

Abstract

Cognitive ability has an significant impact on several socioeconomic outcomes such as health, educational performance, consumption patterns etc. Education length have been shown to have a positive impact on cognitive ability. However, it remains unknown, whether social class affects the impact of education on cognitive ability.

In this paper I provide new evidence suggesting that the causal effect of education on cognition vary in the population across individuals with different social backgrounds. I use SHARE data and schooling reforms in several European countries as instruments for educational attainment. I find that an extra year of education has a significant protective effect on cognition; the cognition score increases by 0,24 standard deviation on average.

I find that heterogeneity and selection play a part in schooling decisions, and that the reforms mainly change the schooling decisions of individuals from the low socioeconomic class.

Using the data available in SHARELIFE on early life conditions of the respondents such as socioeconomic status, health and performance at school, I identify that individuals from a higher socioeconomic class receives a larger marginal return on cognition from education than average, whereas individuals form a lower socioeconomic class receives a smaller marginal return on cognition from education than average.

This implies that the impact of education on cognition is strongly affected by ones social class.

Indholdsfortegnelse

1	Introduktion	4
2	Teori	5
2.1	Familiens betydning for børns kognitive evner	5
3	Den empiriske model	8
3.1	OLS-estimation	8
3.2	IV-estimation	8
4	Data	9
4.1	Uddannelsesreformer som naturligt instrument	10
4.2	Mål for kognitive evner	11
4.3	De forklarende variable	11
5	Empirisk analyse	13
5.1	Effekten af uddannelsesreformer på uddannelseslængde	13
5.2	Effekten af uddannelseslængde på kognitive evner	16
6	Robusthedsanalyse	20
6.1	Metodeændringer og tidsvarierende faktorer	20
6.2	Tilfældig stikprøve	20
6.3	Effekten af uddannelseslængde på kognitive evner som kvadratisk funktion	22
7	Konklusion	23
8	Diskussion	23
	Referencer	25
A	Appendiks: Udeladte figurer	28
A.1	Den kumulative fordeling af kognitive evner i stikprøven	28
A.2	Histogram over uddannelseslængden	28
A.3	De tyske reformers betydning for uddannelseslængden	29
B	Appendiks: Anvendte tests	29
B.1	Endogenitets test	29
B.2	Breusch-Pagan test	30

C	Appendiks: Effekten af uddannelse på kognitive evner - fitted values.	30
C.1	De gennemsnitlige kognitive evner fordelt betinget af uddannelseslængde og barndomsvariable. Fitted values	30

1 Introduktion

Niveauet af vores kognitive evner har en stor betydning for flere socioøkonomiske udfald. Tidligere studier har vist, at niveauet af kognitive evner blandt andet påvirker vores valg af uddannelseslængde [Nagoshi et al., 1993], sundhed [Arendt, 2005], arbejdes performance [Hunter, 1986], forbrugsmønstre [Ward et al., 1997] samt risiko- og tidspræferencer [Burks et al., 2009]. Den signifikante effekt af kognitive evner på socioøkonomiske udfald gør det relevant at undersøge, hvad der bestemmer niveauet af kognitive evner.

Den eksisterende litteratur har primært haft fokus på sammenhængen mellem uddannelseslængde og kognitive evner. [Falch et al., 2011], [Banks et al., 2012], [Crespo et al., 2014] finder alle, at der eksisterer en signifikant positiv sammenhæng mellem uddannelseslængde og kognitive evner. Der har derimod været et mindre fokus på sammenhængen mellem socioøkonomisk klasse og kognitive evner. Ifølge [Haveman & Wolfe, 1995] vil forældre fra en højere socioøkonomisk klasse investere mere tid og flere ressourcer i deres børn. Den øgede investering vil medføre en forskel i børnenes kognitive evner før de starter i skole. [Ermisch et al., 2012] argumenterer for, at forskellen i børnenes kognitive evner øges gennem skoletiden på grund af det endogene valg af videregående uddannelse. Flere forfattere har leveret empiriske beviser for sammenhængen mellem socioøkonomisk klasse og performance i skolen. [Feinstein, 2003] finder, at elever, der kommer fra en højere socioøkonomisk klasse, har lettere ved at indhente sine klassekammerater fagligt efter fravær. [Decker et al., 2017] finder, at børn fra en højere socioøkonomisk klasse klarer sig bedre i intelligenthæftstests. Disse resultater giver anledning til at undersøge, om effekten af uddannelseslængde på kognitive evner varierer i populationen på baggrund af socioøkonomiske karakteristika.

