

Forskjeller og likheter i lønnsavkastning av formell utdanning og kognitive ferdigheter mellom nordiske land i PIAAC-undersøkelsen

Bacheloroppgave i Samfunnsøkonomi

Mai 2019

Forskjeller og likheter i lønnsavkastning av formell utdanning og kognitive ferdigheter mellom nordiske land i PIAAC-undersøkelsen

SØK2901 Bacheloroppgave i samfunnsøkonomi

Sammendrag

En standard Mincer-lønnsfunksjon med formell utdanning som mål på humankapital blir utvidet med observerte kognitive ferdigheter innen tallforståelse, leseferdigheter og problemløsning i IKT-miljø. Ordinær regresjonsanalyse på dataene fra PIAAC-undersøkelsen i de nordiske landene Danmark, Finland, Norge og Sverige avdekker statistisk signifikant positive lønnsavkastninger av formell utdanning og kognitive ferdigheter, hvor de grunnleggende avkastningene varierer marginalt mellom landene. Analysen avdekker derimot forskjeller i avkastningen av kognitive ferdigheter innad i demografiske grupper, der høyere utdannede i Danmark, Finland og Norge får en høyere avkastning av kognitive ferdigheter enn lavere utdannede, samt at eldre arbeidere i Danmark, Finland og Sverige får en høyere avkastning av kognitive ferdigheter enn yngre arbeidere. Avkastning av utdanning og ferdigheter mellom arbeidssektorer varierer i store trekk ikke mellom landene, men noen forskjeller avdekkes for arbeidssektorene med dårligst avkastning av formell utdanning.

Abstract

A standard Mincer earnings function with formal education as a measure of human capital is expanded with observed cognitive skills in numeracy, literacy and problem solving in technology-rich environments. Ordinary regression analysis of the data from the PIAAC survey in the Nordic countries Denmark, Finland, Norway and Sweden reveals statistically significant positive economic returns to formal education and cognitive skills, where the fundamental returns to education and skills vary marginally between countries. The analysis however reveals differences in the return to cognitive skills within demographic groups, where higher educated respondents in Denmark, Finland and Norway receive a higher return to cognitive skills compared to lower educated respondents. Equivalently, older workers in Denmark, Finland and Sweden receive a higher return to cognitive skills compared to younger workers. The return to education and skills between industries and sectors does not vary widely between countries, but some differences are revealed for industries and sectors with the worst return to formal education.

INNHold

1 Innledning	4
2 Teori	5
2.1 Bakgrunn og tidligere studier	5
2.2 Analytisk rammeverk	7
2.2.1 Modell for multippel lineær regresjon	7
2.2.2 Estimering av β med minste kvadraters metode (OLS)	8
2.2.3 Hypotesetester for β	8
2.2.4 Studentiserte residualer og tilhørende tester av modellantagelser	9
2.2.5 Chow-test	9
2.2.6 Linear mixed effects models (LMM)	9
3 Dataanalyse	10
3.1 Om PIAAC-undersøkelsen	10
3.2 Deskriptiv statistikk	11
4 Økonometrisk modell	14
4.1 Konseptuelt rammeverk	14
4.2 Empirisk modell	14
4.3 Utvidelser av empirisk modell	15
5 Regresjonsanalyse	15
5.1 Grunnnavkastning av utdanning og ferdigheter	16
5.2 Avkastning av utdanningsnivå og interaksjon med ferdigheter	17
5.3 Aldersvarierende avkastning av ferdigheter og utdanning	18
5.4 Avkastning av flere typer kognitive ferdigheter	21
5.5 Demografisk varierende avkastningen av ferdigheter og utdanning	23
5.6 Arbeidssektor-varierende avkastningen av ferdigheter og utdanning	24
5.7 Diagnostisk analyse for tilpasning av likning (17)	26
6 Oppsummering	27
Litteraturliste	29
A Appendiks	31

INNLEDNING

Utvikling av arbeidsstyrkens kunnskaper og ferdigheter blir i større og større grad ansett som et sentralt element i moderne økonomisk politikk. Moderne syn på langsiktig økonomisk vekst og stabilitet legger betydelig vekt på det enkelte menneskets iboende kompetanse og produktivitet, dets «humankapital». Disse evnene kan videre styrkes gjennom en økt investering i f.eks. utdanning, opplæring i jobb eller helse, og arbeidsstyrkens humankapital har lenge blitt ansett som en sentral faktor for fordelingen av økonomiske goder.

Helt tilbake til 1700-tallet postulerer Adam Smith at arbeidets lønning reflekterer hvor lett et arbeid er og hvor vanskelig det er å lære det, og om utdanning skriver Smith at lønnen må speile ferdighetsinvesteringen utdanningen utgjør: «The work which he learns to perform, it must be expected, over and above the usual wages of common labour, will replace to him the whole expense of his education, with at least the ordinary profits of an equally valuable capital.» (Smith, 2013)

Ulike myndigheters ønske om å utarbeide komparativ informasjon vedrørende den voksende befolkningens grunnleggende ferdigheter, ferdighetens fordeling i befolkningen, avkastningen av ferdigheter og befolkningens disposisjon for ferdighetervervelse, som kan bistå myndigheter med å evaluere politikk og utforme mer effektive intervensjoner i utdanningssystemet og arbeidsmarkedet, motiverte OECD for å etablere en strategi for å vurdere den voksnes befolkningens kompetanse på tvers av land – Den internasjonale undersøkelsen om lese- og tallforståelse (PIAAC) (Schleicher, 2008). PIAAC-undersøkelsen søker å beskrive og sammenlikne grunnleggende ferdigheter av typen leseferdigheter, tallforståelse og problemløsning i et IKT-miljø mellom voksne individer fra over 40 land. I tillegg er respondentenes respektive ferdighetsscorer supplert med en serie utfyllende bakgrunnsspørsmål om blant annet utdanningsnivå, lønn, arbeidserfaring og demografiske forhold.

Med grunnlag i resultatene fra PIAAC-undersøkelsen ønsker denne rapporten å studere og kartlegge forskjellene og likhetene i lønnsavkastning av formell utdanning og kognitive ferdigheter mellom de nordiske landene Danmark, Finland, Norge og Sverige. Denne analysen blir gjennomført ved å formulere en humankapital-lønnsfunksjon inspirert av Mincer (1974), og estimere denne ved hjelp av klassisk multippel lineær regresjon. Dernest blir tilpasningen studert fra et statistisk standpunkt og avkastningen av ferdigheter og utdanning på lønn blir analysert.

Med tanke på de store likhetene mellom de nordiske landene forventes det at resultatene på tvers av disse kommer til å variere marginalt. Dette kan bl.a. begrunnes med at landenes politiske og økonomiske system har mange fellestrekk. Videre forventes en positiv lønnspremie både for økte kognitive ferdigheter og for økt formell utdanning, samtidig som lønnspremien av kognitive ferdigheter varierer mellom høyere og lavere utdannede respondenter. Videre vil rapporten undersøke hvordan avkastningen varierer på tvers av aldersgrupper, demografier og sektorer i de ulike landene. Her forventes små variasjoner, med grunnlag i de store likhetene mellom landene i undersøkelsen. Særlig for sektorer som i utstrakt grad er internasjonaliserte forventes det at variasjonene er tilnærmet ikke-eksisterende.

Resultatene viser at den statistisk signifikant positive grunnavkastningen av formell utdanning og kognitive ferdigheter varierer lite mellom de nordiske landene, men at landene fremviser noen forskjeller når avkastningen av utdanning og ferdigheter studeres mellom ulike demografiske grupper.

Spesielt Norge betaler en stor lønnspremie for høyere utdanning, mens Sverige i mindre grad premierer høyere utdanning fremfor lavere utdanning; i det videre får høyere utdannede i Danmark, Finland og Norge en høyere avkastning av kognitive ferdigheter enn lavere utdannede, samt at eldre arbeidere i Danmark, Finland og Sverige får en høyere avkastning av kognitive ferdigheter enn yngre arbeidere. Avkastning av utdanning og ferdigheter mellom arbeidssektorer varierer i store trekk ikke mellom landene, men noen forskjeller viser seg for arbeidssektorene med dårligst avkastning av formell utdanning.

TEORI

2.1 Bakgrunn og tidligere studier

Utdanning er en tidkrevende, og i mange land kostnadskrevende, investering som ikke garanterer høyere lønn. Sammenhengen mellom utdanning og lønn, og mellom kognitive ferdigheter og lønn, har derimot blitt veldokumentert i litteraturen, hvor studier generelt har vist at det er en positiv lønnspremie for arbeidere med en høyere grad av utdanning og for arbeidere med høyere observerte kognitive ferdigheter (Gleeson, 2005; Green et al., 2001).

Carnevale et al. (2011) konkluderer med at høyere utdanning er nøkkelen til høyere lønninger og større økonomisk handlingsrom, og viser at medianlønnen til arbeidstakere i det amerikanske arbeidsmarkedet øker for alle økende universitetsgrader, både innad i yrkesgrupper og mellom yrkesgrupper som tradisjonelt sett krever ulik grad av formell utdanning. Tilsvarende viser Robinson and Sexton (1994) på amerikansk folketellingsdata at sannsynligheten for å lykkes som selvstendig næringsdrivende øker for økt utdanning, at denne effekten er større enn effekten av økt erfaring og at lønnen til selvstendig næringsdrivende økte for hvert ekstra år med utdanning.

Hvor i utdanningsløpet avkastningen av utdanning er størst, har vært grunnlag for diskusjon i litteraturen. Der en i perioden 1960–1990 observerte en konkav lønnsavkastning i utdanning, antyder data fra 1990-tallet og tidlig 2000 at avkastning på grunnutdanning nå er lavere enn avkastningen på høyere utdanning, slik at lønnspremien på et ekstra år med utdanning øker for utdanningsnivået (Colclough et al., 2010).

Tilsvarende har en økning i antallet som påbegynner høyere utdanning blitt koblet opp mot en observert økning i lønnsvariasjoner mellom arbeider med lik grad av høyere utdanning. Walker and Zhu (2008) finner ikke en signifikant endring i lønnspremien av utdanning i perioden 1994–2006, men gjennom kvantilregresjon observerer de at lønnspremien av utdanning øker for de i den øvre kvantilen av lønnsfordelingen, mens lønnspremien av utdanning avtar for de i den nedre kvantilen av lønnsfordelingen. Lindley and McIntosh (2015) observerer at den økende lønnsforskjellen blant høyt utdannede briter er relatert til en økt varians i ferdighetene til studentene, der flere studenter har lavere ferdigheter enn tidligere, samt at studentene går over til et større sett med jobber.

For kognitive ferdigheter er lønnspremien tilstedeværende for selv de med lavest grunnleggende ferdigheter i lesing og tallforståelse, mens lønnspremien til de med høyere ferdighetsnivåer varierer fra land til land etter relativt tilbud og etterspørsel etter slike ferdigheter (Vignoles, 2016). Bynner (2002) fant for The 1970 British Cohort Study (BCS70) at lav tallforståelse og lave leseferdigheter reduserte britiske arbeideres ansettelseskapasitet uavhengig av andre kvalifikasjoner som utdanning,

og Shomos (2010) fant en positiv, statistisk signifikant effekt av å bedre lese- og tallforståelsen til australske arbeidere i The Adult Literacy and Lifeskills Survey (ALLS). Heri viste Shomos at å øke et individs ferdigheter fra laveste nivå til det som ble ansett å være minimum for å delta i en kunnskapsbasert økonomi, økte sannsynligheten for at individet deltok i arbeidsstyrken med 15 % for kvinner og 5 % for menn, og timeslønnen økte med omtrent 25 og 30 % for henholdsvis kvinner og menn.

McIntosh and Vignoles (2001) brukte 1991 National Child Development Study (NCDS) og 1994 International Adult Literacy Survey (IALS) til å studere individer på den nedre delen av ferdighetsfordelingen. Gjennom regresjonsanalyse fant de substansielle bevis for en lønnspremie av økte kognitive ferdigheter innen lese- og tallforståelse. Resultatene fra IALS-dataene antyder også at tallforståelse er viktigere enn leseferdigheter for lønn, hvor de estimerte lønnspremien av å øke et individs tallforståelse opp til ferdighetsnivået assosiert med et minimum for å delta i et moderne arbeidsplass til 16–21 % dersom en ikke kontrollerte for andre effekter, men at leseferdigheter er mest utslagsgivende for sannsynligheten av å være i jobb.

Green and Riddell (2003) undersøkte den kanadiske komponenten av 1994 International Adult Literacy Survey (IALS) for å se hvordan observerte leseferdigheter interakterte med uobserverte kognitive ferdigheter. De implementerte en kvantilregresjon med og uten leseferdigheter som kontrollvariabel, og fant at leseferdigheter har en stor effekt på lønn, samt at inkluderingen av leseferdigheter reduserer medianlønnseffekten av utdanning med omtrent 30 %. De fant i det videre at leseferdigheteffekten ikke varierte over kvantilene på lønnsfordelingen, og introduisering av leseferdigheter hadde lite effekt på avkastningen av erfaring.

Det er i det videre velobservert at avkastningen av kognitive ferdigheter varierer for ulike aldersgrupper. Murnane et al. (1995) observerer at grunnleggende kognitive ferdigheter innen matematikk var viktigere prediktorer for lønn seks år etter skoleslutt enn de var to år etter skoleslutt, mens Bishop (1992) ikke finner, når det er kontrollert for utdanning, at akademiske ferdigheter har noen effekt på lønnsraten de første åtte årene med arbeid etter avsluttet videregående utdanning i NLS-Youth-studiet. Bishop konkluderer at markedet har incentiver for å statistisk diskriminere unge arbeidere på basis av attester og enkelt observerbare variabler korrelert med ferdigheter og produktivitet, jf. utdanning. Kompetanse ikke demonstrert gjennom attester går lenge uten å bli kompensert, ettersom markedet bruker lang tid på å oppdage disse ferdighetene. Altonji and Pierret (1998) postulerer tilsvarende at samtidig som markedet lærer om det aktuelle individet over tid, slik at avkastningen av individets ferdigheter øker, burde signaleffekten utdanning utgjør avta.

Samtidig som det er en åpenbar korrelasjon mellom skolegang og resultater på kognitive tester, tyder noen studier på at respondenter med høye kognitive ferdigheter opplever en høyere avkastning høyere utdanning enn respondenter med lave kognitive ferdigheter. På data for hvite menn fra National Longitudinal Survey of Youth (NLSY) fra årene 1979–1987 finner Blackburn and Neumark (1993) at lønnsavkastningen av en universitetsgrad er større for menn som gikk ut fra videregående med sterke kognitive ferdigheter enn for menn som gikk ut med svakere ferdigheter. Murnane et al. (2000) ønsket å se hvorvidt den samme interaksjonen viste seg for andre datasett, hhv. for to longitudinelle studier på amerikansk ungdom, the National Longitudinal Survey of the High School Class of 1972 (NLS72) og High School and Beyond (HS&B). For menn i NLS72 er interaksjonen

positiv og statistisk signifikant: Estimert lønnspremie for menn som begynte på universitetet med en matematikk-score ett standardavvik over gjennomsnittet er 31 %, mens estimert lønnspremie for menn som begynte på universitetet med en matematikk-score ett standardavvik under gjennomsnittet er 21 %. For kvinner i NLS72 og for menn og kvinner i HS&B er dog ikke interaksjonen statistisk signifikant.

