vil jeg studere separate effekter på de forskjellige kohortene.

5.4 Deskriptiv statistikk

Tabell 4: Deskriptiv statistikk

Variabel	Gjennomsnitt	Standardavvik	N
timelønn	235,972	75,304	1521
lese	290,228	41,712	1521
tall	294,248	48,992	1521
data	290,057	38,128	1417
ferdighet	293,759	37,714	1417
utdanning	14,88	2,32	1520
erfaring	22,185	7,385	1520
alder	44,226	5,595	1521
arbeidstimer	39,992	6,513	1521
kvinne	0,44	0,497	1521
offentlig	0,363	0,481	1521

Utvalget er gitt ved individer mellom 35-54 år i full jobb, vektet etter PIAACs utvalgsvekter. Full jobb defineres som 30 arbeidstimer eller mer i uken.

6 Empiriske Resultater

Jeg vil i dette kapittelet presentere mine empiriske resultater. Jeg vil ta utgangspunkt i en MKM-estimering av modellen gitt ved likning (18) presentert i kapittel 4. Jeg vil så så å teste om Mincer-modellen og formuleringen i Hanushek et al. (2013) er gyldige forenklinger av dette eller om det finnes andre, bedre spesifiseringer for å forklare avkastningen av ferdigheter. Jeg vil så i kapittel 6.4 utvide modellen og se nærmere på ferdighetseffektene før jeg i kapittel 6.5 ser nærmere på en instrumentvariabel-estimering. I kapittel 6.6 vil jeg så se nærmere på heterogene effekter på avkastningen av ferdigheter mellom undergrupper i utvalget.

6.1 Grunnmodell

Jeg tar utgangspunkt i grunnmodellen forklart i kapittel 4:

$$\begin{aligned} \ln y_i = & \beta_0 + \beta_1 \textit{lese}_i + \beta_2 \textit{tall}_i + \beta_3 \textit{data}_i + \beta_4 \textit{erfaring}_i + \beta_5 \textit{erfaring}_i^2 + \delta_1 \textit{kvinne}_i \\ & + \delta_2 \textit{kohort2}_i + \delta_3 \textit{kohort3}_i + \delta_4 \textit{kohort4}_i + \delta_5 \textit{kohort5}_i + u_i \end{aligned} \quad (27)$$

Interesseparameterne er β_1 , β_2 og β_3 som angir påvirkningen ferdighetene har på inntekt. Jeg vil hen vise til denne effekten som avkastning av ferdigheter, men det er ikke tatt hensyn til de eventuelle kostnadene som ligger bak å tilegne seg ferdighetene.

I tabell 5 vises resultatene fra MKM-estimeringen av likning (27) som modell 1. Resultatene viser at alle ferdighetene gir positiv effekt på inntekten, men koeffisienten for leseferdigheter, β_1 er ikke signifikant. Dette skyldes sannsynligvis problemet med sterk korrelasjon mellom ferdighetene, slik at avkastningen av leseferdigheter fanges opp av avkastningen av de andre ferdighetene, *tall* og *data*. Koeffisientene for avkastningen av tallforståelse og problemløsning kommer derimot ut som signifikante på 1% signifikansnivå. Estimaten sier oss da at en økning med ett standardavvik i tallforståelse eller problemløsning, gitt de andre kontrollvariablene, vil gi en økt timelønn med henholdsvis 5,3% og 6,6%.

Koeffisienten for arbeidserfaring, β_4 , er positiv og signifikant og sier da at økt arbeidserfaring med ett år, alt annet gitt, vil øke timelønnen med 2%. Jeg har derimot argumentert for at denne effekten ikke er lineær, noe som kommer frem ved det kvadratiske erfaringsleddet, β_5 som er negativ og signifikant lik -0,048. Det tyder altså på at den positive effekten av arbeidserfaring på inntekten, som antatt er avtakende.

Tabell 5: Generell grunnmodell, avkastning av ferdigheter estimert ved MKM.

Variabel	Modell 1	Modell 2	Modell 3	Modell 4
lese	0.014 (0.02)	0.013 (0.02)		
tall	0.053*** (0.02)	0.022 (0.02)		
data	0.066*** (0.01)	0.052*** (0.01)		
erfaring	0.020*** (0.01)	0.025*** (0.01)	0.029*** (0.01)	0.029*** (0.01)
erfaring ²	-0.048*** (0.01)	-0.046*** (0.01)	-0.056*** (0.01)	-0.055*** (0.01)
kvinne	-0.135*** (0.02)	-0.151*** (0.01)	-0.160*** (0.01)	-0.160*** (0.01)
kohort2	0.012 (0.02)	-0.005 (0.02)	-0.012 (0.02)	
kohort3	0.058** (0.03)	0.019 (0.02)	-0.001 (0.02)	
kohort4	0.084*** (0.03)	0.031 (0.03)	0.003 (0.03)	
kohort5	0.102*** (0.03)	0.026 (0.03)	-0.002 (0.03)	
utdanning		0.041*** (0.00)	0.056*** (0.00)	0.056*** (0.00)
konstant	5.243*** (0.063)	5.243*** (0.063)	5.243*** (0.063)	5.243*** (0.063)
N	1417	1416	1519	1519
R ²	0.22	0.29	0.25	0.25

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

Koeffisienten for kjønn, δ_1 , er lik -0,135 og signifikant på 1% signifikansnivå. Alt annet likt, vil altså det å være kvinne gi 13,5% lavere timelønn. Dette er ikke nødvendigvis en kausal virkning av det å være kvinne, men kan skyldes en effekt av uobserverbare variabler som at flere kvinner arbeider i lavere lønnede yrker enn menn.

Kohorteffektene blir større og mer signifikante i senere kohorter. *kohort1* er referansegruppen, slik at effektene av de andre kohort-dummyene må tolkes i forhold til denne. *kohort2* er svakt negativ, men ikke signifikant. *kohort3*, *kohort4* og *kohort5* er gitt ved henholdsvis 0,58, 0,84 og 0,102. Dermed vil, ved alt annet likt, det å tilhøre kohort 3, 4 eller 5 gi henholdsvis 5,8%, 8,4% og 10,2% høyere timelønn enn individer i kohort 1. Det at kohorteffektene er så jevnt stigende kan tyde på at de kun fanger opp en alderseffekt og

at det dermed ikke er noen annen variasjon mellom kohorter. Som foreslått av Boockmann og Steiner (2006) kan dette også komme av rigide lønninger i offentlig sektor, hvor man får økt lønn etter alder. Dermed er det mulig at kohorteffektene kan komme tydeligere frem ved å kun studere individer i privat sektor.

I modell 2 har jeg også inkludert antall års utdanning, *utdanning*. Dette gir en positiv og signifikant koeffisient lik 0.041, som tilsier at ett års ekstra utdanning, alt annet likt, vil gi økt timelønn med 4.1%. Denne koeffisienten angir effekten av antall års utdanning på inntekt til et gitt ferdighetsnivå. Altså tyder dette på en signaleffekt eller andre uobserverbare effekter av utdanning. Inkluderingen av antall år utdanning reduserer koeffisientene for ferdighetene, slik at kun β_3 er signifikant og lik 0.053. Inkluderingen av utdanning øker også koeffisienten for arbeidserfaring noe, fra 0,020 til 0,025. I Modell 2 er heller ingen av kohort-dummyene signifikante. Dette tyder på at *utdanning* og *erfaring* i denne modellen fanger opp en alderseffekt som i Modell 1 inngikk i kohort-dummyene.