Med afsæt i den økonomiske model opstillet i "*Human Capital and the Rise and Fall of Families*" [Becker & Tomes, 1985] vil jeg i den følgende analyse undersøge, om effekten af uddannelseslængde på kognitive evner, hos personer der er 50 år eller ældre, varierer på baggrund af socioøkonomiske karakteristika. Analysen er inspireret af resultaterne fra [Crespo et al., 2014], der finder en signifikant positiv effekt af uddannelseslængde på kognitive evner, men ingen signifikante beviser for, at denne effekt varierer på baggrund af individets barndomsmiljø. Barndomsmiljø anvendes som proxy for den socioøkonomiske klasse, som individet er opvokset under.

For at undersøge om effekten af uddannelse på kognitive evner varierer på baggrund af socioøkonomiske karakteristika, anvender jeg instrumental estimation (IV). Effekten af uddannelse på kognitive evner isoleres ved at anvende uddannelsesreformer, der forlænger den obligatoriske uddannelseslængde, som instrument. Uddannelsesreformerne kan anvendes som instrument, da de forudsiger en eksogen variation i individers valg af uddannelseslængde. Derved har jeg kontrolleret for, at et individs valg af uddannelseslængden

kan være endogen bestemt af uobserverbare faktorer som genetik og barndomsmiljø. Ved at anvende et interaktionsled mellem uddannelseslængde og barndomsvariable isoleres effekten, fra uddannelse på kognitive evner, for grupper med specifikke socioøkonomiske karakteristika. Hvis parameteren for interaktionsleddet indgår signifikant, varierer effekten af uddannelse på kognitive evner i populationen på baggrund af socioøkonomiske klasser.

Jeg anvender den tværnationale paneldatabase: "Survey of Health, Ageing and Retirement in Europe" (SHARE), der indeholder mikrodata om helbred og socioøkonomiske status for 140.000 tilfældigt udvalgte personer, som er 50 år eller ældre, fra 26 europæiske lande. Jeg har heraf sammensat en stikprøve, der indeholder barndomsvariable fra Wave 3 og 7¹ og grunddata fra easySHARE², for de individer der blev påvirket af en uddannelsesreform. De kognitive evner angives i følgende analyse som en samlet score af henholdsvis en hukommelsestest og en fluencytest (en test der måler individets evne til at tænke i semantiske skemaer). Den udvalgte stikprøve anvendt i analysen indeholder information om 8.907 individer.

Jeg finder, at et ekstra års uddannelse har en signifikant og positiv effekt på kognitive evner uanset socioøkonomisk baggrund, men at niveauet af effekten er betinget af socioøkonomiske karakteristika. Individer, der er opvokset i et hjem med færre end 25 bøger, forbedrer deres kognitive evner med 0,021 standardafvigelse per års uddannelse. Individer, der er opvokset i hjem med det maksimale antal brugsgoder,³ forbedrer derimod deres kognitive evner med 0,278 standard afvigelse per års uddannelse. Dette tyder på, at socioøkonomiske klasse påvirker effekten af uddannelse på kognitive evner.

2 Teori

2.1 Familiens betydning for børns kognitive evner

Den økonomiske model opstillet af [Becker & Tomes, 1985] beskriver hvordan lønninger, aktiver og forbrug nedarves mellem generationer. I denne analyse antages det, at kognitive evner nedarves på tilsvarende vis. Denne antagelse kan retfærdiggøres af resultatet i [Tucker-Drop et al., 2014], der viser, at kognitive evner hovedsageligt nedarves genetisk, men påvirkes af barndomsmiljø og uddannelse. I følgende afsnit opstiller jeg en teoretisk model, der beskriver hvorfor børn fra en højere socioøkonomisk klasse får en større effekt af uddannelse. Modellen er inspireret af Becker & Tomes.