2.2 Analytisk rammeverk

2.2.1 Modell for multipel lineær regresjon

La $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^T$ være en kontinuerlig n -dimensjonal stokastisk variabel bestående av enkeltobservasjoner Y_i , for $i = 1, \dots, n$. La \mathbf{X} være en $(n \times p)$ -dimensjonal designmatrise hvor hver rad tilsvarer de observerte verdiene av $k = p - 1$ kontrollvariabler for hver respektive observasjon,

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix},$$

og anta at \mathbf{X} har full rank, $\text{Rank}(\mathbf{X}) = p$. (For det studerte datasettet i denne oppgaven er òg \mathbf{x}_i en stokastiske variabel. Følgelig er observasjonsparet (Y_i, \mathbf{x}_i) en realisering av det stokastiske paret (Y, \mathbf{x}) , og alle modellantakelsene som følger under må gjøres betinget på designmatrisen \mathbf{X} . For notasjonell simplisitet undertrykkes den betingede avhengigheten av \mathbf{X} for alle stokastiske variabler i senere underkapitler.)

For hvert observasjonspaar (Y_i, \mathbf{x}_i) , for $i = 1, \dots, n$, ønsker vi å modellere sammenhengen mellom avhengig variabel og kontrollvariabelvektor gjennom en systematisk komponent $f(\mathbf{x}_i) = \mathbf{x}_i^T \boldsymbol{\beta}$, som er en lineær kombinasjon av kontrollvariablene, slik at $E(Y_i | \mathbf{X}) = f(\mathbf{x}_i)$. Heri er $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^T$ en p -dimensjonal vektor med ukjente regresjonsparametre. Ettersom Y_i er en stokastisk variabel, antar vi at vi har additive feil, slik at

$$Y_i = f(\mathbf{x}_i) + \varepsilon_i, \quad \text{for } i = 1, \dots, n. \quad (1)$$

La $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$ være en n -dimensjonal vektor med homoskedastiske stokastiske feil, slik at feilene har forventningsverdi null og konstant varians σ^2 . Anta i det videre at feilene er ukorrelerte og følger en normalfordeling: $\boldsymbol{\varepsilon} | \mathbf{X} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$, som impliserer at $\boldsymbol{\varepsilon}$ og \mathbf{X} er stokastisk uavhengige. Dersom vi antar at vi har parvis uavhengige observasjonspaar (Y_i, \mathbf{x}_i) , for $i = 1, \dots, n$, er den klassiske multiple lineære regresjonsmodellen følgelig gitt ved

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2)$$

hvor $\mathbf{Y} | \mathbf{X} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$.

2.2.2 Estimering av β med minste kvadraters metode (OLS)

La estimatoren til \mathbf{Y} være $\hat{\mathbf{Y}} = \hat{f}(\mathbf{X}) = \mathbf{X}\hat{\beta}$. Vi ønsker å finne regresjonsparameterene $\hat{\beta}$ som minimerer avvikene $(Y_i - \hat{Y}_i)^2$, for $i = 1, \dots, n$. Definer vektoren av residualer $\hat{\epsilon} := \mathbf{Y} - \mathbf{X}\hat{\beta}$, slik at summen av de kvadrerte residualene er $\hat{\epsilon}^T \hat{\epsilon} = (\mathbf{Y} - \mathbf{X}\hat{\beta})^T (\mathbf{Y} - \mathbf{X}\hat{\beta})$. For å finne $\hat{\beta}$ som minimerer summen av de kvadrerte residualene, deriveres $\hat{\epsilon}^T \hat{\epsilon}$ med hensyn på $\hat{\beta}$ og settes lik null. Dette gir

$$\frac{\partial \hat{\epsilon}^T \hat{\epsilon}}{\partial \hat{\beta}} = -2\mathbf{X}^T \mathbf{Y} + 2\mathbf{X}^T \mathbf{X} \hat{\beta} = 0, \quad (3)$$

hvilket løst for $\hat{\beta}$ gir estimatoren

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}. \quad (4)$$

For å vise at (4) er et minimum, deriveres uttrykket i (3) med hensyn på $\hat{\beta}^T$, slik at den dobbeltderiverte av summen av kvadrerte residualer med hensyn på regresjonsparameteren $\hat{\beta}$ blir funnet til å være

$$\frac{\partial^2 \hat{\epsilon}^T \hat{\epsilon}}{\partial \hat{\beta} \partial \hat{\beta}^T} = 2\mathbf{X}^T \mathbf{X}. \quad (5)$$

Dersom \mathbf{X} har full rank, er $2\mathbf{X}^T \mathbf{X}$ en positivt definit matrise, og følgelig er $\hat{\beta}$ et minimum (Fahrmeir et al., 2013).

Fra (4) er det trivielt å vise at $E(\hat{\beta}) = \beta$ og at $\text{Var}(\hat{\beta}) = \sigma^2(\mathbf{X}^T \mathbf{X})^{-1}$, slik at $\hat{\beta} \sim N(\beta, \sigma^2(\mathbf{X}^T \mathbf{X})^{-1})$. Under noen svake regularitetsvilkår kan det vises at (Agresti, 2015) når $n \rightarrow \infty$, eksisterer $\hat{\beta}$ og er konsistent, samt

$$\hat{\beta} \approx N_p(\beta, \hat{\sigma}^2(\mathbf{X}^T \mathbf{X})^{-1}), \quad (6)$$

der $\hat{\sigma}^2$ er REML-estimatoren for σ^2 gitt ved

$$\hat{\sigma}^2 = \frac{1}{n-1} (\mathbf{Y} - \mathbf{X}\hat{\beta})^T (\mathbf{Y} - \mathbf{X}\hat{\beta}). \quad (7)$$

2.2.3 Hypotesetester for β

Gitt at $\hat{\beta}_j$, $j = 1, \dots, p$, og $\hat{\sigma}^2$ er uavhengige (som kan vises), følger det at

$$T_j = \frac{\hat{\beta}_j - \beta_j}{\sqrt{c_{jj} \hat{\sigma}^2}} \sim t_{n-p}, \quad (8)$$

der c_{jj} er det j -te diagonale elementet i $(\mathbf{X}^T \mathbf{X})^{-1}$. Signifikansen til β_j kan testes gjennom hypotesetesten $H_0 : \beta_j = 0$ mot $H_1 : \beta_j \neq 0$ med test-statistikk gitt i (8) og tilhørende p -verdi $2 \cdot P(T_j > |t_j|)$.

2.2.4 Studentiserte residualer og tilhørende tester av modellantagelser

La $\mathbf{H} := \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$. Det følger at $\hat{\boldsymbol{\varepsilon}} = (\mathbf{I} - \mathbf{H})\mathbf{Y}$, slik at

$$\hat{\boldsymbol{\varepsilon}} \sim N(\mathbf{0}, \sigma^2(\mathbf{I} - \mathbf{H})). \quad (9)$$

Residualene er følgelig, i motsetning til feilene, heteroskedastiske og korrelerte. Residualene brukes i praksis til å teste modellantagelsene om uavhengige og homoskedastiske feil, samt hvorvidt feilene er uavhengige av kovariatene. Følgelig, for å kunne brukes til å vurdere gyldigheten til antagelsen om homoskedastisitet, må residualene først standardiseres. Fahrmeir et al. (2013) viser at det kan konstrueres studentiserte residualer

$$r_i^* = \frac{\hat{\varepsilon}_i}{\hat{\sigma}_{(i)} \sqrt{1 - h_{ii}}} \sim t_{n-p-1} \quad (10)$$

slik at r_i^* kan brukes til å teste ε s homoskedastisitet, der h_{ii} er det i -te diagonalelementet til \mathbf{H} , og $\hat{\sigma}_{(i)}$ er den estimerte variansen i en modell hvor observasjon i er utelatt. Studentiserte residualer er fremdeles korrelerte, men i de fleste tilfeller er korrelasjonen mellom residualene neglisjerbar (Fahrmeir et al., 2013).

De studentiserte residualene kan blant annet bli brukt i følgende tester av modellantagelsene:

- ★ Et plot av (r_i^*, \hat{Y}_i) signaliserer hvorvidt det er en avhengighet mellom residualene og de predikerte verdiene, hvilket i så fall antyder at en må revurdere den estimerte regresjonsmodellen. En ikke-konstant varians kan tyde på at en transformasjon av \mathbf{Y} er nødvendig.
- ★ En trend i et plot av r_i^* mot en kovariat eller en funksjon av kovariater kan avdekke hvorvidt en transformasjon av enkelte kovariater i regresjonen er nødvendig.
- ★ Et QQ-plot av residualene kan brukes for å vurdere antagelsen om normalfordelte feil.

2.2.5 Chow-test

En Chow-test er en statistisk test for hvorvidt koeffisientene i to lineære regresjoner på to ulike datasett er like. La SSE_1 , SSE_2 og SSE_p være summen av de kvadrerte residualene for henholdsvis regresjon på første datasett, regresjon på andre datasett og regresjon på det aggregerte datasettet. Da er Chow-statistikken

$$C = \frac{(SSE_p - (SSE_1 + SSE_2))/k}{(SSE_1 + SSE_2)/(n - 2k)} \sim F_{k, n-2k}, \quad (11)$$

hvor det totalt er n observasjoner i det aggregerte datasettet og k regresjonsparametere i hver modell. p -verdien for Chow-testen er $P(C > F_{k, n-2k})$ (Chow, 1960).

2.2.6 Linear mixed effects models (LMM)

Anta at observasjonsparene (Y_i, \mathbf{x}_i) kommer fra m ulike klustere, og la observasjonene fra kluster i være $\mathbf{Y}_i = (Y_{1i}, \dots, Y_{in_i})^T$. Dersom en ønsker å estimere en regresjonslikning på det aggregerte

datasettet, kan en for å bevare antakelsen om uavhengige observasjonspaar legge til $m - 1$ dummyvariabler for klustrene; dette medfører derimot at en må estimere $m - 1$ ekstra parametere en ikke er veldig interessert i. En kan istedenfor innføre en intrakluster-korrelasjon τ_0^2 , slik kovariansen mellom observasjoner fra like og ulike klustere blir

$$\text{Cov}(Y_{ij}, Y_{kl}) = \begin{cases} \tau_0^2 + \sigma^2 = \text{Var}(Y_{ij}), & \text{for } i = k, j = l, \\ \tau_0^2, & \text{for } i = k, j \neq l, \\ 0, & \text{for } i \neq k, j \neq l. \end{cases} \quad (12)$$

For kluster i følger det da at $\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{U}_i\gamma_i + \boldsymbol{\varepsilon}_i$, med designmatrise $\mathbf{X}_i : n_i \times p$, designvektor for stokastiske effekter $\mathbf{U}_i = \mathbf{1}$ av lengde n_i , samt $\gamma_i \sim N(0, \tau_0^2)$ og $\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \sigma^2\mathbf{I})$.

La i det videre $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_m)^T$, $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_m)^T$, $\mathbf{U} = \text{diag}(\mathbf{U}_1, \dots, \mathbf{U}_m)$, $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_m)^T$ og $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_m)^T$. En kan dermed lage en global modell for alle klusterene gitt ved

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}. \quad (13)$$

Det kan da vises at

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y}, \quad (14)$$

hvor $\mathbf{V} = \text{Cov}(\boldsymbol{\varepsilon}) + \text{Cov}(\boldsymbol{\gamma}\mathbf{U})$ må estimeres, se Fahrmeir et al. (2013). Grunnet problemer med å estimere frihetsgrader i t -fordelingen for LMM-modeller er ikke p -verdiene til regresjonskoeffisientene i LMM-tilpasninger rapportert i denne oppgaven, se Bates et al. (2014).

DATAANALYSE

3.1 Om PIAAC-undersøkelsen

PIAAC-undersøkelsen er iverksatt av Organisasjon for økonomisk samarbeid og utvikling (OECD) for å kartlegge ferdighetsnivået til voksenbefolkningen, og undersøkelsen er gjennomført i over 40 land. Respondentene ble trukket blant bosatte personer i alderen 16–65 år i hvert land. Personer bosatt i kollektive enheter (militærleire, helseinstitusjoner og fengsler) er ikke omfattet av undersøkelsen. Datagrunnlaget ble hentet inn gjennom en kombinasjon av personlig intervjuer, for å kartlegge bakgrunnsinformasjon, og selvutfylling, for å vurdere respondentenes ferdighetsnivå.

For å kartlegge voksenbefolkningens ferdigheter gjennomførte respondentene et sett med tester som besto av oppgaver fra de ulike ferdighetsområdene. Testene i PIAAC-undersøkelsen er basert på Item response theory-paradigmet, hvor respondentens sannsynlighet for å klare oppgaver av ulik vanskelighetsgrad beregnes slik at testene blir mer skreddersydd individets ferdigheter. Der tidligere studier som IALS og ALL benyttet sannsynlighetsverdi 0.80 for at en respondent skulle klare en oppgave på det ferdighetsnivået hen tilhørte, benyttet PIAAC sannsynlighetsverdi 0.67. Denne endringen vanskeliggjør en direkte sammenlikning av ferdighetsscorene mellom disse undersøkelsene.

PIAAC-undersøkelsen kartlegger befolkningens ferdigheter innenfor tre områder: leseferdighet («literacy»), tallforståelse («numeracy») og problemløsning i IKT-miljø («problem solving in technology-

rich environments»):

- ★ Lesferdigheter er definert som evnen til å forstå, vurdere og benytte skrevet tekst for å delta i et moderne samfunn. Lesferdighet-testene søker å vurdere et spektrum fra evnen til å tyde skrevne ord og setninger til evnen til analysere og fortolke kompliserte tekster.
- ★ Tallforståelse er definert som evnen til å tilegne seg, bruke, tolke og formidle matematisk informasjon, fremstilt som tekst og grafikk, og bruke dette til å håndtere og mestre hverdagslige situasjoner.
- ★ Problemløsning i IKT-miljø er definert som evnen til å bruke digital teknologi, kommunikasjonsverktøy, datamaskiner og datanettverk for å erverve informasjon, kommunisere, og løse og utføre praktiske oppgaver tilknyttet privatliv, arbeid og deltakelse i sosiale sammenhenger.