I modell 3 og 4 har jeg fjernet variablene for ferdighetsmålene. Modell 4 inneholder heller ikke kohort-dummys, og angir dermed den klassiske Mincer-modellen. I modell 3 er kohort-effektene inkludert, men ingen av dummyvariablene er signifikant ulik 0 og gir dermed veldig liten forskjell mellom modellene. Ved å fjerne ferdighetsmålene øker koeffisienten for både utdanning og erfaring sammenliknet med modell 2, til henholdsvis 0,029 og 0,056. Ved å ikke ta hensyn til ferdighetene og alt annet likt vil altså ett år ekstra med utdanning føre til en økt timelønn med 5,6%.

Det er også interessant å se effekten av arbeidserfaring sammenliknet med modellene hvor ferdigheter er inkludert. Ettersom koeffisienten for *erfaring* reduseres med 31% fra modell 3 og 4 til modell 1 ved å inkludere ferdighetene, tyder det på at 31% av avkastningen av økt arbeidserfaring kommer av økte ferdigheter. Dette strider mot resultatene fra Green og Riddel (2001) som konkluderte med at arbeidserfaring hadde svært liten effekt på de kognitive ferdighetene. På samme måte reduseres koeffisienten for utdanning med 26,8% når ferdighetene inkluderes i modellen. Dette tyder på at utdanning har en stor effekt på ferdighetene, men også at en stor del av avkastningen av utdanning er en direkte effekt på inntekt som ikke virker gjennom de kognitive ferdighetene.

Problemet med kollineariteten mellom ferdighetene i modell 1 kan kontrolleres for ved å kun benytte ett ferdighetsmål. I Hanushek et al. (2013) benyttes kun tallforståelse, noe som blant annet begrunnes med at denne ferdigheten lettest lar seg sammenlikne på tvers av landegrenser. For å teste om dette er en gyldig forenkling benytter jeg da en F-test på hypotesen:

$$H_0 : \beta_1 = \beta_2 = \beta_3$$

Dette gir en F-verdi på $F(2, 1406) = 1,70$. og en kritisk verdi ved 5% signifikansnivå $F_{crit} = 3$. H_0 kan forkastes dersom $F_{obs} > F_{crit}$. $1,70 < 3$ altså kan jeg ikke forkaste

Hypotesen om at alle de kognitive ferdighetene gir lik avkastning.

Ettersom jeg ikke kan si at det er noen forskjell i avkastning av de forskjellige ferdighetene innfører jeg derfor en variabel, *ferdighet* for gjennomsnittsscore.

$$ferdighet_i = \frac{lese_i + tall_i + data_i}{3}$$

Ved å benytte denne variabelen kan jeg lettere få frem avkastningen av kognitive ferdigheter. Jeg standardiserer også denne variabelen på samme måte som jeg gjorde for de enkelte ferdighetene slik at koeffisienten angir den prosentvise endringen i inntekt ved ett standardavviks endring i gjennomsnittsresultatet.

I Modell 1 i tabell 6 benyttes gjennomsnittsresultat på ferdighetstestene som forklaringsvariabel.

Tabell 6: Avkastning av ferdigheter med gjennomsnittsresultat som ferdighetsmål

Variabel	Modell 1	Modell 2
ferdighet	0.117*** (0.007)	0.078*** (0.008)
erfaring	0.021*** (0.006)	0.025*** (0.006)
erfaring ²	-0.049*** (0.012)	-0.047*** (0.012)
kvinne	-0.137*** (0.015)	-0.151*** (0.014)
kohort2	0.008 (0.023)	-0.010 (0.022)
kohort3	0.052** (0.025)	0.013 (0.024)
kohort4	0.079*** (0.027)	0.025 (0.026)
kohort5	0.096*** (0.033)	0.017 (0.032)
utdanning		0.041*** (0.003)
konstant	5.247*** (0.062)	4.567*** (0.081)
N	1417	1416
R ²	0.217	0.286

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

Koeffisienten for *ferdighet* er lik 0,117 og signifikant på 1% signifikansnivå. Dette vil si at, alt annet likt, vil en økt gjennomsnittsscore med ett standardavvik øke timelønnen med 11,7%. De andre parameterne er tilnærmet uendret sammenliknet med tabell 5, men kohort-effektene er noe reduserte.

I modell 2 har jeg igjen inkludert antall års utdanning. Koeffisienten for *utdanning* er lik 0,041 og signifikant. Inkluderingen av utdanning reduserer avkastningen av ferdigheter noe, til 0,078. Dette tyder på at koeffisienten for *utdanning* fanger opp variasjon som ble forklart av *ferdighet* i modell 1. Koeffisienten for utdanning er uendret sammenliknet med modell 2 i tabell 5, men det er nå tydeligere hvor stor effekt utdanningen har sammenliknet med ferdighetene. Til et gitt nivå av ferdigheter, og de andre kontrollvariablene, vil altså ett års økt utdanning øke timelønnen med 4,1%.

Videre tester jeg også for om kohort-effektene har noen individuell, signifikant effekt på avkastningen, eller om man kan anta at effektene er like. Jeg vil ved en F-test teste hypotesen:

$$H_0 : \delta_2 = \delta_3 = \delta_4 = \delta_5 = 0$$

Dette gir en F-verdi på $F(4, 1408) = 3,38$. og en kritisk verdi ved 5% signifikansnivå $F_{crit} = 2,6$. H_0 kan forkastes dersom $F_{obs} > F_{crit} = 3,38 > 2,6$, altså kan jeg forkaste hypotesen om at kohorteffektene er lik null. Det er forskjell i lønn til forskjellige kohorter, selv når det tas hensyn til arbeidserfaring og utdanningsnivå.

6.2 Kohortspesifikk effekt av ferdigheter

For å se nærmere på hvordan kohorteffektene påvirker avkastning av ferdigheter vil jeg også inkludere interaksjonsledd mellom kohortene og testresultatet, $kohort * ferdighet_i$. Interaksjonsleddene vil få frem en eventuell forskjell i avkastning av ferdigheter mellom kohorter. Jeg vil derfor estimere likningen:

$$\begin{aligned} \ln y_i = & \beta_0 + \beta_1 ferdighet + \beta_4 erfaring_i + \beta_5 erfaring_i^2 + \delta_1 kvinne_i \\ & + \delta_2 kohort2_i + \delta_3 kohort3_i + \delta_4 kohort4_i + \delta_5 kohort5_i + \delta_6 kohort2 * ferdighet_i \\ & + \delta_7 kohort3 * ferdighet_i + \delta_8 kohort4 * ferdighet_i + \delta_9 kohort5 * ferdighet_i + u_i \end{aligned} \quad (28)$$

Resultatene fra denne estimeringen er gjengitt i tabell 7. Koeffisienten for *ferdighet*, $\beta_1 = 0,112$, angir her avkastningen av ferdigheter for individer i kohort1. En endring i gjennomsnittsscoren med ett standardavvik vil endre timelønnen med 11,2% altså noe lavere enn for hele utvalget i tabell 6. Koeffisientene for interaksjonsleddene, $kohort * ferdighet_i$, angir nå differansen i avkastningen av ferdigheter for de andre kohor-

tene sammenliknet med referansegruppen. Disse effektene er små og positive bortsett fra for kohort3 hvor den er negativ, men hver for seg er ingen av koeffisientene for interaksjonsleddene signifikante, selv på 10% signifikansnivå.