¹Wave 3 og 7 adskiller sig fra de øvrige spørgerunder (wave's), ved at indeholde mikrodata om individets barndom. Kilde: [SHARE release guide 7.0.0, 2019].

²easySHARE er et simplificeret datasæt, der samler mikrodata for alle adspurgte individer i de syv spørgerunder. Kilde: [SHARE release guide 7.0.0, 2019].

³Antal brugsgoder=5.

Nogle børn har den fordel, at de er født ind i familier hvor forældrene har en høj intelligens og et større fokus på børnenes læring. Det antages, at sådanne genetiske og kulturelle fordele videregives fra forældre til børn gennem DNA og familiekultur. Denne intergenerationelle overdragelse af familiespecifikke evner kan beskrives med ligningen:

$$F_{it} = \alpha_t + hF_{i(t-1)} + v_{it} \quad (1)$$

Hvor F_{it} er en vektor af familiespecifikke evner i den t 'ende generation. h er en vektor af konstanter, der angiver hvor stor en andel af evnerne, der nedarves mellem generationerne. α_t angiver de samfundsspecifikke evner og v_{it} er et stokastisk fejld.

Forældre videregiver ikke kun familiespecifikke evner til børn, men investerer også i børnenes humankapital ved at dække deres omkostninger til eksempelvis skole, bøger, helbred mm. Følgende ligning kan anvendes til at beskrive, hvordan forældre influerer børnenes udvikling af kognitive evner:

$$CA_t = \gamma H_t + l_t \quad (2)$$

Hvor CA_t angiver barnets kognitive evner som voksen. H_t angiver barnets akkumulerede humankapital. γ angiver effekten af en enhed humankapital på kognitive evner og l angiver held. Da den sande værdi af γ er ukendt antages det at $\gamma = 1$.

Det antages, at et individs humankapital er bestemt af nedarvede familiespecifikke evner (F_{it}), forældres investeringer i børnenes læring (X_{t-1}), og samfundets investeringer i uddannelse (s_{t-1}). Dette kan beskrives ved:

$$H_t = \psi(F_{it}, X_{t-1}, s_{t-1}) \quad (3)$$

Det marginale afkast af forældrenes investering på børnenes kognitive evner kan beskrives ved:

$$\frac{\partial CA_t}{\partial X_{t-1}} = \frac{\partial H_t}{\partial X_{t-1}} = \psi_X = 1 + r(F_{it}, X_{t-1}, s_{t-1}) \quad (4)$$

Det ses heraf, at fordelagtige familiespecifikke evner (F_{it}) øger det marginale afkast af både familiens og det offentlige investeringer i barnets humankapital. Det betyder, at børn der har arvet genetiske og kulturelle fordele fra deres forældre modtager et relativt større humankapitalafkast, når familien og/eller det offentlige investerer i dem. Dette kan ses ved:

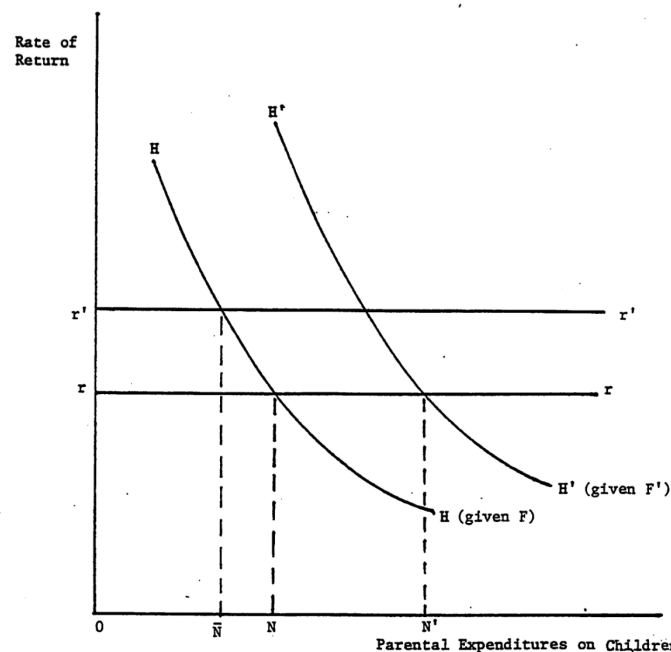
$$\frac{\partial^2 H_t}{\partial y_{t-1} \partial E_{it}^i} = \frac{\partial r}{\partial E_{it}^i} = \psi_{yE} > 0 \quad , \text{ hvor } y = X, s \quad (5)$$