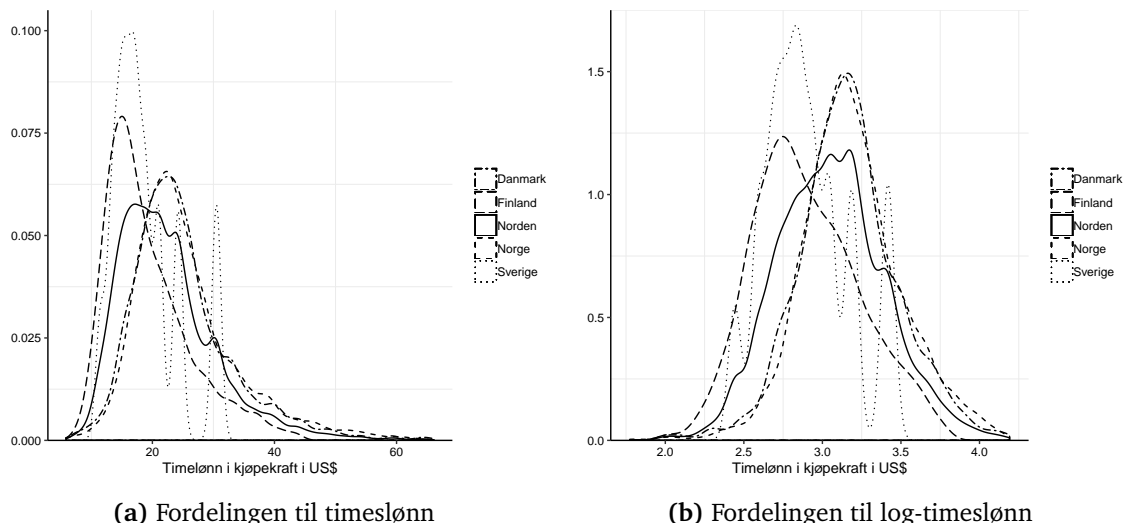
3.2 Deskriptiv statistikk

Ettersom denne rapporten blant annet søker å studere lønnsavkastningen av formell utdanning, er studiet i denne rapporten avgrenset til respondenter over 25 år. For å enklere kunne vurdere avkastningen av ferdigheter, er de tre ferdighetsscorene standardisert til standardavvik 1 innad i hvert land. Deskriptiv statistikk for datasettet, når ugyldige observasjoner er fjernet, er presentert i tabell 1. Brutto timeslønn er oppgitt i kjøpekraft i amerikanske dollar for å bedre kunne studere lønnsforskjeller mellom land. Utdanning er gitt i antall år formell utdanning, og erfaring er gitt i antall år betalt arbeidserfaring. Kjønn er kodet som 0 for menn og 1 for kvinner.

Tabell 1: Deskriptiv statistikk

	<i>(Norden)</i>	<i>(Danmark)</i>	<i>(Finland)</i>	<i>(Norge)</i>	<i>(Sverige)</i>
Brutto timelønn	22.87 (8.14)	25.46 (8.24)	19.60 (6.96)	25.99 (8.58)	19.09 (5.32)
Tallforståelse	0.28 (0.93)	0.31 (0.82)	0.28 (0.88)	0.26 (0.83)	0.26 (0.80)
Lesferdigheter	0.25 (0.83)	0.29 (0.80)	0.24 (0.89)	0.21 (0.86)	0.26 (0.79)
Problemløsning	0.044 (0.98)	0.05 (0.98)	0.06 (0.98)	0.02 (0.99)	0.04 (0.98)
Utdanning	13.98 (2.55)	13.82 (2.49)	13.75 (2.79)	15.09 (2.21)	13.19 (2.30)
Erfaring	22.33 (11.63)	25.56 (11.78)	19.44 (11.04)	21.12 (10.75)	22.03 (11.84)
Kjønn	0.51	0.52	0.52	0.49	0.50
Observasjoner	11325	3571	2597	2734	2423

Note: Deskriptiv statistikk med gjennomsnitt (og tilhørende standardavvik i parentes under) for sentrale variabler for hvert land. De kognitive ferdighetene tallforståelse, lesferdigheter og problemløsning er standardisert til standardavvik 1 innad i hvert land. Respondenter under 25 år er utelatt. Brutto timelønn er oppgitt i kjøpekraft i US\$. For kjønnsindikatoren er kvinner kodet som 1 og menn kodet som 0. Formell utdanning og betalt arbeidserfaring er gitt i antall år. *Data fra:* PIAAC.



Figur 1: Lønnsfordelingen for nordiske landene og for norden som helhet. *Note:* Timeslønn er målt i kjøpekraft i US\$. (a) viser fordelingen av timeslønn og (b) viser fordelingen av log-transformert timeslønn. Bare respondenter eldre enn 25 år er inkludert. *Data fra:* PIAAC.

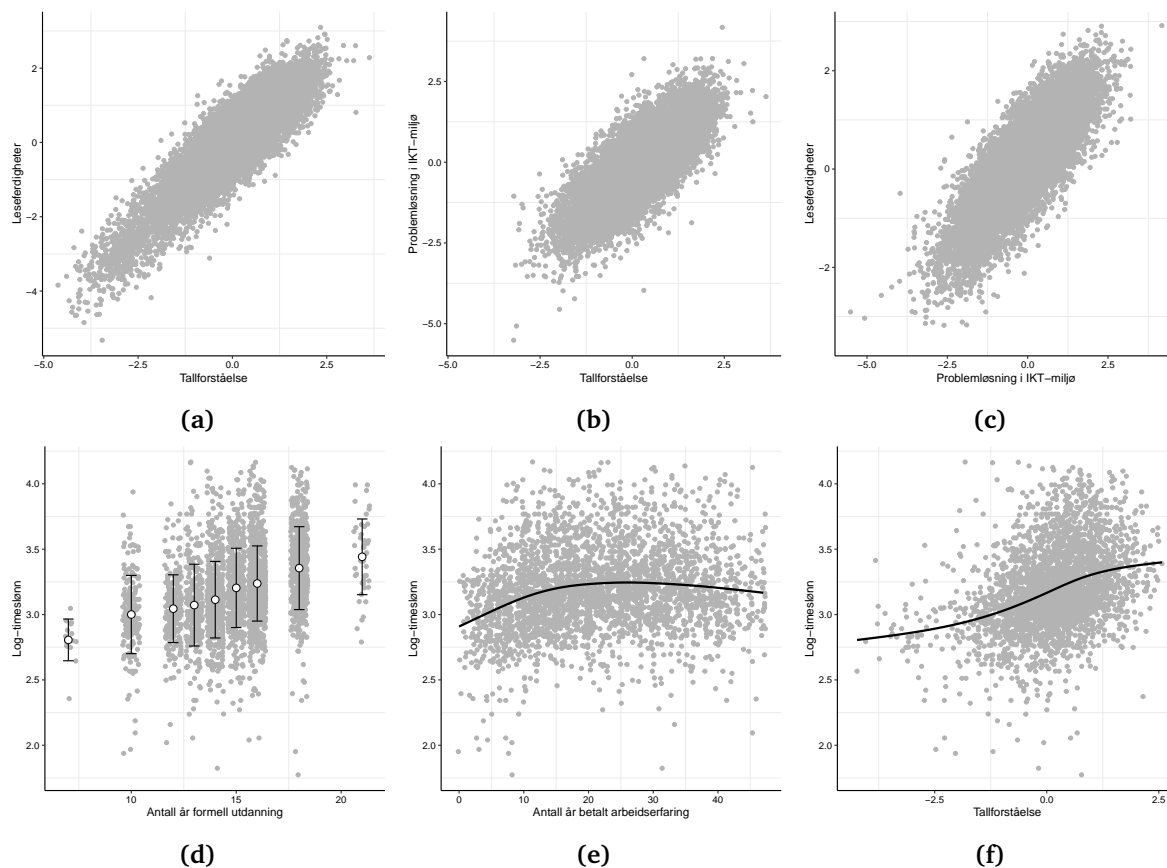
I figur 1 er fordelingen av timeslønn målt i kjøpekraft i US\$ for de nordiske landene og for norden som helhet presentert. Vi observerer at fordelingen av timeslønnen er skjev, mens fordelingen av log-transformert timeslønn mer likner en normalfordeling. For å støtte antakelsen om normalfordelte feil vil vi i den senere regresjonsanalysen følgelig benytte log-timeslønn som avhengig variabel. Dermed vil regresjonskoeffisienten til standardisert testscore tilnærmet kunne tolkes som prosentvis lønnsøkning av et standardavviks økning i test-score. Tilsvarende vil regresjonskoeffisienten til formell utdanning kunne tolkes som prosentvis lønnsøkning av et år ekstra formell utdanning.

I figur 2 er korrelasjonen mellom log-timeslønn og test-score i tallforståelse, antall år formell utdanning og antall år betalt arbeidserfaring, samt korrelasjon mellom test-score i tallforståelse og leseferdigheter, test-score i tallforståelse og problemløsning i IKT-miljø og test-score i leseferdigheter og problemløsning i IKT-miljø, for Norge presentert. Liknende korrelasjoner fremgår for de tre andre nordiske landene.

Vi observerer at det fremgår en kvadratisk sammenheng mellom log-timeslønn og antall år betalt arbeidserfaring. Vi ser i det videre at det fremgår en sammenheng mellom antall år formell utdanning og log-timeslønn, samt mellom test-score i tallforståelse og log-timeslønn, som likner et polynom av tredje grad. I disse to tilfellene virker dog en lineær sammenheng som en adekvat approksimasjon, og antakelsen om en lineær sammenheng mellom formell utdanning og log-timeslønn, og mellom kognitive ferdigheter og log-timeslønn, er sedvanen i litteraturen.

Vi ser i figur 2a–2c at ferdighet-scorene er svært korrelerte: For Norden som helhet er korrelasjonen mellom tallforståelse og leseferdigheter 0.84, korrelasjonen mellom tallforståelse og problemløsning i IKT-miljø er 0.77, og korrelasjonen mellom leseferdigheter og problemløsning i IKT-miljø er 0.81. Korrelasjonen mellom ferdighet-scorene er mindre enn en observerte i f.eks. tidligere IALS-studier, men den høye korrelasjonen vanskeliggjør dog en simultan estimering av avkastningen til ulike ferdighet-scoringer.

Vi vet fra litteraturen at desto mindre den lineære avhengigheten mellom kontrollvariabelen x_j er



Figur 2: Korrelasjon mellom log-timeslønn og kovariater, og mellom ulike ferdighet-scorer, i Norge. *Note:* (a) viser parvisplott av kognitive ferdigheter innen tallforståelse og leseferdigheter, (b) viser parvisplott av kognitive ferdigheter innen tallforståelse og problemløsning i IKT-miljø, (c) viser parvisplott av kognitive ferdigheter innen leseferdigheter og problemløsning i IKT-miljø, og (d)–(f) viser parvisplott av log-transformert timeslønn mot hhv. antall år formell utdanning, antall år betalt arbeidserfaring og kognitive ferdigheter innen tallforståelse. Alle ferdighet-scorer er standardisert til standardavvik 1. Bare respondenter eldre enn 25 år er inkludert. *Data fra:* PIAAC.

med andre kontrollvariabler, desto mindre er variansen til den tilhørende regresjonsparameteren $\hat{\beta}_j$. For svært høye korrelasjoner blir estimatene svært upresise og $\text{Var}(\hat{\beta}_j) \rightarrow \infty$ (Fahrmeir et al., 2013). Grunnet problemer med multikollinearitet vil vi følgelig i stor grad utelukkende benytte kognitive ferdigheter innen tallforståelse som et mål på et individs ferdigheter, da innledende analyser har vist at tallforståelse er den sterkeste predikatoren på lønnsdannelse av ferdighet-scorene.

 ØKONOMETRISK MODELL

4.1 Konseptuelt rammeverk

Standard humankapital-tilnærming for bestemmelse av individuelle lønninger, y_i for individ i , antar

$$y_i = \beta_0 + \beta_1 H_i + \varepsilon_i, \quad (15)$$

for humankapital H_i og additiv stokastisk feil ε_i . Mincer (1974) estimerer humankapitalen gjennom antall år formell utdanning (S_i), slik at $H_i = S_i$. Det er derimot rimelig å forvente at relevant humankapital for lønnsbestemmelse avhenger av flere variabler, som kvaliteten på utdanningen, individuelle ferdigheter og demografiske effekter. Det er derfor ikke rimelig å forvente at ε_i er stokastisk uavhengig av $H_i = S_i$. En eventuell tilnærming er å estimere humankapitalen med en testscore for kognitive ferdigheter, C_i . Kombinert med en lønnseffekt av utdanning gir dette

$$y_i = \beta_0 + \beta_1 C_i + \beta_2 S_i + \varepsilon_i. \quad (16)$$

Det er dog sannsynlig at testscoren er et ufullstendig mål på humankapitalen, slik at $C = H + \nu$, for stokastisk feil ν . Hvis S_i og C_i er positivt korrelert og S_i er målt uten feil, viser Hanushek et al. (2015) at C_i er asymptotisk biased mot null og S_i er biased oppover. En samfunnsøkonomisk forklaring på dette, tilsvarende analysene til Altonji and Pierret (1998), er at arbeidsgivere har manglende informasjon om arbeidstakeres ferdigheter, slik at arbeidsgivere velger å statistisk diskriminere arbeidstakerne med grunnlag i deres formelle utdanning.

4.2 Empirisk modell

Inspirert av humankapital-lønnsfunksjonen til Mincer (1974), utvidet med målte kognitive ferdigheter, ønsker vi å studere en funksjon for individuelle lønninger

$$\ln y_i = \beta_0 + \beta_1 C_i + \beta_2 S_i + \beta_3 E_i + \beta_4 E_i^2 + \beta_5 G_i + \varepsilon_i, \quad (17)$$

hvor $\ln y_i$ er logaritmen av timeslønnen til individ i , E_i er antall år med betalt erfaring i arbeidsmarkedet, G_i er en kjønnsindikator lik 1 for kvinner og lik 0 for menn og ε_i er en additiv stokastisk feil. Grunnet den sterke korrelasjonen mellom de tre ferdighetsmålene i PIAAC-undersøkelsen vil vi i de fleste analysene i denne oppgaven sette $C_i = C_i^{(T)}$, hvor $C_i^{(T)}$ er målte kognitive ferdigheter innen tallforståelse (standardisert til standardavvik 1 innad i hvert land), da innledende analyser har vist at tallforståelse er den sterkeste indikatoren på lønnsdannelse av ferdighetsmålene. Dette har òg vært den vanlige fremgangsmåten i litteraturen, se f.eks. Murnane et al. (2000).

Logaritmisk-tranformert timeslønn er en vanlig funksjonsform i litteraturen som gjør det mulig å sammenlikne estimerte lønnsgep på tvers av studier og land. Vi ser òg i figur 1 at en log-tranformasjon er hensiktsmessig for å tilfredsstille OLS-antakelsen om normalfordelte feil.

4.3 Utvidelser av empirisk modell

For å undersøke ulike dimensjoner av lønnsavkastningen av formell utdanning og kognitive ferdigheter, ønsker vi å estimere et sett med likninger som komplementerer likning (17). For å studere interaksjonen mellom utdanningsnivå og avkastningen av kognitive ferdigheter, deles utdanningsnivået inn i de tre indikatorene $S_i^{(1)}$, $S_i^{(2)}$ og $S_i^{(3)}$, for henholdsvis grunnskole (6–9 år), videregående (10–12 år) og høyere utdanning (13–20 år). Dette gir likningen

$$\ln y_i = \beta_0 + \beta_1 C_i + \beta_2 S_i^{(2)} + \beta_3 S_i^{(3)} + \beta_4 C_i \times S_i^{(2)} + \beta_5 C_i \times S_i^{(3)} + \beta_6 X_i + \varepsilon_i, \quad (18)$$

der $A \times B$ denoterer en toveis-interaksjonseffekt mellom A og B , og for notasjonell simplisitet lar vi $X_i := \gamma_1 E_i + \gamma_2 E_i^2 + \gamma_3 G_i$.