Tabell 7: Avkastning av ferdigheter inkludert kohorteffekter

Variabel	Koeffisient
ferdighet	0.112*** (0.019)
erfaring	0.020*** (0.006)
erfaring ²	-0.049*** (0.013)
kvinne	-0.137*** (0.015)
kohort2	0.007 (0.026)
kohort3	0.051* (0.027)
kohort4	0.077*** (0.028)
kohort5	0.096*** (0.034)
kohort2*ferdighet	0.003 (0.025)
kohort3*ferdighet	-0.010 (0.023)
kohort4*ferdighet	0.013 (0.025)
kohort5*ferdighet	0.025 (0.027)
konstant	5.251*** (0.065)
N	1417
R ²	0.218

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

Ved en F-test vil jeg teste om kohorteffektene er signifikant forskjellige fra 0 med hypotesen:

$$H_0 : \delta_6 = \delta_7 = \delta_8 = \delta_9 = 0$$

Dette gir en F-verdi på $F(4, 1404) = 0,71$. og en kritisk verdi ved 5% signifikansnivå $F_{crit} = 2,37$. H_0 kan forkastes dersom $F_{obs} > F_{crit}$ $0,71 < 2,37$ Dermed kan jeg med god

margin beholde hypotesen om at interaksjonsleddene ikke er signifikant forskjellig fra null. Da kan jeg også konkludere med at det ikke er noen forskjell i avkastning av ferdigheter mellom kohorter. Som diskutert i kapittel 2 kan kohort-effektene peke i forskjellige retninger avhengig av variasjon i tilbud og etterspørsel i arbeidsmarkedet, men dette tyder altså på at slike effekter ikke har innvirkning på avkastningen av ferdigheter i utvalget.

6.3 Alderseffekter

Jeg konkluderte etter resultatene i tabell 6 med at det er variasjon i inntekt mellom kohorter, men at den jevne stigningen over kohorter kan tyde på en ren alderseffekt. Jeg inkluderer derfor en variabel $alder_i$ for individets alder i stedet for kohortdummyene.

Tabell 8: Avkastning av ferdigheter inkludert alder som kontrollvariabel

Variabel	Modell 1	Modell 2
ferdigheter	0.117*** (0.007)	0.077*** (0.008)
erfaring	0.020*** (0.005)	0.025*** (0.006)
erfaring ²	-0.049*** (0.012)	-0.047*** (0.012)
kvinne	-0.138*** (0.015)	-0.150*** (0.014)
alder	0.006*** (0.002)	0.001 (0.002)
utdanning		0.041*** (0.003)
konstant	5.011*** (0.086)	4.520*** (0.089)
N	1417	1416
R ²	0.216	0.285

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

Koeffisienten for *alder* angir da endringen i timelønn ved endring i alder. I modell 1 er koeffisienten lik 0.006 og signifikant på 1% signifikansnivå. Det tyder på at en økt alder med ett år gir en økt timelønn med 0,6%. Dette virker konsistent med koeffisientene for kohorter i tabell 6. Koeffisienten for ferdigheter er uendret sammenliknet med modell 1 i tabell 6. Det er også relativt små endringer i de andre kontrollvariablene.

I modell 2 i tabell 8 inkluderes igjen utdanning som kontrollvariabel. Denne spesifiseringen gjør at koeffisienten for *alder* ikke lenger har signifikant effekt på inntekten.

6.4 Ikke-linearitet i ferdighetseffekter

Jeg har til nå konkludert med at de forskjellige ferdighetene har tilnærmet lik påvirkning på inntekt og at det dermed er bedre å benytte en gjennomsnittsscore for å studere avkastningen. Det er derimot ikke gitt at denne avkastningen er lineær. Flere tidligere studier har tydet på en avtakende avkastning, blant annet i (Murnane et al., 2000). Dermed er det interessant å utvide modellen med alternative spesifikasjoner av ferdighetene.

En mulig utvidelse for å ta høyde for denne ikke-lineariteten er å legge til høyere polynom av ferdighetsmålet:

Tabell 9: Avkastning av ferdigheter, ikke-lineære effekter

Variabel	Modell 1	Modell 2
ferdighet	0.117*** (0.007)	0.133*** (0.010)
ferdighet ²	-0.006 (0.006)	-0.011** (0.005)
ferdighet ³		-0.005*** (0.002)
erfaring	0.020*** (0.006)	0.020*** (0.020)
erfaring ²	-0.048*** (0.012)	-0.048*** (0.012)
kvinne	-0.139*** (0.015)	-0.138*** (0.015)
alder	0.007*** (0.002)	0.006*** (0.002)
konstant	5.019*** (0.087)	5.029*** (0.086)
N	1417	1417
R ²	0.216	0.219

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

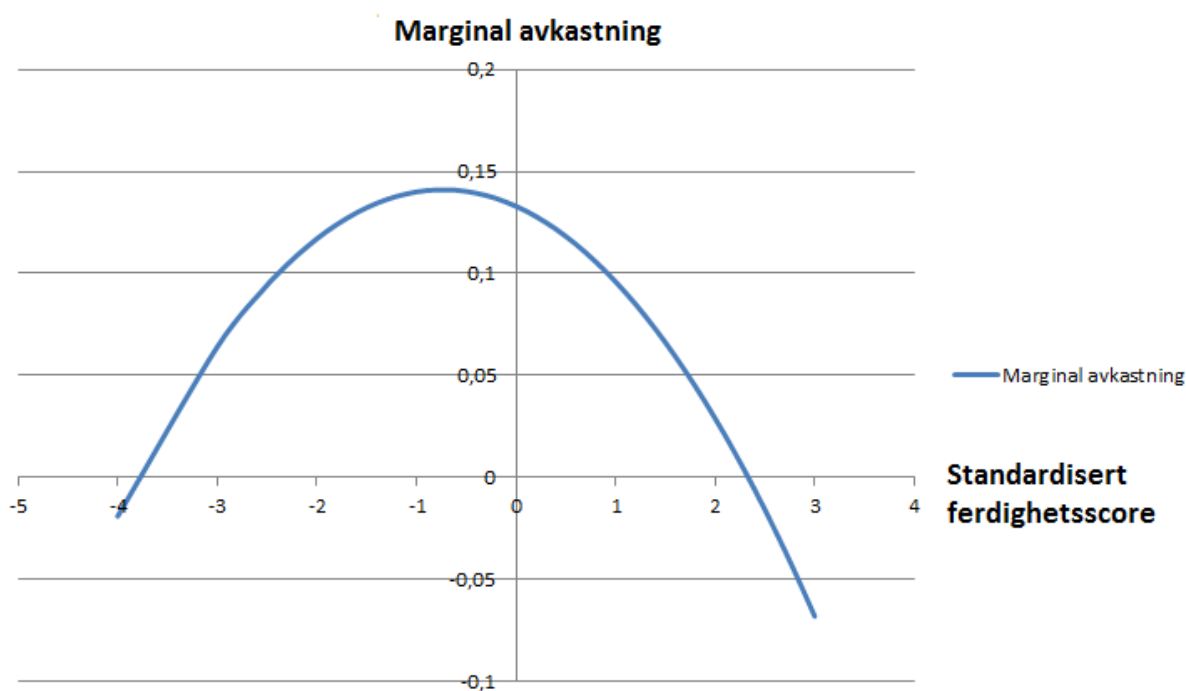
Signifikansnivå: *: 10% **: 5% ***: 1%.