For å studere en aldersvarierende avkastning av kognitive ferdigheter og formell utdanning, deles arbeidstakerne opp i tre grupper $A_i^{(1)}$, $A_i^{(2)}$ og $A_i^{(3)}$, for henholdsvis unge arbeidere (25–34 år), modne arbeidere (35–54 år) og avsluttende arbeidere (55–65 år). Dette gir likningene

$$\ln y_i = \beta_0 + \beta_1 C_i + \beta_2 S_i + \beta_3 A_i^{(2)} + \beta_4 A_i^{(3)} + \beta_5 C_i \times A_i^{(2)} + \beta_6 C_i \times A_i^{(3)} + \beta_7 X_i + \varepsilon_i, \quad (19a)$$

$$\ln y_i = \beta_0 + \beta_1 C_i + \beta_2 S_i + \beta_3 A_i^{(2)} + \beta_4 A_i^{(3)} + \beta_5 S_i \times A_i^{(2)} + \beta_6 A_i \times S_i^{(3)} + \beta_7 X_i + \varepsilon_i. \quad (19b)$$

For å studere varierende avkastning av formell utdanning og kognitive ferdigheter mellom demografiske grupper (f.eks. mellom kvinner og menn eller migranter og ikke-migranter), la D_i være en demografi-indikator med nivåer 0 og 1, som gir likningene

$$\ln y_i = \beta_0 + \beta_1 C_i + \beta_2 S_i + \beta_3 D_i + \beta_4 C_i \times D_i + \beta_5 X_i + \varepsilon_i, \quad (20a)$$

$$\ln y_i = \beta_0 + \beta_1 C_i + \beta_2 S_i + \beta_3 D_i + \beta_4 S_i \times D_i + \beta_5 X_i + \varepsilon_i. \quad (20b)$$

For å studere varierende avkastning av formell utdanning og kognitive ferdigheter mellom k ulike arbeidssektorer, la $\xi_i^{(j)}$ være en indikator for arbeidssektor j . Dersom $\xi_i := \gamma_1 \xi_i^{(2)} + \dots + \gamma_{k-1} \xi_i^{(k)}$, blir likningene dermed

$$\ln y_i = \beta_0 + \beta_1 C_i + \beta_2 S_i + \beta_3 \xi_i + \beta_5 C_i \times \xi_i + \beta_5 X_i + \varepsilon_i, \quad (21a)$$

$$\ln y_i = \beta_0 + \beta_1 C_i + \beta_2 S_i + \beta_3 \xi_i + \beta_4 S_i \times \xi_i + \beta_5 X_i + \varepsilon_i. \quad (21b)$$

REGRESJONSANALYSE

For å undersøke hvordan lønnsavkastning av kognitive ferdigheter og formell utdanning varierer på tvers av nordiske land, begynner vi med å estimere likning (17) for å studere den grunnleggende avkastningen av utdanning og ferdigheter. Siden ferdighetsmålet er standardisert til standardavvik 1, vil dens tilhørende regresjonsparameter kunne tolkes som prosentvis lønnsøkning ved en ett standar-

davviks økning i gitt kognitiv ferdighet. Tilsvarende vil regresjonsparameteren til formell utdanning kunne tolkes som prosentvis lønnsøkning ved ett ekstra år utdanning. Derneft komplementeres analysen gjennom å tilpasse likninger som søker å forklare en variasjon i avkastningen av utdanning og ferdigheter mellom ulike aldere, utdanningsnivåer, demografiske grupper og arbeidssektorer.

5.1 Grunnnavkastning av utdanning og ferdigheter

For å kartlegge den grunnleggende lønnsavkastningen av kognitive ferdigheter og formell utdanning estimeres først likning (17) for hvert land med OLS, i tillegg til at (17) estimeres for norden som helhet med LMM med respondentenes nasjonalitet som stokastisk effekt. Basisresultatene indikerer konsekvent at både kognitive ferdigheter innen tallforståelse og formell utdanning er relatert til høyere lønn i samtlige land, se tabell 2 og figur 3. Vi ser at Sverige har lavest avkastning i både utdanning og ferdigheter, mens Norge og Finland har høyest avkastning i henholdsvis tallforståelse og formell utdanning.

Tabell 2: Avkastning av kognitive ferdigheter og formell utdanning i norden

	<i>(Norden)</i>	<i>(Danmark)</i>	<i>(Finland)</i>	<i>(Norge)</i>	<i>(Sverige)</i>
Tallforståelse	0.067 (0.003)	0.067*** (0.005)	0.066*** (0.006)	0.076*** (0.006)	0.056*** (0.006)
Utdanning	0.043 (0.001)	0.043*** (0.002)	0.054*** (0.002)	0.042*** (0.002)	0.030*** (0.002)
Erfaring	0.019 (0.005)	0.018*** (0.009)	0.022*** (0.010)	0.018*** (0.010)	0.015*** (0.009)
Erfaring ²	-0.0003 (0.001)	-0.0002*** (0.002)	-0.0003*** (0.002)	-0.0003*** (0.002)	-0.0002*** (0.002)
Kjønn	-0.120 (0.000)	-0.076*** (0.000)	-0.172*** (0.000)	-0.129*** (0.000)	-0.119*** (0.000)
R ²	0.396	0.277	0.373	0.266	0.228
Observasjoner	12643	4050	2945	3020	2628

Note: Regresjonskoeffisientene er estimert ved å tilpasse en OLS-regresjon av log-timeslønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring og antall år arbeidserfaring kvadrert. For det aggregerte datasettet «Norden» er regresjonen gjennomført med LMM med respondentenes nasjonalitet som stokastisk effekt. Tilhørende standardavvik står i parentes under. Grunnet problemer med å estimere frihetsgrader i *t*-fordelingen for LMM er ikke *p*-verdiene til regresjonskoeffisientene i norden-tilpasningen rapportert. Bare respondenter eldre enn 25 år er inkludert.

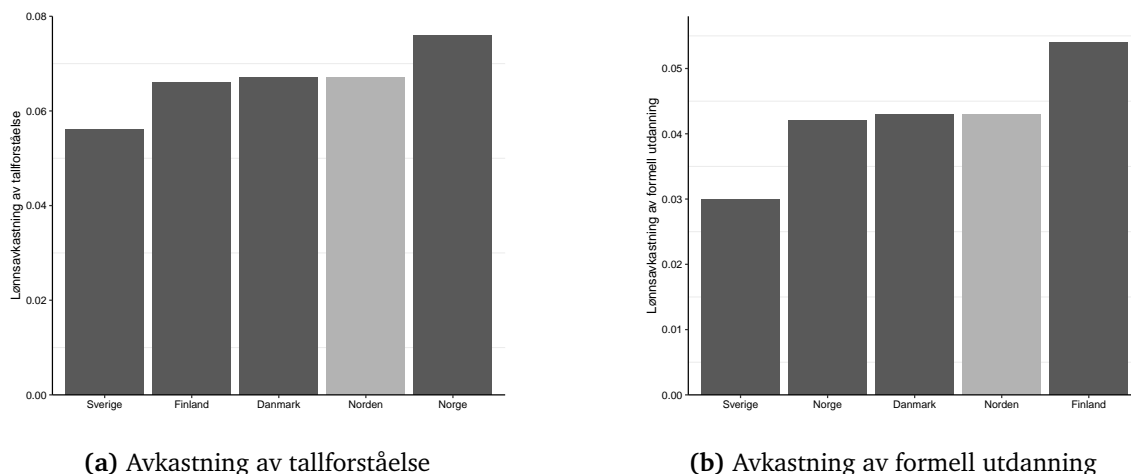
* $p < 0.10$

** $p < 0.05$

*** $p < 0.01$

Data fra: PIAAC.

Dersom en ikke kontrollerer for respondentenes tallforståelse eller andre kognitive ferdigheter, tilsvarende analysene i Mincer (1974), er lønnsavkastningen av et ekstra år med utdanning estimert til å være 5.3 % for norden som helhet; betinget på individenes tallforståelse reduseres lønnsavkastningen av utdanning til 4.3 %. Lønnspremien av ett ekstra år med utdanning er i ganske lik for de nordiske landene, og lønnspremien varierer fra 3.0 % i Sverige til 5.4 % i Danmark.



Figur 3: Estimert avkastning av (a) kognitive ferdigheter innen tallforståelse og (b) formell utdanning i nordiske land. *Note:* Regresjonskoeffisientene er estimert med en OLS-regresjon av log-timeslønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring og antall år arbeidserfaring kvadrert. For det aggregerte datsettet «Norden» er regresjonen gjennomført med LMM med respondentenes nasjonalitet som stokastisk effekt. Bare respondenter eldre enn 25 år er inkludert. Resultatene replikerer tilpasningen fra tabell 2. *Data fra:* PIAAC.

Gjennomsnittlig lønnspremie for ett standardavviks økning i tallforståelse er estimert til 6.7 % i norden, med forholdsvis små variasjoner innad i landene. (Vi ser at den største variasjonen mellom de nordiske landene er den negative lønnspremien for kvinner, som varierer fra -7.6% i Danmark til -17.2% i Finland.) Utføring av Chow-tester mellom de ulike landenes estimerte lønnslikninger viser dog at tilpasningene er statistisk ulike.

5.2 Avkastning av utdanningsnivå og interaksjon med ferdigheter

For å undersøke hvordan lønnspremien av høyere utdanning varierer mellom nordiske land, og hvorvidt høyere utdannede blir bedre premiert for høye kognitive ferdigheter enn lavere utdannede blir, tilpasser vi likning (18) for hvert land med OLS. Heri deler vi opp respondentene i de tre utdanningsnivåene grunnskoleutdanning (6–9 år utdanning), videregående (10–12 år) og høyere utdanning (13–20 år). Regresjonstilpasningene er presentert i tabell 3.¹

I Danmark, Finland og Norge ser vi en interaksjonseffekt hvor lønnavkastningen av kognitive ferdigheter er statistisk signifikant høyere for arbeidere med en universitetsgrad enn for arbeidere for øvrig; norske arbeidere med universitetsgrad har en estimert gjennomsnittlig 15.3 % lønnspremie av ett standardavviks økning i kognitive ferdigheter vis-à-vis arbeidere med bare grunnskoleutdanning, og denne premien er 3.8 ganger større i Norge enn i Danmark (4.0 % lønnspremie) og 2.0 ganger større enn i Finland (7.5 % lønnspremie). For Sverige er interaksjonseffekten mellom utdanningsnivå og lønnsavkastning av kognitive ferdigheter veldig liten og ikke statistisk signifikant.

I det videre, der vi tidligere i tabell 2 og figur 3 observerte marginale forskjeller i lønnavkastningen i formell utdanning mellom de nordiske landene, forteller tabell 3 en noe ulik historie. For samtlige

¹Murnane et al. (2000) og Blackburn and Neumark (1993) gjennomfører et tilsvarende studie hvor ferdighetene til respondentene var målt etter at de var ferdige på videregående og før de eventuelt begynte på en høyere utdanning. Her ser vi derimot på ferdigheter målt på et «tilfeldig» tidspunkt etter ferdig utdanning.

Tabell 3: Avkastning av utdanningsnivå og kognitive ferdigheter gitt utdanningsnivå

	<i>(Norden)</i>	<i>(Danmark)</i>	<i>(Finland)</i>	<i>(Norge)</i>	<i>(Sverige)</i>
Tallforståelse	0.0475 (0.010)	0.0596*** (0.015)	0.0330 (0.019)	-0.0418 (0.068)	0.0660*** (0.019)
×Videregående	0.0162 (0.011)	-0.0038 (0.017)	0.0360 (0.022)	0.1300 (0.070)	-0.0058 (0.021)
×Universitet	0.0523 (0.010)	0.0400** (0.017)	0.0751*** (0.021)	0.1523** (0.068)	-0.0054 (0.070)
Videregående	0.0844 (0.013)	0.1106*** (0.019)	0.0885*** (0.023)	0.2190 (0.180)	0.0012 (0.024)
Universitet	0.2443 (0.013)	0.2623*** (0.020)	0.3060*** (0.023)	0.3727** (0.170)	0.1258*** (0.024)
R^2	0.376	0.251	0.318	0.243	0.217
Observasjoner	12643	4050	2945	3020	2628

Note: Regresjonskoeffisientene er estimert ved å tilpasse en OLS-regresjon for hvert land for log-timeslønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, utdanningsnivå (grunnskole (6–9 år), videregående (10–12 år) og høyere utdanning (13–20 år)) og en interaksjonseffekt mellom tallforståelse og utdanningsnivå. Tilhørende standardavvik står i parentes under. For det aggregerte datasettet «Norden» er regresjonen gjennomført med LMM med respondentenes nasjonalitet som stokastisk effekt. Grunnet problemer med å estimere frihetsgrader i t -fordelingen for LMM er ikke p -verdiene til regresjonskoeffisientene i norden-tilpasningen rapportert. Bare respondenter eldre enn 25 år er inkludert.

* $p < 0.10$

** $p < 0.05$

*** $p < 0.01$

Data fra: PIAAC.

land er det en statistisk signifikant lønnseffekt av å ha en universitetsgrad, og sammenliknet med å ha grunnskole som høyeste utdanning varierer denne lønnspremien fra 12.6 % i Sverige til 37.3 % i Norge. I Danmark og Finland er det en statistisk signifikant lønnspremie av videregående utdanning vis-à-vis bare grunnskole på henholdsvis 11.1 og 8.9 %, mens Norge har en tilsvarende (dog ikke statistisk signifikant) lønnspremie på 21.9 %. I Sverige er lønnspremien av videregående utdanning vis-à-vis grunnskoleutdanning liten og ikke statistisk signifikant.

5.3 Aldersvarierende avkastning av ferdigheter og utdanning

For å studere om lønnsavkastningen av kognitive ferdigheter varierer mellom ulike aldersgrupper, deles arbeidstyrken opp i oppstartsalder (25–34 år), moden alder (35–54 år) og avsluttende alder (55–65 år), tilsvarende analysene i Hanushek et al. (2015). Dernest tilpasses likning (19a), se tabell 4. I figur 4 er en tilsvarende analyse for 5-års aldersgrupper gjennomført, og lønnsavkastningen av kognitive ferdigheter for hver aldersgruppe er rapportert. Altonji and Pierret (1998) forklarer økt avkastning av ferdigheter for økt alder med at arbeidsgiver statistisk diskriminerer unge arbeidere i deres utdanning, ettersom arbeidsgiver ikke kan observere arbeidernes ferdigheter; for økt alder er det derfor rimelig å forvente at avkastningen av formell utdanning avtar samtidig som arbeidsgiveren i større grad

Tabell 4: Avkastning av kognitive ferdigheter for aldersgruppe

	(Norden)	(Danmark)	(Finland)	(Norge)	(Sverige)
Tallforståelse	0.0480 (0.006)	0.0536*** (0.009)	0.0278** (0.012)	0.0703*** (0.010)	0.0153 (0.010)
×Moden alder	0.0352 (0.008)	0.0273** (0.013)	0.0605*** (0.013)	0.0150 (0.012)	0.0600*** (0.012)
×Avsluttende alder	0.0132 (0.012)	-0.0010 (0.013)	0.0331** (0.017)	-0.0072 (0.016)	0.0502*** (0.015)
Moden alder	0.0369 (0.012)	0.0537*** (0.016)	0.0203 (0.019)	0.0400** (0.017)	0.0094 (0.017)
Avsluttende alder	0.0237 (0.009)	0.0421** (0.021)	-0.0085 (0.027)	0.0400 (0.025)	0.0079 (0.025)
Utdanning	0.0424 (0.001)	0.0428*** (0.002)	0.0537*** (0.002)	0.0413*** (0.002)	0.0287*** (0.002)
R^2	0.3994	0.281	0.379	0.298	0.236
Observasjoner	12643	4050	2945	3020	2628

Note: Regresjonskoeffisientene er estimert ved å tilpasse en OLS-regresjon for hvert land for log-timeslønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, aldersgruppe (oppstartsalder (25–34 år), moden alder (35–54 år) og avsluttende alder (55–65 år)) og en interaksjonseffekt mellom tallforståelse og aldersgruppe. Tilhørende standardavvik står i parentes under. For det aggregerte datsettet «Norden» er regresjonen gjennomført med LMM med respondentenes nasjonalitet som stokastisk effekt. Grunnet problemer med å estimere frihetsgrader i t -fordelingen for LMM er ikke p -verdiene til regresjonskoeffisientene i norden-tilpasningen rapportert. Bare respondenter eldre enn 25 år er inkludert.

* $p < 0.10$

** $p < 0.05$

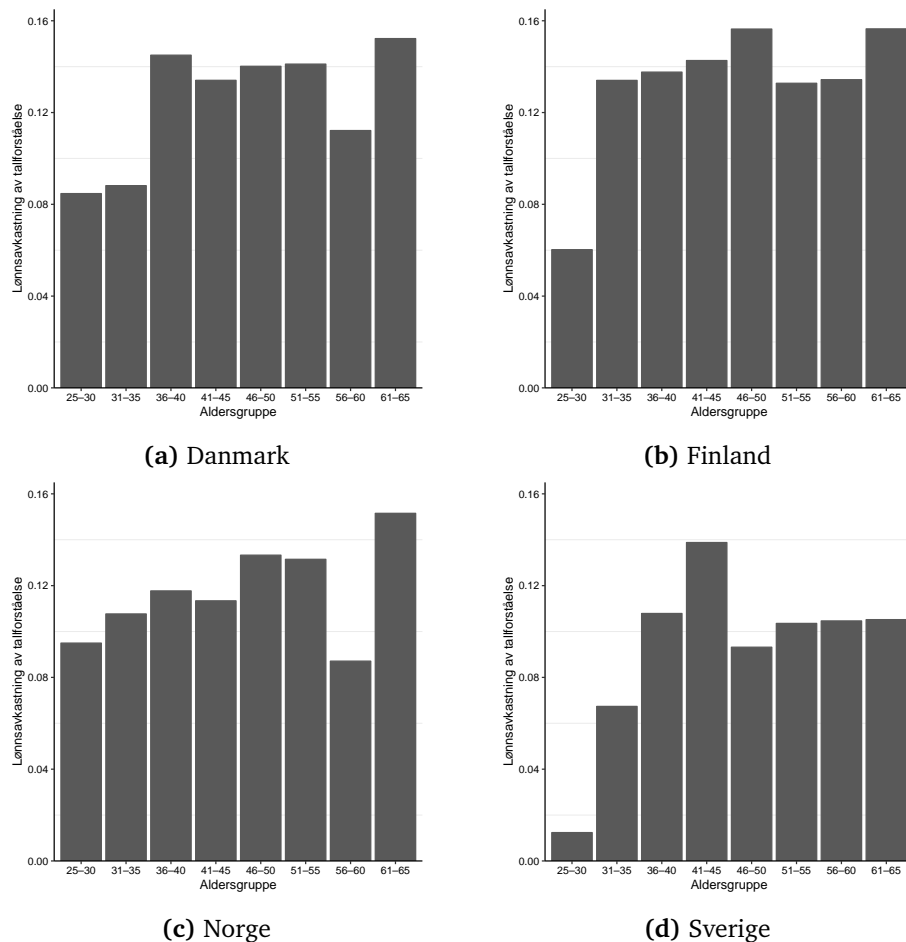
*** $p < 0.01$

Data fra: PIAAC.

observerer arbeidstakerens ferdigheter. Tilpasning av likning (19b) med samme aldersgrupper gir derimot ingen statistisk signifikante alderseffekter på avkastningen av formell utdanning i nordiske land.

For Finland og Sverige ser vi at betinget på aldersgruppe reduseres basiseffekten av tallforståelse med henholdsvis 58.2 og 72.5 %. Sverige og Finland opplever derimot store lønnspremier av tallforståelse for eldre aldersgrupper: Modne arbeidere i Sverige har en 6.0 % premie på økt tallforståelse sammenliknet med arbeidere i oppstartsalder, og arbeidere i avsluttende alder har en 5.0 % premie på økt tallforståelse sammenliknet med arbeidere i oppstartsalder. I Danmark har modne arbeidere en 2.7 % lønnspremie av økt tallforståelse sammenliknet med arbeidere i oppstartsalder, mens lønnspremien av økt tallforståelse for gamle arbeidere sammenliknet med unge arbeidere er smått negativ og ikke signifikant. For Norge fremgår det at eldre arbeidere har en høyere lønn, men det observeres ingen statistisk signifikant alderseffekter på avkastningen av kognitive ferdigheter.

Med grunnlag i de observerte aldersvariasjonene på avkastningen av ferdigheter kan en spørre seg om analysen i det tidligere avsnitt bryter antakelsen om uavhengige observasjonspar: Står unge, modne og avsluttende arbeidere ovenfor samme lønnslikning? Dersom vi tilpasser en basislikning



Figur 4: Avkastning av kognitive ferdigheter innen tallforståelse for ulike aldersgrupper i nordiske land. *Note:* Regresjonskoeffisientene er estimert i en OLS-regresjon, for 5-års aldersgrupper, av log-lønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), kjønn, antall år formell utdanning, antall år arbeidserfaring og antall år arbeidserfaring kvadrert. Bare fulltidsarbeidere er inkludert. Bare respondenter eldre enn 25 år er inkludert. *Data fra:* PIAAC.

(17) for hver aldersgruppe i de ulike landene, viser Chow-tester at i Sverige står unge ovenfor en annen lønnslikning enn modne og avsluttende, i Danmark står modne ovenfor en annen lønnslikning enn unge, i Norge står modne ovenfor en annen lønnslikning enn avsluttende, og i Finland står unge ovenfor en annen lønnslikning enn modne og avsluttende. Dersom en komplementerer likning (17) med respondentens alder, og legger til en toveis-interaksjonseffekt mellom alder og avkastningen av kognitive ferdigheter, forblir resultatene av Chow-testene de samme, se tabell 5.

Det observeres altså en aldersvariasjon på timeslønn vi ikke klarer å fange opp i likning (17), selv om vi kontrollerer for alder. En potensiell løsning er å tilpasse likning (17) med LMM for hvert land med aldersgruppe som stokastisk effekt; dette gir en avkastningen av kognitive ferdigheter og en avkastningen av formell utdanning som er så godt som lik tilpasningen i tabell 2. Det fremgår dog en liten økning (22.5, 15.5, 12.5 og 7.1 % i henholdsvis Danmark, Finland, Norge og Sverige²) i lønnsavkastningen av antall år betalt arbeidserfaring.

²For Sverige er LMM-tilpasningen gjort uten arbeidserfaring kvadrert, da dette ga en singularer modell.

Tabell 5: Hypotesetest om likhet av lønnslikninger for aldersgrupper

(Danmark)	(Finland)	(Norge)	(Sverige)
$\hat{f}_D^u \neq \hat{f}_D^{m***}$	$\hat{f}_F^u \neq \hat{f}_F^{m***}$	–	$\hat{f}_S^u \neq \hat{f}_S^{m***}$
–	$\hat{f}_F^u \neq \hat{f}_F^{g***}$	–	$\hat{f}_S^u \neq \hat{f}_S^{g***}$
–	–	$\hat{f}_N^m \neq \hat{f}_N^{g**}$	–

Note: Forkastede hypoteser i Chow-tester på hypotesene $H_0 : \hat{f}_i^j = \hat{f}_k^j$ mot $H_1 : \hat{f}_i^j \neq \hat{f}_k^j$, for aldersgruppe $i \in \{\text{oppstartsald} (25\text{--}34 \text{ år}), \text{moden alder} (35\text{--}54 \text{ år}), \text{avsluttende alder} (55\text{--}65 \text{ år})\}$ og land $j \in \{\text{Danmark, Finland, Norge, Sverige}\}$, der \hat{f} er estimert lønnslikning av log-timeslønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år betalt arbeidserfaring, antall år betalt arbeidserfaring kvadrert, alder og en toveis-interaksjonseffekt mellom alder og tallforståelse. Bare respondenter eldre enn 25 år er inkludert.

* $p < 0.10$

** $p < 0.05$

*** $p < 0.01$

Data fra: PIAAC.

5.4 Avkastning av flere typer kognitive ferdigheter

Til nå i denne oppgaven er utelukkende tallforståelse brukt til å studere lønnsavkastningen av kognitive ferdigheter i nordiske land. Sammenliknet med tidligere kartleggingstudier av voksenbefolkningens ferdigheter, som f.eks. International Adult Literacy Survey (IALS), formidler PIAAC en rikere kartlegging av ferdigheter gjennom mindre korrelerte ferdighetsscorer i leseferdigheter og tallforståelse. I tillegg inkluderer PIAAC den nye ferdighetskategorien problemløsning i IKT-miljø. Mindre korrelerte ferdighetsscorer muliggjør identifisering og analyser av hver ferdighets respektive lønnsavkastning.

Først tilpasser vi likning (17) for hvert land med bare én av ferdighetsscorene inkludert. Derneft tilpasses likningen med alle ferdighetsscorene inkludert som egne kontrollvariabler. Tilpasningene er rapportert i tabell 6.

Når ferdighetene er inkludert hver for seg, ser vi at avkastningen av tallforståelse og leseferdigheter et tilnærmet like i Danmark og Sverige, mens avkastningen av tallforståelse er 19.4 og 15.8 % større enn avkastningen av leseferdigheter i Finland og Norge. Avkastningen av problemløsning i IKT-miljø er generelt noe mindre enn avkastningen av tallforståelse og leseferdigheter i alle nordiske land. Vi ser at avkastningen av formell utdanning er tilnærmet lik uansett hvilken ferdighetsscore som inkluderes i likning (17).

En analyse av den simultane tilpasningen med alle tre ferdighetsscorer inkludert må gjøres med forsiktighet, ettersom en høy lineær korrelasjon mellom kontrollvariabler medfører mindre presise parameterestimer og høyere estimerte standardavvik for regresjonsparameterene. Kvalitativt ser vi derimot at i Danmark, Finland og Norge dominerer tallforståelse leseferdigheter, slik at avkastningen av leseferdigheter ikke er statistisk signifikant. (Avkastningen av høyere leseferdigheter er sågar smått negativ i Norge, hvilket ikke er intuitivt rimelig og nødvendigvis skyldes den høye korrelasjonen mellom testscoren i leseferdigheter og testscoren i tallforståelse.) I Sverige er situasjonen annerledes, hvor leseferdigheter er statistisk signifikant og høyere enn den ikke-signifikante avkastningen av

Tabell 6: Avkastning av formell utdanning og kognitive ferdigheter for ulike ferdighetsområder: tallforståelse, leseferdigheter og problemløsning i IKT-miljø

	(1)	(2)	(3)	(4)
Tallforståelse	0.065*** (0.005)			0.035*** (0.011)
Leseferdigheter		0.067*** (0.005)		0.018 (0.012)
Problemløsning			0.057*** (0.005)	0.022** (0.009)
Utdanning	0.043*** (0.002)	0.043*** (0.002)	0.046*** (0.002)	0.043*** (0.002)
R ²	0.277	0.277	0.262	0.267
Observasjoner	4052	4052	3571	3571

(Danmark)

	(1)	(2)	(3)	(4)
Tallforståelse	0.067*** (0.006)			0.054*** (0.011)
Leseferdigheter		0.054*** (0.006)		0.018 (0.013)
Problemløsning			0.044*** (0.006)	-0.003 (0.010)
Utdanning	0.053*** (0.002)	0.056*** (0.002)	0.058*** (0.002)	0.054*** (0.002)
R ²	0.373	0.364	0.363	0.375
Observasjoner	2945	2945	2597	2597

(Finland)

	(1)	(2)	(3)	(4)
Tallforståelse	0.076*** (0.006)			0.051*** (0.013)
Leseferdigheter		0.065*** (0.006)		-0.012 (0.013)
Problemløsning			0.063*** (0.006)	0.039*** (0.010)
Utdanning	0.042*** (0.002)	0.044*** (0.002)	0.045*** (0.003)	0.043*** (0.003)
R ²	0.296	0.286	0.282	0.287
Observasjoner	3020	3020	2734	2734

(Norge)

	(1)	(2)	(3)	(4)
Tallforståelse	0.056*** (0.006)			0.017 (0.013)
Leseferdigheter		0.059*** (0.006)		0.028** (0.013)
Problemløsning			0.052*** (0.006)	0.025*** (0.009)
Utdanning	0.030*** (0.002)	0.030*** (0.002)	0.030*** (0.002)	0.028*** (0.002)
R ²	0.234	0.230	0.220	0.226
Observasjoner	2628	2628	2423	2423

(Sverige)

Note: Regresjonskoeffisientene er estimert ved å tilpasse en OLS-regresjon for hvert land av log-timeslønn på kognitive ferdigheter (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring og antall år arbeidserfaring kvadrert. Tilhørende standardavvik står i parentes bak hver koeffisient. Bare respondenter eldre enn 25 år er inkludert.

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Data fra: PIAAC.

tallforståelse.

Med unntak av i Finland er dessuten avkastningen av problemløsning i IKT-miljø positiv og statistisk signifikant i den simultane tilpasningen for nordiske land. Avkastningen av problemløsningsferdigheter er 37.1 og 23.5 % lavere enn avkastningen av tallforståelse i Danmark og Norge, mens avkastningen av problemløsningsferdigheter er 10.7 % lavere enn avkastningen av leseferdigheter i Sverige. I den simulatene tilpasning reduseres i det videre avkastningen av tallforståelse med 46.2, 19.4 og 32.9 % i henholdsvis Danmark, Finland og Norge, og avkastningen av leseferdigheter reduseres tilsvarende med 52.5 % i Sverige. Den simultane tilpasningen antyder følgelig at kognitive ferdigheter innen problemløsning i IKT-miljø tilfører en unik ferdighetslønnsavkastning ikke fanges opp av kognitive ferdigheter innen tallforståelse eller leseferdigheter i Danmark, Norge og Sverige.