I modell 1 i tabell 9 har jeg lagt til et kvadratisk ledd av ferdighetsscoren. Koeffisienten for *ferdighet* er nå lik 0,117 som vil si at en endring i gjennomsnittsscoren på ferdighetene med ett standardavvik vil endre timelønnen med 11,7%. Dette er ingen endring sammenliknet med tabell 6. Koeffisienten for *ferdighet*² er lik -0,006, noe som kan tyde på at mistanken om en positiv men avtakende avkastning av ferdigheter stemmer, men effekten er ikke signifikant.

I modell 2 i tabell 9 tar jeg høyde for en mer komplisert funksjonsform, ved å også inkludere

et tredjegradspolynom av ferdighetene. Ved å inkludere variabelen *ferdighet*³ inngår både denne og *ferdighet*² signifikant i modellen. Dette er altså en mer presis spesifikasjon av modellen.

Koeffisienten for *ferdighet* er i modell 2 i tabell 9 gitt ved 0,133 som er en økning sammenliknet med resultatene fra tabell 6 og 9. Med spesifiseringen jeg har benyttet hvor forklaringsvariabelen er standardisert, angir nå koeffisienten, β_1 , avkastningen av økte ferdigheter for individer med gjennomsnittsscoren for utvalget. Alle ferdighetsnivå som er lavere enn gjennomsnittet vil ha negative verdier, og høyere ferdighetsnivå vil gi positive verdier. Ved å sette inn for verdier for ferdighetsnivåene kan vi se nærmere på marginaleffektene av økte ferdigheter ved forskjellige ferdighetsnivå.



Figur 1: Marginal avkastning av ferdigheter

Gjennomsnittsscoren på ferdighetstestene for utvalget er 293,8, markert ved 0 i grafen. Deres avkastning av økte ferdigheter med ett standardavvik er gitt ved koeffisienten for *ferdighet* i tabell 9, 13,3% økt timelønn. X-aksen angir standardavvik fra gjennomsnittet. Estimaten tyder altså på at individer med veldig lav ferdighetsscore vil ha relativt lav marginal avkastning. Den marginale avkastningen er høyest ved en score 0,7 standardavvik lavere enn gjennomsnittsscoren, altså en score på 267,41. Ved ferdighetsnivå høyere enn dette vil den marginale avkastningen avta.

En annen mulig formulering for å ta høyde for ikke-lineariteten i ferdigheter er å innføre dummyvariabler for forskjellige nivå på ferdighetene. På denne måten kan jeg kontrollere for eventuell variasjon i avkastning av ferdigheter til forskjellige ferdighetsnivå. Jeg deler gjennomsnittsscoren opp i 5 nivåer slik som leseferdigheter og tallforståelse ble inndelt

i PIAAC-undersøkelsen, som forklart i kapittel 5, og inkluderer så dummyvariabler for de forskjellige nivåene i stedet for den kontinuerlige variabelen *ferdighet*. De forskjellige ferdighetsnivåene er gitt ved: *ferdighet*₀: 0-176. *ferdighet*₁: 176-225. *ferdighet*₂: 226-275. *ferdighet*₃: 276-325. *ferdighet*₄: 326-375. *ferdighet*₅: 376-500.

Tabell 10: Avkastning av ferdigheter til forskjellige ferdighetsnivå

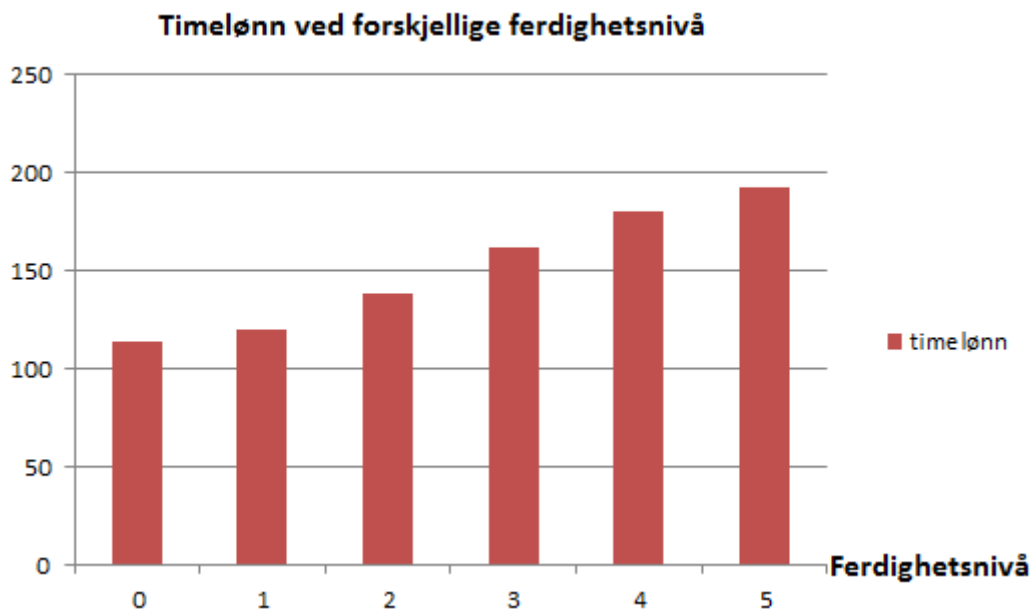
Variabel	Koeffisient
<i>ferdighet</i> ₁	0.050 (0.088)
<i>ferdighet</i> ₂	0.196** (0.084)
<i>ferdighet</i> ₃	0.353*** (0.083)
<i>ferdighet</i> ₄	0.457*** (0.083)
<i>ferdighet</i> ₅	0.523*** (0.091)
erfaring	0.021*** (0.006)
erfaring ²	-0.049*** (0.012)
kvinne	-0.142*** (0.015)
alder	0.006*** (0.002)
konstant	4.735*** (0.115)
N	1417
R ²	0.211

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

I tabell 10 er *ferdighet*₀ referansegruppen, altså individer med en gjennomsnittsscore på under 176. Koeffisientene for de andre ferdighetsnivåene angir da den prosentvise forskjellen i inntekt mellom dette ferdighetsnivået og referansegruppen. I følge resultatene i tabell 10 vil individene med ferdighetsnivå *ferdighet*₁, alt annet likt, ha en 5% høyere inntekt enn individene med ferdighetsnivå 0. Tilsvarende vil individer med ferdighetsnivå *ferdighet*₂, *ferdighet*₃, *ferdighet*₄ og *ferdighet*₅ ha en henholdsvis 19,6%, 35,3%, 45,7% og 52,3% høyere timelønn enn referansegruppen, alt annet likt.

Differansen mellom avkastningen til de forskjellige ferdighetsnivåene er da størst mellom *ferdighet*₂ og *ferdighet*₃. 35,3 – 19,6 = 15,7 prosentpoeng. Dette tyder, som resultatene



Figur 2: Timelønn ved forskjellige ferdighetsnivå, gitt kontrollvariablene

fra tabell 9, på at avkastningen av ferdigheter først er tiltakende, for så å bli avtakende etter et visst ferdighetsnivå.

6.5 Instrumentvariabel-estimering

Som diskutert i kapittel 2 og 4 er det flere potensielle kilder til skjevhet i modellen. Både utelatt variabel og målefeil kan føre til at MKM-estimatet feilestimerer avkastningen av ferdigheter. For å forsøke å kontrollere for dette benytter jeg meg av en instrumentvariabel-estimering. Jeg benytter meg av variabler for foreldres utdanning som instrumentvariabel, som forklart i kapittel 5.3.4.