5.5 Demografisk varierende avkastningen av ferdigheter og utdanning

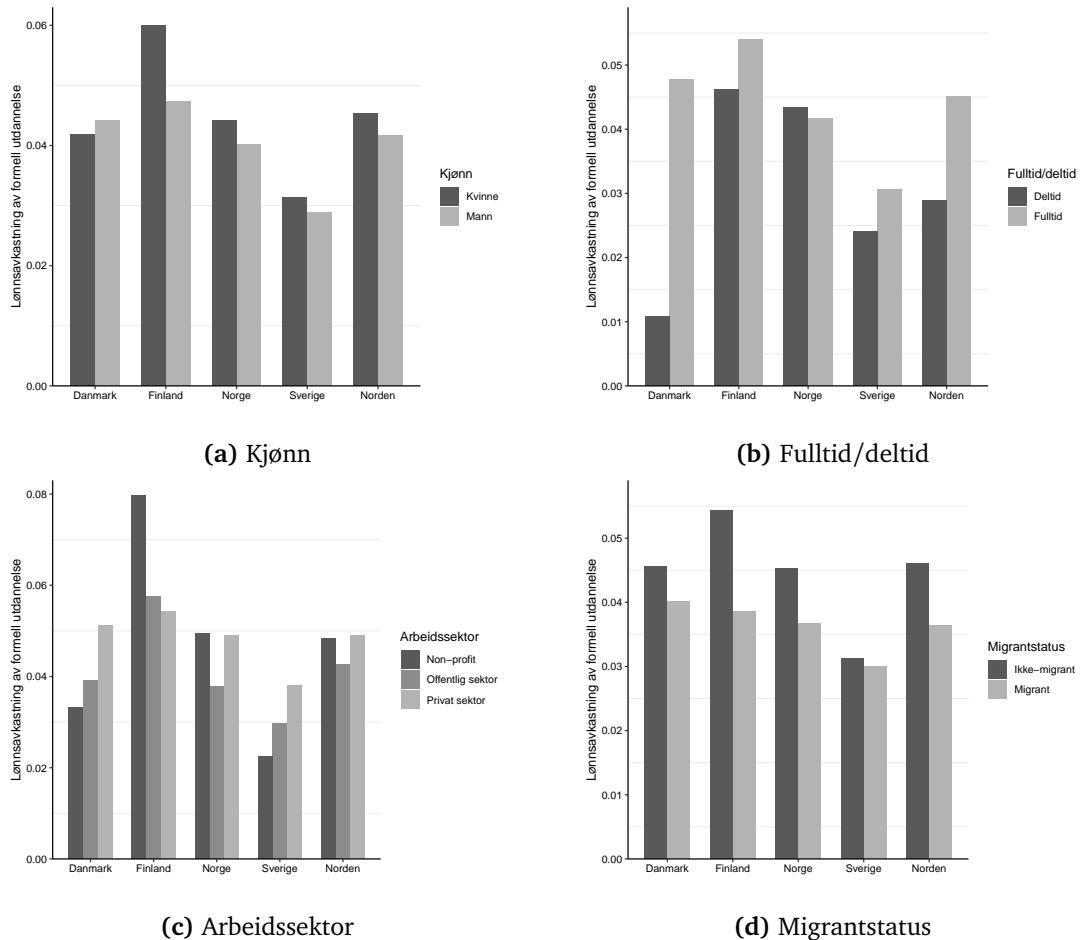
For å studere hvorvidt avkastningen av kognitive ferdigheter og formell utdanning varierer mellom demografiske inndelinger av befolkningen, ble likning (20) tilpasset med OLS for hvert nordiske land. Her ble det studert forskjeller mellom menn og kvinner, fulltidsarbeidere og deltidsarbeidere, arbeidssektorer (privat sektor, offentlig sektor og ideell organisasjon), samt mellom migranter og ikke-migranter. Avkastning av formell utdanning er presentert i tabell A1 i appendikset og i figur 5, og avkastningen av kognitive ferdigheter innen tallforståelse mellom demografiske grupper er presentert i tabell A2 i appendikset og i figur 6.

I Finland har kvinner en statistisk signifikant 1.3 % lønnspremie på avkastning av formell utdanning sammenliknet med menn. I Danmark og Norge har menn en høyere avkastning av kognitive ferdigheter enn kvinner (1.7 og 2.0 % lønnspremie), mens finske kvinner har en 3.8 % lønnspremie på avkastning av kognitive ferdigheter sammenliknet med finske menn. I Sverige fremgår det ingen statistisk signifikante kjønnseffekter på avkastningen av formell utdanning eller kognitive ferdigheter.

I Danmark har fulltidsarbeidere en 3.7 % lønnspremie på avkastning av formell utdanning sammenliknet med deltidsarbeidere. Alle nordiske land har en statistisk signifikant høyere lønnsavkastning av kognitive ferdigheter for fulltidsarbeidere enn for deltidsarbeidere, og denne premien varierer fra 3.0 og 3.3 % i Norge og Finland til 7.0 og 7.5 % i Sverige og Danmark.

Danmark, Norge og Sverige observerer alle en liknende lønnseffekt mellom offentlig og privat sektor. I Danmark, Norge og Sverige har offentlig sektor en lavere avkastning av formell utdanning enn arbeidere har i privat sektor (en lønnspremie på henholdsvis -4.3 , -4.7 og -3.5 %), og danske, norsk og svenske arbeidere har tilsvarende en lavere avkastning av kognitive ferdigheter i offentlig sektor enn i privat sektor (en lønnspremie på henholdsvis -1.2 , -1.1 , -0.8 %). Det observeres ellers ingen statistisk signifikante variasjoner i avkastningen av formell utdanning og kognitive ferdigheter mellom privat sektor og ideelle organisasjoner i norden, med unntak av i Finland hvor arbeidere i ideelle organisasjoner en (svakt statistisk signifikant) 3.0 % lønnspremie av formell utdanning sammenliknet med arbeidere i privat sektor.

Det virker som om migranter generelt har en noe lavere avkastning på formell utdanning og kognitive ferdigheter enn ikke-migranter i norden. Den eneste signifikante forskjellen er i Finland, hvor migranter har en -5.1 % lønnspremie av kognitive ferdigheter sammenliknet med ikke-migranter.

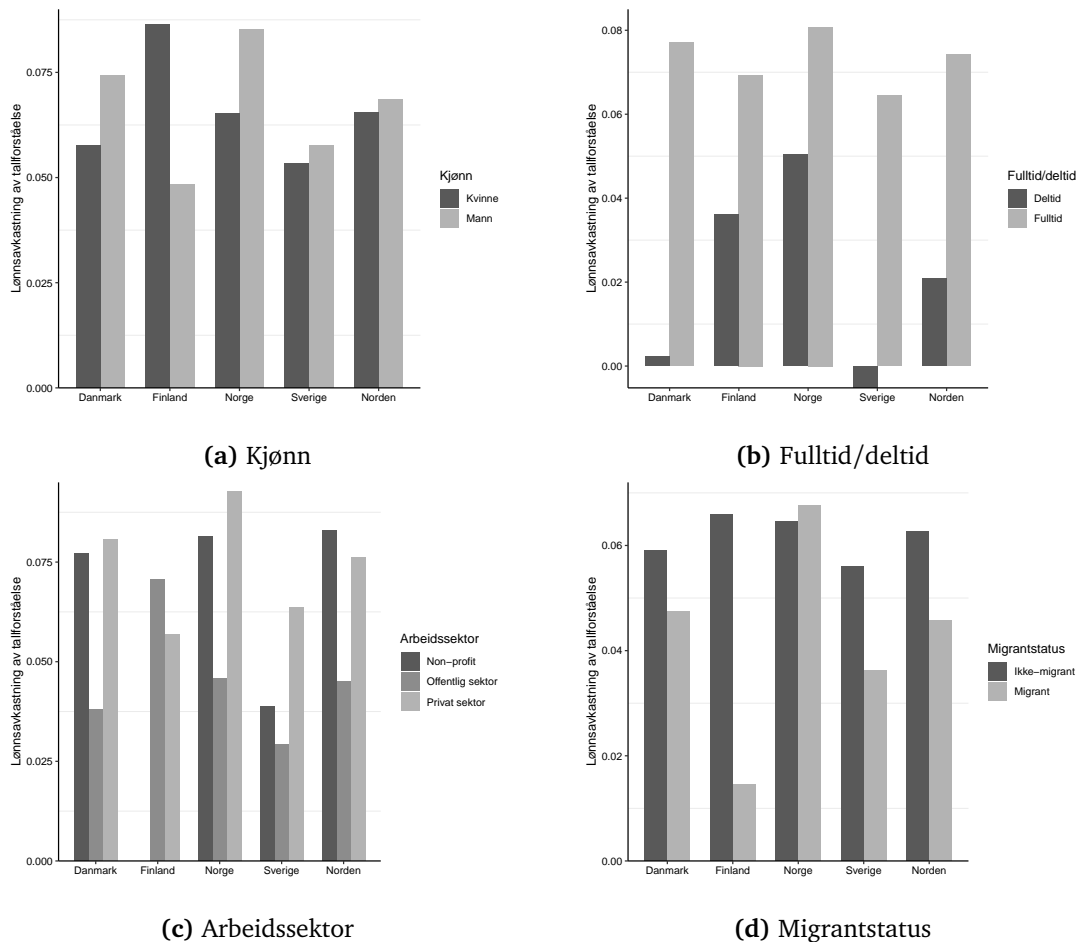


Figur 5: Avkastning av formell utdanning for ulike demografiske grupper, for (a) kjønn, (b) fulltid/deltid, (c) arbeidssektor og (d) migrantstatus. *Note:* Regresjonskoeffisientene er estimert i en OLS-regresjon av log-lønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, demografisk gruppe og en interaksjonseffekt mellom tallforståelse og demografisk gruppe. Resultatene replikerer tilpasningen fra tabell A1. Bare respondenter eldre enn 25 år er inkludert. *Data fra:* PIAAC.

5.6 Arbeidssektor-varierende avkastningen av ferdigheter og utdanning

Vi observerte i seksjon 5.5 at avkastningen av formell utdanning og kognitive ferdigheter varierte mellom privat og offentlig sektor i Danmark, Norge og Sverige. For å videre studere hvordan avkastningen av utdanning og ferdigheter varierer mellom ulike arbeidssektorer definert i henhold til 2008 International Standard Classification of Occupations (ISCO), tilpasses likning (21a) og (21b) med OLS for hvert land. Avkastningen av kognitive ferdigheter innen tallforståelse mellom ulike jobbsektorer gir svært få signifikante estimater, så de blir ikke diskutert her, men regresjonsestimatene kan oppspores i tabell A4 og figur A1 i appendikset. Avkastningen av formell utdanning er presentert i tabell A3 i appendikset og i figur 7.

Administrative ledere og politikere, akademikere og teknikere virker generelt å ha høyest avkastning av formell utdanning i samtlige nordiske land. I Danmark er ikke avkastningsnivået signifikant forskjellig mellom de tre yrkessektorene, og i Norge er ikke avkastningsnivået signifikant forskjellig

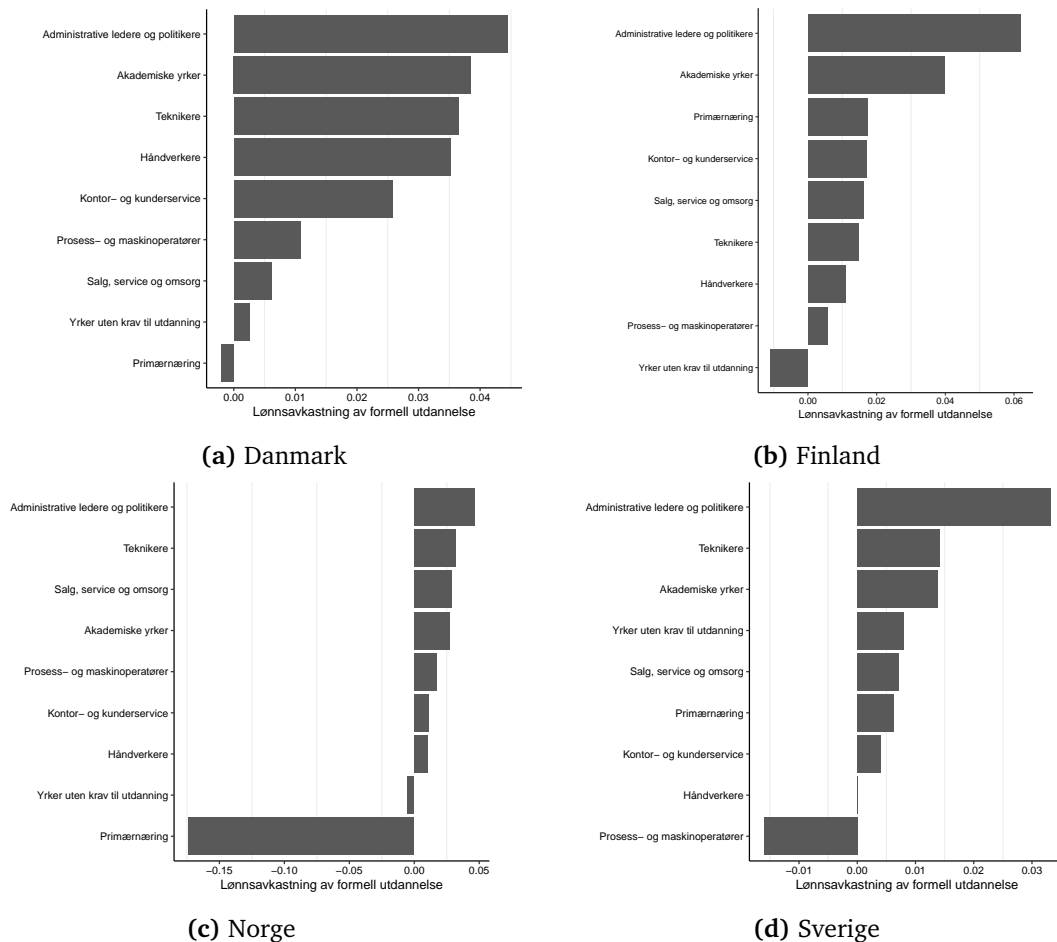


Figur 6: Avkastning av kognitive ferdigheter innen tallforståelse for ulike demografiske grupper, for (a) kjønn, (b) fulltid/deltid, (c) arbeidssektor og (d) migrantstatus. *Note:* Regresjonskoeffisientene er estimert i en OLS-regresjon av log-lønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, demografisk gruppe og en interaksjonseffekt mellom tallforståelse og demografisk gruppe. Resultatene replikerer tilpasningen fra tabell A2. Bare respondenter eldre enn 25 år er inkludert. *Data fra:* PIAAC.

mellom adm. ledere og politikere og teknikere. Finland skiller seg noe ut, hvor teknikeres avkastning av formell utdanning vis-à-vis adm. ledere og politikere er -4.7% og litt under sektor-gjennomsnittet.

I Norge har arbeidere i primærnæringen en voldsom statistisk signifikant negativ avkastning av formell utdanning på -17.4% . I Danmark har òg arbeidere i primærnæringen en smått negativ avkastning av formell utdanning på -0.3% , mens i Sverige og Finland er avkastningen av formell utdanning i primærnæringen positiv og litt under sektor-gjennomsnittet. Tilsvarende har svenske prosess- og maskinoperatører en negativ premie av formell utdanning på -1.6% ; for de andre nordiske landene befinner sektoren seg noe under gjennomsnittet, men opplever en positiv lønnspremie.

Spredningen i avkastning av formell utdanning mellom sektorene er størst i Norge og Sverige, og minst i Danmark. Dersom vi fjerner sektoren med lavest avkastning fra hvert av landene (slik at spredningen ikke skal bli veldig påvirket av primærnæring i Norge og av prosess- og maskinoperatører i Sverige), forblir rangeringen den samme. Standardavviket blant norske sektorer er 2.5 større enn standardavviket blant svenske sektorer og 7.6 ganger større enn standardavviket blant danske sektorer.



Figur 7: Avkastning av formell utdanning for ulike arbeidssektorer, definert i henhold til 2008 ISCO, i nordiske land. *Note:* Regresjonskoeffisientene er estimert i en OLS-regresjon av log-lønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, arbeidssektor og en interaksjonseffekt mellom formell utdanning og arbeidssektor. Resultatene replikerer tilpasningen fra tabell A3. Bare respondenter eldre enn 25 år er inkludert. *Data fra:* PIAAC.

Standardavviket blant finske sektorer er 24.5 % større enn standardavviket blant danske sektorer.

5.7 Diagnostisk analyse for tilpasning av likning (17)

Ingen regresjonsmodell er korrekt, men antakelsene tilpassede modeller hviler på bør tilnærmet møtes. For OLS-regresjon er antakelsene om de stokastiske feilene viktige for å kunne estimere signifikansen til regresjonsparameterene. For basislikningen (17) ønsker vi følgelig å studere hvorvidt feilene er homoskedastiske, normalfordelte og stokastisk uavhengige av kontrollvariablene. Liknende studier kan gjøres for de andre tilpassede likningene og gir tilsvarende resultater.

I figur A2 i appendikset er QQ-plot av de studentiserte residualene for tilpasningen til de nordiske landene i tabell 2 presentert. Residualene for Finland- og Sverige-tilpasningene virker tilnærmet å følge en normalfordeling, mens Norge- og Danmark-tilpasningene lider av svært tunge haler. Tunge haler problematiserer resultatene av t -testene for signifikansnivå av estimerte regresjonsparametere, men ettersom tilpasningene er gjort med henholdsvis 4050 og 3020 observasjoner er det rimelig at

den asymptotiske fordelingen til regresjonsparameterene i likning (6) er mer eller mindre gyldig, slik at de estimerte signifikansnivåene ikke er helt urimelige.

I figur A3 i appendikset er de studentiserte residualene for Danmark-tilpasningen i tabell 2 plottet opp mot estimert log-timeslønn og de brukte kontrollvariablene. Plott for de tre andre nordiske landene fremviser samme natur. Kjønn, antall år betalt arbeidserfaring og antall år betalt arbeidserfaring kvadrert virker å være ukorrelert med residualene. Vi observerer derimot en hetroskedastisk varians for estimert log-timeslønn, kognitive ferdigheter innen tallforståelse og antall år formell utdanning. Dette er ikke veldig overraskende da vi over blant annet har observert intraalder- og intrasektorvariasjon i avkastningen av formell utdanning og kognitive ferdigheter, i tillegg til at vi har observert en toveis-interaksjonseffekt mellom avkastningen av utdanning og ferdigheter.

Summa summarum fremgår det fra asymptotiske resultater at de estimerte signifikansnivåene i denne rapporten ikke er helt urimelige, men figur A2 og A3 antyder vi burde revurdere formuleringen av lønnslikningen i kontrollvariablene utdanning og ferdigheter, for å kunne kreere en mer adekvat og overbevisende statistisk modell.

OPPSUMMERING

Vi har i denne rapporten gjennom statistiske analyser av PIAAC-datasettet for de nordiske landene Danmark, Finland, Norge og Sverige undersøkt sammenhengen mellom lønnsdannelse og formell utdanning og kognitive ferdighetsmål. Det har videre blitt sett på variasjoner på tvers av aldersgrupper, demografier og arbeidssektorer for hvert enkelt land.

Grunnavkastningen av formell utdanning og kognitive ferdigheter ble funnet til å være statistisk signifikant positiv i alle nordiske land, og denne avkastningen varierer lite mellom landene. Norge virker å premiere ferdigheter høyest, samtidig som Finland virker å premiere utdanning høyest. Kognitive ferdigheter innen tallforståelse er den sterkeste predikatoren for lønnsdannelse av de tre ferdighetsmålene i PIAAC-undersøkelsen i alle land med unntak av i Sverige, hvor leseferdigheter er sterkere. Kognitive ferdigheter innen problemløsning i IKT-miljø virker å fange opp en unik lønnseffekt ikke omfattet av tallforståelse og leseferdigheter i alle land med unntak av i Finland.

Norge virker å premiere grader fra høyere utdanning mer enn de andre landene, og i Norge, Finland og Danmark får respondenter med høyere kognitive ferdigheter signifikant høyere avkastning av høyere utdanning. Denne effekten er ikke signifikant i Sverige, som virker å betale betraktelig mindre for høyere utdanninger enn de andre nordiske landene.

I alle land, med unntak av Norge, observeres i det videre en aldersvarierende avkastning av kognitive ferdigheter hvor eldre arbeidere premieres bedre for kognitive ferdigheter enn yngre arbeidere. Denne effekten er størst i Sverige og Finland. En tilsvarende negativ effekt på avkastningen av formell utdanning for økt alder, som blant annet Altonji and Pierret (1998) postulerer, observeres ikke i noen av landene.

I Danmark, Norge og Sverige er avkastningen av utdanning og ferdigheter i offentlig sektor signifikant lavere enn i privat sektor, og denne effekten er omtrent like stor i de tre landene. Mellom ulike arbeidssektorer er fordelingen og størrelsen av avkastningen av utdanning ganske lik mellom landene, men det observeres noen store forskjeller blant arbeidssektorene med lavest avkastning

av utdanning: I Norge har primærnæringen en stor negativ lønnspremie av formell utdanning, og i Sverige har prosess- og maskinoperatører en stor negativ lønnspremie av formell utdanning. Avkastningen av kognitive ferdigheter mellom ulike arbeidssektorer er i store trekk ikke statistisk signifikant for noen av de nordiske landene.

Fremtidige studier kan gå i mange ulike retninger for å forbedre undersøkelsene foretatt i denne rapporten. Enkelte ting er mer åpenbare, slik som å inkludere flere kontrollvariabler eller interaksjonseffekter. Diagnostisk analyse av residualene til de tilpassede lønnslikningene i denne rapporten har problematisert bruken av den utvidede Mincer-modellen som utgangspunkt for analyser på PIAAC-datasettet. Det er derimot mulig at andre faktorer enn de som er inkorporert i rapportens modeller vil kunne forklare lønnsforskjeller og produsere mer adekvate statistiske modeller.

Datasettet rapporten bygger på skiller ikke mellom ulike typer utdanning, og således åpnes det ikke opp for muligheten til å se på variasjoner mellom ulike retninger innen akademia. Det kan være naturlig å anta at utdanninger som er spesielt etterspurt vil gi en høyere lønnspremie, og også at lønnspremien på enkelte studieretninger vil være negativ. En undersøkelse av disse spørsmålene vil være særdeles interessant i et samfunn der svært mange velger høyere utdanning, samtidig som arbeidsmarkedet er preget av høy konkurranse.

Videre kan det være av interesse å undersøke hvilke bakenforliggende forklaringer som manifesterer seg i resultatene som kommer frem av denne rapporten. En økonometrisk sammenheng er interessant i seg selv, men det vil være naturlig å finne kvalitative forklaringsmodeller for hvorfor vi observerer disse effektene. Dette vil innebære studier som ikke kan sies å falle innenfor økonometri, men som likeledes vil kunne bidra til å underbygge eller kaste nytt lys på resultatene fra denne rapporten.

Til slutt er det verdt å nevne at økte lønninger ikke nødvendigvis er den eneste årsaken til at personer velger å ta høyere utdanning; i vår moderne kunnskapsøkonomi vil flere år med studier kunne åpne opp for et bredere spekter av jobbmuligheter. Mer kvalitative mål på f.eks. jobbtilfredshet og livskvalitet vil kanskje være interessante å studere i sammenheng med høyere utdanning, så vel som kognitive ferdigheter.

LITTERATURLISTE

- Alan Agresti. *Foundations of linear and generalized linear models*. John Wiley & Sons, 2015.
- Joseph G Altonji and Charles R Pierret. Employer learning and the signalling value of education. In *internal Labour markets, incentives and Employment*, pages 159–195. Springer, 1998.
- Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*, 2014.
- John Bishop. The impact of academic competencies on wages, unemployment, and job performance. In *Carnegie-Rochester Conference Series on Public Policy*, volume 37, pages 127–194. Elsevier, 1992.
- McKinley L Blackburn and David Neumark. Omitted-ability bias and the increase in the return to schooling. *Journal of labor economics*, 11(3):521–544, 1993.
- John Bynner. Literacy, numeracy and employability. 2002.
- Anthony P Carnevale, Ban Cheah, and Stephen J Rose. The college pay off. 2011.
- Gregory C Chow. Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, pages 591–605, 1960.
- Christopher Colclough, Geeta Kingdon, and Harry Patrinos. The changing pattern of wage returns to education and its implications. *Development Policy Review*, 28(6):733–747, 2010.
- Ludwig Fahrmeir, Thomas Kneib, Stefan Lang, and Brian Marx. *Regression: models, methods and applications*. Springer Science & Business Media, 2013.
- Lynne M Gleeson. *Economic returns to education and training for adults with low numeracy skills*. NCVET, 2005.
- David A Green and W Craig Riddell. Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation. *Labour Economics*, 10(2):165–184, 2003.
- David Alan Green, William Craig Riddell, et al. *Literacy, numeracy and labour market outcomes in Canada*. Statistics Canada Ottawa, 2001.
- Eric A Hanushek, Guido Schwerdt, Simon Wiederhold, and Ludger Woessmann. Returns to skills around the world: Evidence from piae. *European Economic Review*, 73:103–130, 2015.
- Joanne Lindley and Steven McIntosh. Growth in within graduate wage inequality: The role of subjects, cognitive skill dispersion and occupational concentration. *Labour Economics*, 37:101–111, 2015.
- Steven McIntosh and Anna Vignoles. Measuring and assessing the impact of basic skills on labour market outcomes. *Oxford Economic Papers*, 53(3):453–481, 2001.
- Jacob Mincer. Schooling, experience, and earnings. *human behavior & social institutions* no. 2. 1974.

Richard J Murnane, John B Willett, and Frank Levy. The growing importance of cognitive skills in wage determination. Technical report, National Bureau of Economic Research, 1995.

Richard J Murnane, John B Willett, Yves Duhaldeborde, and John H Tyler. How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management*, 19(4):547–568, 2000.

Peter B Robinson and Edwin A Sexton. The effect of education and experience on self-employment success. *Journal of business Venturing*, 9(2):141–156, 1994.

Andreas Schleicher. Piac: A new strategy for assessing adult competencies. *International Review of Education*, 54(5-6):627–650, 2008.

Anthony Shomos. Links between literacy and numeracy skills and labour market outcomes. 2010.

Adam Smith. *On the wealth of nations*. Simon and Schuster, 2013.

Anna Vignoles. What is the economic value of literacy and numeracy? *IZA World of Labor*, 2016.

Ian Walker and Yu Zhu. The college wage premium and the expansion of higher education in the uk. *Scandinavian Journal of Economics*, 110(4):695–709, 2008.

APPENDIKS

Tabell A1: Avkastning av formell utdanning fordelt på demografiske grupper

	<i>(Norden)</i>	<i>(Danmark)</i>	<i>(Finland)</i>	<i>(Norge)</i>	<i>(Sverige)</i>
Menn	0.042 (0.001)	0.044*** (0.003)	0.047*** (0.003)	0.040*** (0.003)	0.029*** (0.003)
Δ Kvinner	0.004 (0.003)	-0.002 (0.003)	0.013*** (0.003)	0.004 (0.004)	0.002 (0.004)
Privat sektor	0.049 (0.001)	0.051*** (0.002)	0.054*** (0.002)	0.049*** (0.003)	0.038*** (0.003)
Δ Offentlig sektor	-0.006 (0.002)	-0.012*** (0.003)	0.003 (0.004)	-0.011* (0.004)	-0.008* (0.004)
Δ Non-profit	-0.001 (0.007)	-0.018 (0.013)	0.030* (0.012)	0.001 (0.019)	-0.016 (0.016)
Deltid	0.029 (0.001)	0.011*** (0.002)	0.046*** (0.002)	0.043*** (0.003)	0.024*** (0.002)
Δ Fulltid	0.016 (0.002)	0.037*** (0.004)	0.009 (0.010)	-0.002 (0.005)	0.006 (0.005)
Ikke-migrant	0.046 (0.001)	0.046*** (0.002)	0.054*** (0.002)	0.045 (0.003)	0.0312*** (0.002)
Δ migrant	-0.010 (0.002)	-0.005 (0.004)	-0.016 (0.010)	-0.009 (0.005)	-0.001 (0.005)

Note: Horisontale linjer separerer uavhengig estimerte modeller. Regresjonskoeffisientene er estimert ved å tilpasse en OLS-regresjon for log-lønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, indikator for demografisk gruppe og en interaksjonseffekt mellom formell utdanning og demografisk gruppe. Hver linje prefigert med en Δ rapporterer den marginale forskjellen i avkastningen av utdanning for gitt demografisk gruppe, som estimert med interaksjonseffekten mellom utdanning og indikatoren for gitt demografisk gruppe. Tilhørende standardavvik står i parentes under. Bare respondenter eldre enn 25 år er inkludert.

* $p < 0.10$

** $p < 0.05$

*** $p < 0.01$

Data fra: PIAAC.

Tabell A2: Avkastning av kognitive ferdigheter innen tallforståelse fordelt på demografiske grupper

	<i>(Norden)</i>	<i>(Danmark)</i>	<i>(Finland)</i>	<i>(Norge)</i>	<i>(Sverige)</i>
Menn	0.069 (0.004)	0.074*** (0.007)	0.048*** (0.008)	0.085*** (0.007)	0.058*** (0.007)
ΔKvinner	-0.003 (0.005)	-0.017* (0.009)	0.038*** (0.011)	-0.020* (0.010)	-0.004 (0.010)
Privat sektor	0.076 (0.003)	0.081*** (0.006)	0.057*** (0.007)	0.093*** (0.007)	0.064*** (0.007)
ΔOffentlig sektor	-0.031 (0.005)	-0.043*** (0.009)	0.014 (0.010)	-0.047*** (0.011)	-0.035*** (0.010)
ΔNon-profit	0.007 (0.019)	-0.004 (0.037)	0.054 (0.035)	-0.011 (0.045)	-0.025 (0.035)
Deltid	0.021 (0.007)	0.002 (0.013)	0.036** (0.018)	0.050*** (0.014)	-0.005 (0.015)
ΔFulltid	0.053 (0.008)	0.075*** (0.014)	0.033* (0.019)	0.030** (0.015)	0.070*** (0.016)
Ikke-migrant	0.063 (0.003)	0.0590*** (0.006)	0.066*** (0.007)	0.065*** (0.007)	0.056*** (0.007)
ΔMigrant	-0.017 (0.006)	-0.012 (0.010)	-0.051* (0.025)	0.003 (0.012)	-0.020 (0.012)

Note: Horisontale linjer separerer uavhengig estimerte modeller. Regresjonskoeffisientene er estimert ved å tilpasse en OLS-regresjon for log-lønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, indikator for demografisk gruppe og en interaksjonseffekt mellom tallforståelse og demografisk gruppe. Hver linje prefigert med en Δ rapporterer den marginale forskjellen i avkastningen av tallforståelse for gitt demografisk gruppe, som estimert med interaksjonseffekten mellom tallforståelse og indikatoren for gitt demografisk gruppe. Tilhørende standardavvik står i parentes under. Bare respondenter eldre enn 25 år er inkludert.

* $p < 0.10$

** $p < 0.05$

*** $p < 0.01$

Data fra: PIAAC.