Med denne spesifiseringen får koeffisienten for *ferdighet* over dobbelt så høy verdi som i modell 6. β_1 er nå gitt ved 0,261, altså vil en økning i gjennomsnittsscoren på ferdighetstestene med ett standardavvik, alt annet gitt, vil øke timelønna med 26,1%. En svakhet med modellformuleringen i modell 1 er at utdanning ikke er inkludert som kontrollvariabel. Ettersom foreldres påvirkning sannsynligvis vil virke gjennom at individer vil velge høyere utdanning, vil dermed også foreldres utdanning, *pared* kunne være korrelert med restleddet i formuleringen. For ta høyde for dette vil jeg derfor også inkludere antall års utdanning som en forklaringsvariabel. Resultatene fra dette estimatet er gitt ved IV-2.

Koeffisienten *utdanning* i IV-2 i tabell 11 er lik 0,019, men ikke signifikant. Derimot har antall års utdanning stor innvirkning på instrumentet via førstestegslikningen. Der gir antall års utdanning en positiv og signifikant effekt på ferdigheter. Med denne modellformuleringen er fremdeles avkastningen av ferdigheter mye høyere enn ved grunnmodellen,

Tabell 11: Instrumentvariabel-regresjon

Variabel	IV-1	IV-2
ferdighet	0.261*** (0.039)	0.213*** (0.070)
erfaring	0.016** (0.007)	0.019*** (0.007)
erfaring ²	-0.033** (0.014)	-0.036** (0.015)
kvinne	-0.106*** (0.019)	-0.118*** (0.023)
alder	0.008*** (0.002)	0.005*** (0.003)
utdanning		0.017 (0.013)
konstant	4.912*** (0.102)	4.727*** (0.154)
N	1400	1399
R ²	0.023	0.142

Tabellen presenterer resultater fra IV-regresjoner, vektet ved PIAACs utvalgs-vekter. Foreldres utdanning benyttes som instrument for ferdigheter. Avhengig variabel er logaritmisk timelønn.

Estimerte robuste standardavvik er oppgitt i parentes. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC. Signifikansnivå: *: 10% **: 5% ***: 1%.

tabell 6. Koeffisienten for *ferdighet* er nå lik 0,213 som vil si at en endring i gjennomsnittsscoren med ett standardavvik vil endre timelønnen med 21,3%. Dette tyder på skjevheter i MKM-estimeringen

Ved å inkludere variabelen for utdanning i modellen vil det føre til at jeg har en ny potensielt endogen variabel i likningen. Som diskutert tidligere er ikke utdanning eksogent bestemt, men avhenger av individuelle utdanningsvalg. Begge de to IV-formuleringene har dermed fremdeles potensielle kilder til skjevhet. Det optimale ville være å instrumentere både utdanning og ferdigheter med separate instrumenter, men dette er ikke noe jeg har tilgjengelig.

Som nevnt i kapittel 4 må variabelen *pared* oppfylle relevans- og eksklusjonskriteriene. Relevanskriteriet kan testes ved hjelp av førstestegslikningen, gitt i vedlegg A. Her vises foreldres utdannings påvirkning på ferdigheter, samt de andre kontrollvariablene. Foreldres utdanning er inkludert ved dummyvariabler, slik at referansegruppen er *paredlo*. Koeffisientene for *paredmed* og *paredhi* er på henholdsvis 0,166 og 0,298 og angir da økningen i ferdighets-scoren sammenliknet med *paredlo*, gitt de andre kontrollvariablene. Begge koeffisientene er positive og signifikante og har dermed tydelig korrelasjon med ferdighetsnivået. En F-test på de to instrumentvariablene gir en F-verdi på 33,58 og 10,33

for henholdsvis. IV-1 og IV-2, altså holder relevanskriteriet med god margin.⁹

Eksklusjonskriteriet sier at instrumentene ikke kan påvirke den avhengige variabelen gjennom noen annen effekt enn via forklaringsvariabelen og kontrollvariabelen. Dette kan ikke testes direkte, så resultatene vil gjelde under antakelsen om at restriksjonen holder. Resultatene må derfor tolkes under forutsetningen om at det ikke eksisterer noen direkte effekt av foreldres utdanning på individets inntekt.

6.5.1 Hvorfor er IV-estimatene høyere enn ved MKM?

I kapittel 2 og 4 pekte jeg på flere mulige grunner til at MKM-estimeringen av modellen kunne gi skjeve estimater. Resultatet fra IV-2-estimeringen tyder på en 82% høyere avkastning av ferdigheter enn MKM-estimeringen fra tabell 6. Hanushek et al. (2013) pekte på mulige grunner til at MKM ville underestimere avkastningen av ferdigheter og at resultatene derfor burde tolkes som en nedre grense. Dette er på grunn av at ferdighetsmålet ikke kan tolkes som et perfekt mål på individenes humankapital og at det dermed er tilstede en målefeil. En målefeil vil, som nevnt i kapittel 2 og 4, gjøre estimatoren skjev mot null. At resultatene fra IV-estimeringen er mye høyere enn ved MKM-estimeringen tyder på at dette er tilfelle.

Som nevnt i kapittel 2 vil inkluderingen av ferdigheter også ha innvirkninger på utdanningsvalget og dermed kunne føre til skjevheter i MKM-estimatet. Denne underestimeringen vil da kunne skyldes en negative korrelasjonen mellom diskonteringsrente og optimal utdanning (Harmon et al., 2003).

I Card (1999) sammenliknes flere tilfeller hvor foreldres utdanning er benyttet som instrumentvariabel på avkastning av utdanning. I modellene varierer avkastningen av utdanning mellom individer, og den marginale avkastningen av antall års skolegang er en avtakende funksjon av skolegang. Harmon et al. (2003) argumenterer for at når instrumentet baseres på grunnlag av "deltakelse" i en behandlingsgruppe, (høyere utdannede foreldre gir redusert diskonteringsrente) vil IV-estimatoren på avkastning av utdanning være forskjellen i forventet inntekt mellom gruppene delt på forskjellen i forventet utdanning. Med denne tolkningen er det klart at IV-estimatoren kan overstige MKM-estimatoren dersom instrumentet påvirker en gruppe med relativt høy marginal avkastning. Dersom IV-estimeringen påvirker individer med et utdannings- eller ferdighetsnivå under gjennomsnittet, vil IV-estimatoren være høyere enn MKM-resultatet og reflekterer at individer med lavere utdanning har høyere diskonteringsrate

⁹referer til F-verdi > 10-regel

6.6 Analyse på delutvalg

Jeg har til nå studert avkastningen av ferdigheter til hele delutvalget og ikke differensiert mellom grupper. Ved å differensiere delutvalget i mindre grupper vil jeg kunne studere nærmere hvilke heterogene effekter som eksisterer i avkastningen av ferdigheter i arbeidsmarkedet.

6.6.1 Avkastning av ferdigheter etter sektor

Jeg vil først studere forskjeller i offentlig og privat sektor. Offentlig sektor har mer rigide lønninger¹⁰, noe som kan bety at individer med høye ferdigheter ikke belønnes like godt i form av lønn her som i privat sektor. Dette kan også føre til at de klokeste hodene velger seg til privat sektor. Som nevnt tidligere kan lønnssettingen også føre til at kohorteffektene kommer klarere frem i privat sektor.