Tabell A3: Avkastning av formell utdanning fordelt på jobbsektor

	<i>(Danmark)</i>	<i>(Finland)</i>	<i>(Norge)</i>	<i>(Sverige)</i>
Administrative ledere og politikere	0.044*** (0.002)	0.062*** (0.002)	0.047*** (0.003)	0.033*** (0.002)
Δ Akademiske yrker	-0.006 (0.022)	-0.022** (0.029)	-0.019** (0.025)	-0.019** (0.024)
Δ Teknikere	-0.008 (0.024)	-0.047*** (0.029)	-0.015 (0.026)	-0.019** (0.026)
Δ Kontor og kundertjeneste	-0.019** (0.027)	-0.045*** (0.034)	-0.035*** (0.032)	-0.030** (0.035)
Δ Salg, service og omsorg	-0.038*** (0.024)	-0.046*** (0.030)	-0.018* (0.026)	-0.026*** (0.024)
Δ Primærnæringsmiddelindustri	-0.047*** (0.039)	-0.046** (0.042)	-0.221*** (0.034)	-0.027 (0.041)
Δ Håndverkere	-0.009 (0.026)	-0.051*** (0.031)	-0.036*** (0.030)	-0.033*** (0.028)
Δ Prosess- og maskinoperatører	-0.034*** (0.028)	-0.052*** (0.033)	-0.030** (0.032)	-0.049*** (0.028)
Δ Yrker uten krav til utdanning	-0.042*** (0.024)	-0.073*** (0.034)	-0.052*** (0.031)	-0.025** (0.030)

Note: Regresjonskoeffisientene er estimert ved å tilpasse en OLS-regresjon for log-lønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, indikator for jobbsektor og en interaksjonseffekt mellom formell utdanning og jobbsektor. Hver linje prefigert med en Δ rapporterer den marginale forskjellen i avkastningen av formell utdanning for gitt sektor, som estimert av interaksjonseffekten mellom utdanning og indikatoren for gitt sektor. Tilhørende standardavvik står i parentes under. Bare respondenter eldre enn 25 år er inkludert.

* $p < 0.10$

** $p < 0.05$

*** $p < 0.01$

Data fra: PIAAC.

Tabell A4: Avkastning av kognitive ferdigheter innen tallforståelse fordelt på jobbsektor

	<i>(Danmark)</i>	<i>(Finland)</i>	<i>(Norge)</i>	<i>(Sverige)</i>
Administrative ledere og politikere	0.042** (0.020)	0.079*** (0.027)	0.054** (0.025)	0.072*** (0.022)
Δ Akademiske yrker	0.029 (0.021)	-0.053* (0.029)	-0.004 (0.026)	-0.034 (0.024)
Δ Teknikere	0.029 (0.024)	-0.014 (0.029)	0.030 (0.032)	-0.029 (0.026)
Δ Kontor og kundertjeneste	0.027 (0.027)	-0.022 (0.034)	0.001 (0.026)	-0.017 (0.035)
Δ Salg, service og omsorg	-0.017 (0.024)	-0.059* (0.030)	0.003 (0.034)	-0.069*** (0.024)
Δ Primærnæringsmiddel	-0.046 (0.039)	-0.025 (0.042)	0.846** (0.034)	-0.027 (0.041)
Δ Håndverkere	0.007 (0.026)	-0.064** (0.031)	0.021 (0.030)	-0.045 (0.028)
Δ Prosess- og maskinoperatører	-0.025 (0.028)	-0.070** (0.033)	0.007 (0.032)	-0.034 (0.028)
Δ Yrker uten krav til utdanning	-0.025 (0.024)	-0.059* (0.034)	-0.001 (0.031)	-0.046 (0.030)

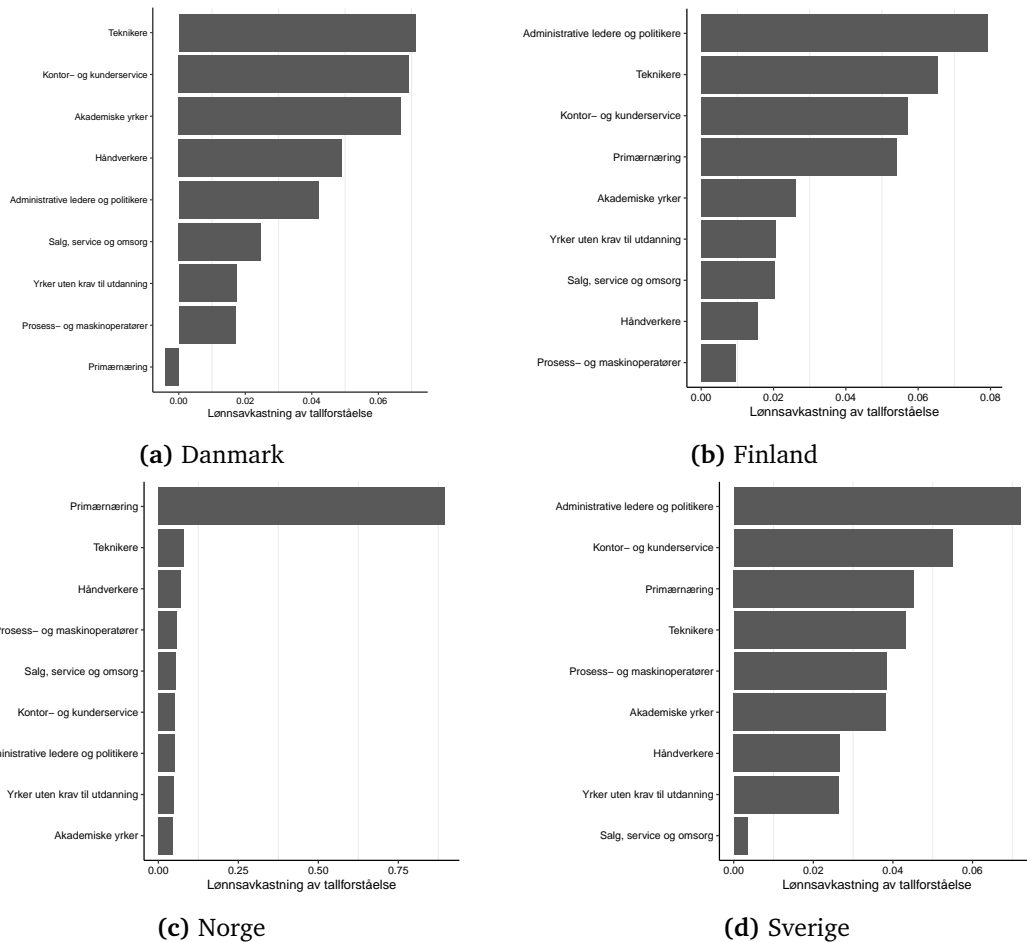
Note: Regresjonskoeffisientene er estimert ved å tilpasse en OLS-regresjon for log-lønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, indikator for jobbsektor og en interaksjonseffekt mellom tallforståelse og jobbsektor. Hver linje prefigert med en Δ rapporterer den marginale forskjellen i avkastningen av tallforståelse for gitt sektor, som estimert av interaksjonseffekten mellom tallforståelse og indikatoren for gitt sektor. Tilhørende standardavvik står i parentes under. Bare respondenter eldre enn 25 år er inkludert.

* $p < 0.10$

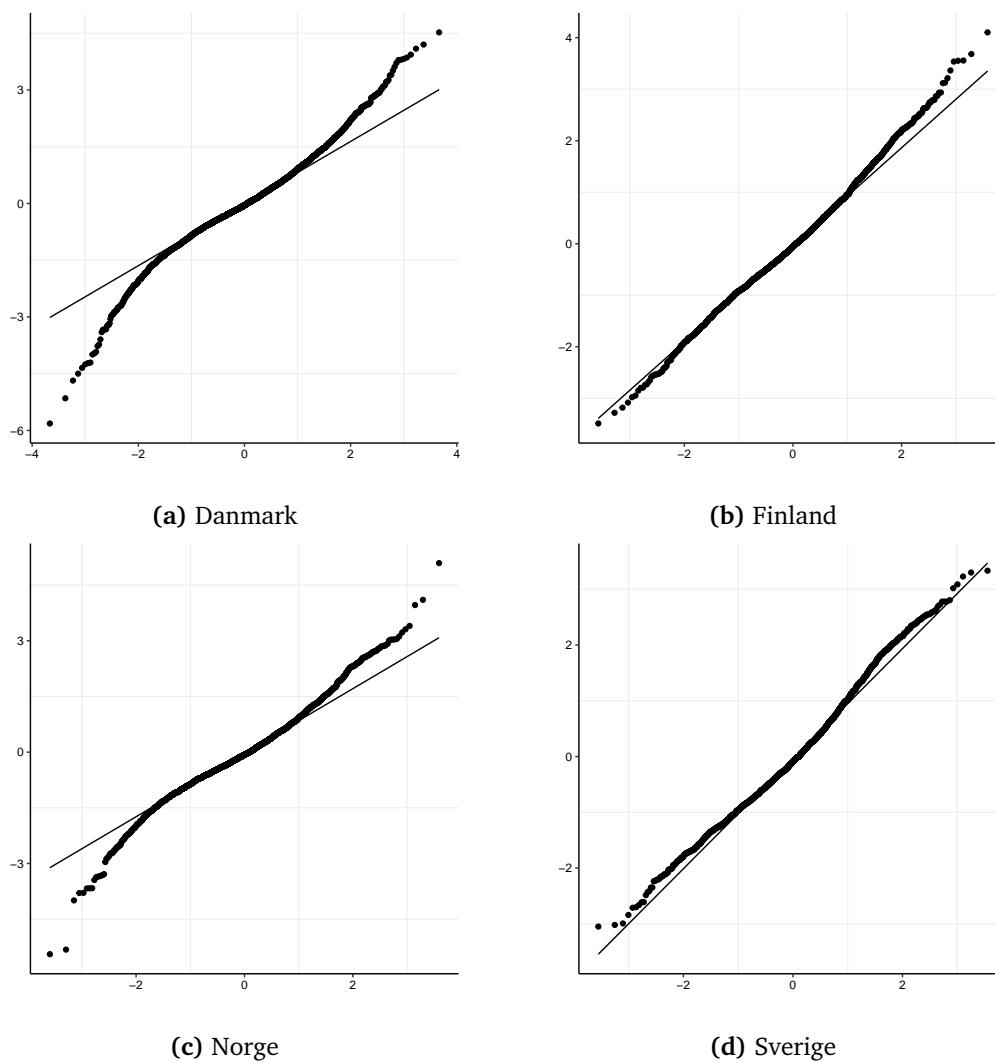
** $p < 0.05$

*** $p < 0.01$

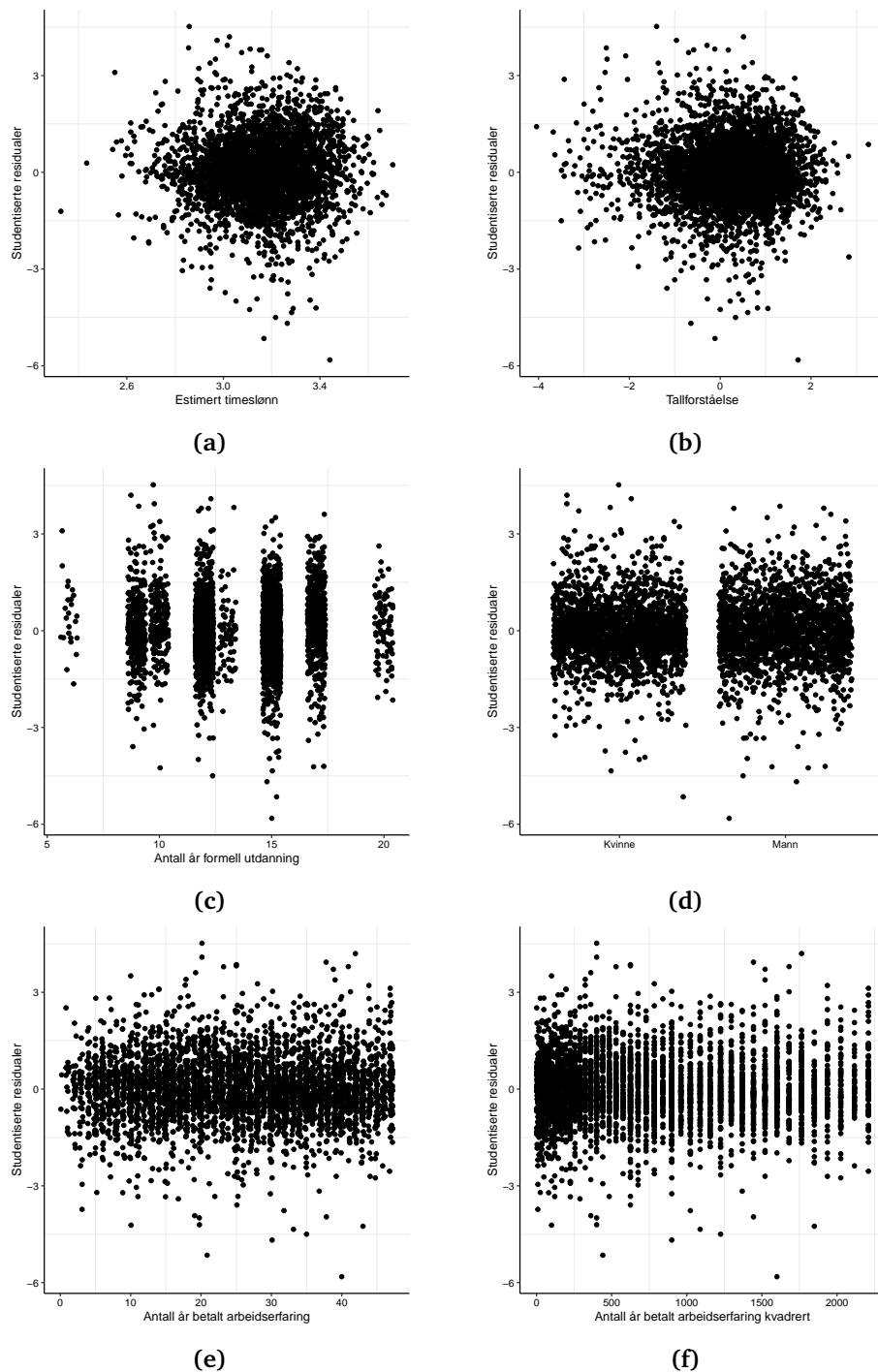
Data fra: PIAAC.



Figur A1: Avkastning av kognitive ferdigheter innen tallforståelse for ulike arbeidssektorer, definert i henhold til 2008 ISCO, i nordiske land. *Note:* Regresjonskoeffisientene er estimert i en OLS-regresjon av log-lønn på tallforståelse (standardisert til standardavvik 1 innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring, antall år arbeidserfaring kvadrert, arbeidssektor og en interaksjonseffekt mellom tallforståelse og arbeidssektor. Bare respondenter eldre enn 25 år er inkludert. Resultatene replikerer tilpasningen fra tabell A4. *Data fra:* PIAAC.



Figur A2: QQ-plot av de studentiserte residualene fra tilpasningene i tabell 2. *Note:* Residualene er beregnet for en OLS-regresjon for hvert land av log-timeslønn på kognitive ferdigheter innen tallforståelse (standardisert til standardavvik innad i hvert land), antall år formell utdanning, kjønn, antall år arbeidserfaring og antall år arbeidserfaring kvadrert. Bare respondenter eldre enn 25 år er inkludert. *Data fra:* PIAAC.



Figur A3: Diagnostisk plott av studentiserte residualer for Danmark-tilpasningen i tabell 2 mot (a) estimert log-timeslønn, (b) kognitive ferdigheter innen tallforståelse (standardisert til standardavvik 1 innad i hvert land), (c) antall år formell utdanning, (d) kjønn, (e) antall år arbeidserfaring og (f) antall år arbeidserfaring kvadrert. *Note:* I figur (c)–(e) er punktene tilført en liten horisontal spredning. Bare respondenter eldre enn 25 år er inkludert. *Data fra:* PIAAC.