Tabell 12: Avkastning av ferdigheter i offentlig og privat sektor

Variabel	Koeffisient	Koeffisient
ferdighet	0.134*** (0.009)	0.087*** (0.010)
offentlig	-0.051*** (0.015)	-0.099*** (0.015)
offentlig*ferdighet	-0.040*** (0.015)	-0.036** (0.014)
erfaring	0.019*** (0.006)	0.024*** (0.006)
erfaring ²	-0.047*** (0.012)	-0.044*** (0.012)
kvinne	-0.124*** (0.015)	-0.126*** (0.015)
alder	0.008*** (0.002)	0.002 (0.003)
utdanning		0.047*** (0.003)
konstant	4.992*** (0.087)	4.418*** (0.091)
N	1398	1397
R ²	0.23	0.32

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

¹⁰Referer til Bjarnes artikkel

I tabell 12 har jeg inkludert en dummyvariabel, *offentlig* lik 1 dersom individet arbeider i offentlig sektor. Jeg har også inkludert et interaksjonsledd, *offentlig*ferdighet* for å fange opp effekten av forskjeller i avkastning av ferdigheter i offentlig og privat sektor. Dette medfører at koeffisienten for *ferdighet* nå angir avkastning av ferdigheter for individer som arbeider i privat sektor. Denne koeffisienten er lik 0,134, altså vil en endring i gjennomsnittscoren på ferdigheter for individer i privat sektor endre timelønnen med 13,4%. Dette er 1,4 prosentpoeng høyere enn avkastningen i grunnmodellen, tabell 6. Avkastningen av ferdigheter for individer i offentlig sektor vil da være gitt ved koeffisienten for *ferdighet* i tillegg til interaksjonsleddet *offentlig*ferdighet* som angir tilleggseffekten av endringer i ferdigheter dersom individet arbeider i offentlig sektor. Dermed er den totale effekten gitt ved $0,134 - 0,04 = 0,094$. Altså vil en endring i gjennomsnittresultatet med ett standardavvik for individer i offentlig sektor endre timelønnen med 9,4%. Det er altså en tydelig lavere avkastning av ferdigheter i offentlig enn i privat sektor på hele 4 prosentpoeng.

Interaksjonsleddet viser også forskjell i inntekt ved å jobbe i offentlig sektor til et gitt ferdighetsnivå. Ved å benytte gjennomsnittsnivået for ferdigheter, tyder dette på at til et gitt ferdighetsnivå og de andre kontrollvariablene, vil det å arbeide i offentlig sektor gi en 5,7% lavere inntekt enn i privat sektor.

For å tydeligere få frem forskjellene i avkastning mellom privat og offentlig sektor kjører jeg også regresjonen separat på de to gruppene. Her inkluderer jeg igjen kohorteffektene for å undersøke om disse til større grad er til stede i privat sektor.

Resultatene Offentlig 1 og Privat 1 i tabell 13 tyder på, som jeg konkluderte med fra tabell 12, at avkastningen av ferdigheter er omtrent 4 prosentpoeng høyere i privat enn i offentlig sektor. Det er også store forskjeller i avkastningen av de andre kontrollvariablene. Koeffisienten for *erfaring* er på 0,021 og 0,16 for henholdsvis offentlig og privat sektor. Den avtakende marginalavkastningen av arbeidserfaring gitt ved andregradspolynommet er også større i offentlig sektor enn privat, henholdsvis -0,047 og -0,044.

I tabell 13 har jeg igjen inkludert kohort-dummys i stedet for aldersvariabelen for å se nærmere på om kohorteffektene kommer klarere frem i privat enn offentlig sektor. Kohorteffektene på inntekt varierer noe mellom sektorene. *kohort2* kommer ut med en høyere og mer signifikant effekt på inntekten i offentlig enn privat sektor. *kohort3* og *kohort4* derimot er klart høyere og med signifikante for privat sektor.

I Offentlig 2 og Privat 2 i tabell 13 har jeg igjen inkludert antall års utdanning som kontrollvariabel. Dette reduserer forskjellen i avkastning av ferdigheter mellom offentlig og privat sektor til 2,8 prosentpoeng. Koeffisienten for *utdanning* kommer ut høyere for privat enn offentlig sektor. Ett års økt utdanning til gitte ferdigheter og andre kontrollvariabler vil føre til en økt lønn med 4,2% og 5% i henholdsvis offentlig og privat sektor. Dette kan

Tabell 13: Avkastning av ferdigheter i offentlig og privat sektor

Variabel	Offentlig 1	Privat 1	Offentlig 2	Privat 2
ferdighet	0.094*** (0.01)	0.134*** (0.01)	0.057*** (0.014)	0.085*** (0.010)
erfaring	0.021** (0.01)	0.160** (0.01)	0.023** (0.011)	0.024*** (0.007)
erfaring ²	-0.047** (0.02)	-0.044*** (0.02)	-0.041* (0.023)	-0.046*** (0.014)
kvinne	-0.121*** (0.02)	-0.127*** (0.02)	-0.121*** (0.020)	-0.130*** (0.020)
kohort1	0.019 (0.03)	0.014 (0.03)	-0.003 (0.033)	0.000 (0.030)
kohort2	0.076** (0.04)	0.055 (0.03)	0.043 (0.036)	0.011 (0.033)
kohort3	0.085** (0.04)	0.108*** (0.04)	0.037 (0.038)	0.052 (0.035)
kohort4	0.065 (0.04)	0.148*** (0.05)	0.003 (0.040)	0.061 (0.048)
utdanning			0.042*** (0.005)	0.050*** (0.005)
konstant	5.180*** (0.11)	5.319*** (0.07)	4.493*** (0.124)	4.483*** (0.108)
N	554	844	554	843
R ²	0.24	0.21	0.33	0.30

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

tyde på at en signaliseringseffekt av utdanning i større grad er til stede i privat sektor.

Alle disse koeffisientene peker i retning av at det er mer fleksible lønninger i privat sektor som gjør at man får større avkastning av både ferdigheter og erfaring i privat enn i offentlig sektor. Inkludering av interaksjonsledd mellom kohort og ferdighet gir ingen signifikante resultater for hverken privat eller offentlig sektor, noe som tyder på at avkastningen av ferdigheter i Norge har vært stabil over tid.

Jeg kan teste om modellene for privat og offentlig sektor er tilstrekkelig forskjellige ved en Chow-test. Ved å ta utgangspunkt i regresjonenes SSR (sum of squared residuals), kan jeg sammenlikne forklaringskraften til de to modellene for enkelte sektorer med forklaringskraften til grunnmodellen, tabell 6 hvor hele delutvalget er inkludert. I praksis gjøres dette ved å inkludere interaksjonsledd mellom dummyvariabelen for offentlig sektor og alle kontrollvariablene for så å teste hypotesen om disse er signifikant forskjellig fra null.

(Resultatene vises i vedlegg C)

$$\begin{aligned}H_0 : \text{offentlig} &= \text{offentlig} * \text{ferdighet} = \text{offentlig} * \text{erfaring} \\ &= \text{offentlig} * \text{erfaring}^2 = \text{offentlig} * \text{kvinne} = \text{offentlig} * \text{alder} = 0\end{aligned}$$

Dette gir en F-verdi på $F(6, 1386) = 4,53$ og en kritisk verdi ved 5% signifikansnivå $F_{crit} = 2,10$. H_0 forkastes dersom $F_{obs} > F_{crit} = 4,53 > 2,10$ altså kan jeg forkaste hypotesen og konkludere med at det er signifikante forskjeller i avkastningen av ferdigheter i offentlig og privat sektor. En viktig begrensning i Chow-testen er at nullhypotesen ikke tillater noen forskjell mellom gruppene i det hele tatt. (Wooldridge, 2009, s. 245)

6.6.2 Avkastning av ferdigheter etter aldersgrupper

Jeg har til nå kun basert estimatene på individer i sin beste alder i arbeidslivet, *prime* mellom 35-54 år. For å undersøke om dette faktisk er det beste utgangspunktet for å estimere avkastning av ferdigheter på livstidsinntekt vil jeg også inkludere høyere og lavere aldersgrupper. *gammel* for individer mellom 55-64 og *ung* for individer mellom 20-34. Jeg inkluderer også interaksjonsledd mellom aldersgruppene og ferdighetene, *prime*ferdighet* og *gammel*ferdighet*.

Koeffisienten for *ferdighet* angir nå avkastning av ferdigheter for den yngste aldersgruppen. Denne er gitt ved 0,103 som vil si at for individer mellom 20-34 år vil en endring i ferdighetene med ett standardavvik øke timelønnen med 10,3%. Dette er 1,4 prosentpoeng lavere enn resultatet i tabell 6. Altså tyder det på at antakelsen jeg tidligere har gjort stemmer. Estimer basert på yngre, nyutdannede individer vil undervurdere avkastningen av ferdigheter på livstidsinntekt. For individer i "prime age" vil da avkastningen av økte ferdigheter med ett standardavvik være $0,103 + 0,017 = 0,12$, en 12% økning i timelønn. Dette er en noe høyere koeffisient enn i tabell 6, men der var også alder inkludert som kontrollvariabel. For individene i den høyeste aldersgruppen vil avkastningen av økte ferdigheter med ett standardavvik være $0,103 + 0,003 = 0,106$, en 10,6% økning i timelønn. Ingen av interaksjonsleddene er derimot signifikante, så noen direkte forskjeller i avkastningen av ferdigheter mellom aldersgrupper kan jeg ikke konkludere med. På samme måte som i tabell 7 tyder dette på at forskjellene i inntekt i stor grad skyldes aldersforskjeller og ikke andre kohort-effekter.

Jeg kan også her dele opp analysen ved å kjøre separate regresjoner på de enkelte aldersgruppene for å tydeligere få frem forskjellene. Avkastningen av økte ferdigheter med ett standardavvik for de forskjellige aldersgruppene er på henholdsvis 8,5%, 11,7% og 11,2% økt timelønn. Differansen mellom unge og gamle er altså noe større ved denne spesifiseringen enn i tabell 14. Dette kommer av større forskjeller som fanges opp av koeffisientene

Tabell 14: Avkastning av ferdigheter til forskjellige aldersgrupper

Variabel	Koeffisient
erfaring	0.103*** (0.011)
prime	0.097*** (0.020)
prime*ferdighet	0.017 (0.013)
gammel	0.146*** (0.029)
gammel*ferdighet	0.003 (0.019)
erfaring	0.021*** (0.002)
erfaring ²	-0.036*** (0.005)
kvinne	-0.126*** (0.011)
konstant	5.114*** (0.020)
N	2563
R ²	0.28

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

for erfaring, alder og kjønn. Koeffisienten for *erfaring* for unge og gamle er på henholdsvis 0,016 og 0,024 og kvadratleddet er på henholdsvis -0,077 og -0,030. Koeffisienten for arbeidserfaring for den yngre gruppen er lavere og mindre signifikant enn de andre, men koeffisienten for alder er mye sterkere. Alt annet likt, vil en økning i alder med ett år føre til en økt timelønn med 2,8% Dette tyder på at for de yngre spiller det en større rolle for arbeidsgiver at man har blitt eldre og mer voksen enn den faktiske erfaringen. Koeffisienten til kjønnsforskjellen, ved *kvinne*, er mindre for den yngre aldersgruppen enn for den eldre gruppen.

Tabell 15: Avkastning av ferdigheter til forskjellige aldersgrupper

Variabel	ung	prime	gammel
ferdigheter	0.085*** (0.01)	0.117*** (0.01)	0.112*** (0.02)
erfaring	0.016* (0.01)	0.020*** (0.01)	0.024*** (0.01)
erfaring ²	-0.077* (0.04)	-0.049*** (0.01)	-0.030** (0.01)
kvinne	-0.117*** (0.00)	-0.138*** (0.00)	-0.133*** (0.01)
alder	0.028*** (0.00)	0.006*** (0.00)	-0.002 (0.01)
konstant	4.418*** (0.07)	5.011*** (0.09)	5.221*** (0.34)
N	771	1417	375
R ²	0.30	0.21	0.22

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 20-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

Oppsummering: Jeg finner klare beviser for en positiv sammenheng mellom ferdigheter og inntekt. Ved å benytte et gjennomsnitt av de forskjellige ferdighetsmålene tyder resultatene på at økt ferdighetsscore med ett standardavvik vil øke inntekten med 11,7%. Ved å teste forskjellige spesifiseringer og ta høyde for flere ulike kontrollvariabler synes resultatet å være robust.

7 Oppsummering og konklusjon

Den klassiske Mincer-modellen for humankapital har vært referansepunkt i utdannings- og arbeidsmarkedsøkonomi de siste tiårene. Ved å ta utgangspunkt i denne, har jeg undersøkt avkastningen av ferdigheter i arbeidslivet i Norge i dag. Jeg har også testet utvidelser av modellen for å undersøke om andre spesifiseringer bedre fanger opp individuelle variasjoner i humankapital og inntekt.

Ved å benytte antall år utdanning som et mål på humankapital antar man at alle typer utdanning gir samme avkastning, at kvaliteten på utdanning er lik mellom skoler, og at alle individer får like mye utbytte av skolegang. Jeg har sett nærmere på tidligere studier med utvidelser av Mincer-modellen og andre teorier som forklarer hvilke effekter som påvirker individuelle forskjeller i humankapital.

Data fra PIAAC-undersøkelsen har gitt meg testresultater på individuelle kognitive ferdigheter som jeg har benyttet som direkte ferdighetsmål i min analyse. Ved å benytte dette som forklaringsvariabel, har jeg tatt høyde for at utdanningskvalitet og -utbytte varierer mellom individer og heller fokusert på resultatet av utdanningen. Dette vil kunne gi et tydeligere bilde på hvilke faktorer som påvirker individuell humankapital og inntekt.

I grunnmodellen hvor jeg tar utgangspunkt i Mincers humankapitalmodell, men inkluderer ferdigheter fremfor utdanning, får jeg at økt ferdighetsscore med ett standardavvik vil øke inntekten med 11,7%, når det kontrolleres for arbeidserfaring, kjønn og alder. Inkludering av utdanning i modellen har en tydelig direkte effekt på inntekt utenom ferdighetene. Til et gitt ferdighetsnivå vil ett års ekstra skolegang gi 4,1% høyere inntekt. Dette tyder på klare effekter av uobserverbare ferdigheter fra skolegang, som sosiale ferdigheter og evnen til å lære. Dette kan også tyde på en signaliseringseffekt av utdanning. Resultatet tyder også på en økning i ferdigheter fra arbeidserfaring.

Jeg utvider den klassiske Mincer-modellen ved å ta høyde for at inntekt og avkastning av ferdigheter kan variere mellom kohorter i utvalget. Jeg finner en jevn, økende inntekt over kohortene, noe som tyder på å være en effekt av alder og erfaring. Jeg finner ingen signifikante forskjeller i avkastning av ferdigheter mellom kohorter, selv ikke når jeg kun ser på privat sektor. Dette tyder på at arbeidsmarkedet i Norge har vært stabilt over tid. Dette kan også tyde på sterke institusjoner og fagforeninger som beskytter arbeideres rettigheter.

For å ta høyde for en endogenitet i målet på ferdigheter estimerer jeg også modellen ved en instrumentvariabel-metode. Ved å benytte foreldres utdanning som instrument gir estimeringen et mye høyere estimat på avkastningen av ferdigheter. Ett standardavviks økte ferdigheter vil, med alt annet gitt, gi en økt inntekt med 21,3%. Denne økte effekten kan begrunnes med at ferdighetene inneholder en målefeil som gir en skjevhet ved MKM-

estimering.

Ved å studere separate regresjoner for individer i privat og offentlig sektor finner jeg klare indikatorer på en høyere avkastning av ferdigheter i privat sektor. Ett standardavvik økning i ferdigheter vil gi 4 prosentpoeng høyere lønnsøkning i privat enn i offentlig sektor. Ved å sammenlikne med individer i lavere og høyere aldersgrupper finner jeg også at avkastningen av ferdigheter er høyest for individer mellom 35-54 år.

Estimatene mine tyder på at avkastningen av ferdigheter ikke er lineær. Ved å benytte flere funksjonsformer finner jeg at den marginale avkastningen vil være høyest ved et ferdighetsnivå like under gjennomsnittet.

Ved alle mine estimeringer tyder resultatene på en klar positiv sammenheng mellom ferdigheter og inntekt. Mulige utvidelser av arbeidet kan være å se nærmere på hvilke effekter som påvirker ferdighetene i hvilken grad og dermed hva som er den beste investeringen i ferdigheter.

Referanser

- Bjørkeng, B. (2013): "Ferdigheter i voksenalderen - Resultater fra den internasjonale undersøkelsen om lese- og tallforståelse (PIAAC)." Statistisk sentralbyrå, Rapporter 42.
- Boockmann, B. og Steiner, V. (2006): "Cohort effects and the returns to education in West Germany." *Applied Economics* 38:10, 1135–1152.
- Card, D. (1999): "The Causal effect of education on earnings." *Handbook of Labor Economics* 3, 1802–1863.
- Card, D. og Krueger, A. (1992): "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Quarterly Journal of Economics* 100, 1–40.
- Gravem, D. F. og Lagerstrøm, O. (2013): "Den internasjonale undersøkelsen og lese- og tallforståelse - PIAAC, dokumentasjonsrapport." Statistisk sentralbyrå, Notater 47/2013.
- Green, D. og Riddell, W. C. (2001): "Literacy, Numeracy and Labour Market Outcomes in Canada, Working Paper." *Canadian International Labour Network* 01-05.
- Hægeland, T., Klette, T. J. og Salvanes, K. G. (1999): "Declining Returns to Education in Norway? Comparing Estimates across Cohorts, Sectors and over Time." *Scandinavian Journal of Economics* 101(4), 555–576.
- Haider, S. og Solon, G. (2006): "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96(4), 1308–1320.
- Hanushek, E. A., Schwerdt, G., Wiederhold, S. og Woessmann, L. (2013): "Returns to skills around the world: Evidence from PIAAC." *National Bureau of Economic Research Working Paper* 19762.
- Harmon, C., Oosterbeek, H. og Walker, I. (2003): "The Returns to Education: Microeconomics." *Journal of Economic Surveys* 17(2), 115–156.
- Lemieux, T. (2006): *The "Mincer Equation" Thirty Years After Schooling, Experience, and Earnings*. Springer US.
- Mincer, J. (1974): *Schooling, Experience and Earnings*. Columbia University Press, New York.
- Murnane, R. J., Willett, J. B., Duhalsky, Y. og Hoxby, C. M. (2000): "How important are the cognitive skills of teenagers in predicting subsequent earnings?" *Journal of Policy Analysis and Management* 19(4), 547–568.

- Murnane, R. J., Willett, J. B. og Levy, F. (1995): "The Growing Importance of Cognitive Skills in Wage Determination." *The Review of Economics and Statistics* 77(2), 251–266.
- Spence, M. A. (1973): "Job Market Signaling." *Quarterly Journal of Economics* 87(3), 355–374.
- Varian, H. R. (1992): *Microeconomic analysis*, bind 3. Norton New York.
- Wooldridge, J. M. (2009): *Introductory Econometrics: A modern Approach (4e utgave)*. South-Western, Canada.

A Førstestegslikninger

Førstestegslikninger for instrumentvariabel-estimeringene i kapittel 6.

Variabel	IV-1	IV-2
erfaring	0.030 (0.021)	0.045** (0.018)
erfaring ²	-0.107** (0.042)	-0.087** (0.038)
kvinne	-0.212*** (0.050)	-0.236*** (0.046)
alder	-0.008 (0.007)	-0.028*** (0.007)
paredmed	0.276*** (0.063)	0.166*** (0.059)
paredhi	0.566*** (0.070)	0.298*** (0.066)
utdanning		0.164*** (0.012)
konstant	0.210 (0.326)	-1.656*** (0.326)
N	1400	1399
R ²	0.118	0.241

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel gjennomsnittstestresultater. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

B Grunnmodell med foreldres utdanning inkludert som kontrollvariabel

Variabel	Koeffisient
ferdighet	0.075*** (0.009)
erfaring	0.026*** (0.006)
erfaring ²	-0.047*** (0.012)
utdanning	0.040*** (0.004)
kvinne	-0.151*** (0.014)
alder	0.001 (0.002)
paredmed	0.031* (0.018)
paredhi	0.041** (0.019)
konstant	4.492*** (0.090)
<hr/>	
N	1399
R ²	0.284

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.

C Forskjeller i avkastning mellom privat og offentlig sektor

Grunnmodell inkludert interaksjonsledd mellom dummyvariabel for offentlig sektor og alle kontrollvariabler. Denne danner grunnlaget for Chow-testen i kapittel 6.6.2.

Variabel	Koeffisient
ferdighet	0.134*** (0.009)
erfaring	0.014** (0.006)
erfaring ²	-0.039*** (0.014)
kvinne	-0.127*** (0.021)
alder	0.010*** (0.003)
offentlig	-0.008 (0.175)
offentlig*ferdighet	-0.041*** (0.015)
offentlig*erfaring	0.011 (0.011)
offentlig*erfaring ²	-0.017 (0.024)
offentlig*kvinne	0.006 (0.030)
offentlig*alder	-0.005 (0.004)
konstant	4.984*** (0.116)
N	1398
R ²	0.231

Tabellen presenterer resultater fra MKM-regresjoner, vektet etter PIAACs utvalgsvekter. Avhengig variabel er logaritmisk timelønn. Estimerte robuste standardavvik er oppgitt i parentes. erfaring² er delt på 100. Testresultatene er standardiserte til standardavvik 1. Utvalget består av individer mellom 35-54 år i full jobb i Norge 2013. Datakilde: PIAAC.

Signifikansnivå: *: 10% **: 5% ***: 1%.