Det antages, at det marginale afkast af humankapital på kognitive evner aftager med mængden af investeringer i barnet [Becker & Tomes, 1985]. De aftagene efterspørgselsskurver H og H' er illustreret i Figur 1. H' ligger på et højere niveau end H , fordi de

familiespecifikke evner F' er relativt fordelagtige i forhold til F . I denne model antages det, at forældrenes omkostning ved at investere i deres børn (r) er den samme uanset hvilken socioøkonomisk klasse man tilhøre. Dette er en streng antagelse, da den individuelle omkostning kan variere på baggrund af socioøkonomisk klasse [Ermisch et al., 2012].

Forældrenes optimale investering i humankapital er i Figur 1 angivet ved skæringen mellem den vandrette udbudskurve (r) og de aftagende efterspørgselskurver (H og H'). Det ses heraf, at børn med bedre familiespecifikke evner ($F' > F$) modtager mere investering fra deres forældre ($N' > N$). De fordelagtige familiespecifikke evner, og den øgede investering fra forældrene, medfører, at disse børn opbygger mere humankapital gennem barndommen. Ifølge ligning (2) betyder det, at børn med en genetisk og kulturel initialfordel opnår et højere niveau af kognitive evner som voksen.

Figur 1: Det marginale afkast på kognitive evner givet forældrenes investering i barnet. [Becker & Tomes, 1985]



Hvis denne model er retvisende, da vil børn der er født med fordelagtige familiespecifikke evner få et større humankapital afkast pr. enhed der investeres i dem. Anses uddannelse som en investering, da vil afkastet af uddannelse på kognitive evner være større, for de børn der er født med fordelagtige familiespecifikke evner. Det er denne hypotese der undersøges i følgende analyse.

Om den højere socioøkonomisk klasse investere mere tid og ressourcer i deres børn, og om de har relativt fordelagtige familiespecifikke evner, undersøges ikke i denne analyse, men antages på baggrund af [Haveman & Wolfe, 1995].

3 Den empiriske model

For at undersøge om uddannelseslængdens effekt på kognitive evner betinget af socioøkonomiske karakteristika, vil der i den empiriske analyse blive fokuseret på nedenstående multiple lineære model (Second Stage):

$$CA_i = \gamma_1 + \gamma_2 E_i + \gamma'_{31} X_{1i} + \gamma'_{32} X_{2i} + \gamma'_4 E_i X_{2i} + v_i \quad (\text{Second Stage})$$

CA_i angiver individ i 's kognitive evner. E_i angiver uddannelseslængden. X_{1i} er en vektor af baggrundsvARIABLE, der beskriver køn, alder, land, og landespecifikke kohorte trends. X_{2i} er en vektor af barndomsvariable, der beskriver relative evner og barndomsmiljø. Interaktionsleddet $E_i X_{2i}$ isolerer den ekstra effekt, som individer med bestemte barndomsvariable modtager ved et års længere uddannelse på kognitive evner. v_i er et stokastisk fejllid. Koefficienten γ_2 angiver den gennemsnitlige effekt af et ekstra års uddannelse på kognitive evner, mens koefficienten $\psi = \gamma_2 + \gamma_4$ angiver effekten af et års ekstra uddannelse på kognitive evner for individer med specifikke barndomsvariable. Effekten af uddannelse på kognitive evner varierer på baggrund af socioøkonomisk karakteristika hvis $\gamma_4 \neq 0$.

I analysen anvendes både OLS og IV som estimationsmetode. Ved at bruge begge metoder er det muligt at undersøge styrken af bias. Nedenfor beskrives de to estimationsmetoder.

3.1 OLS-estimation

OLS er den bedste lineære unbiased estimator, hvis Gauss-Markov antagelserne er overholdt. Gauss-Markov antagelserne⁴ er (1) at modellen er lineær i parametre, (2) at stikprøven er tilfældigt udvalgt, (3) at der ikke er perfekt kollinearitet og (4) at den betingede middelværdi af fejllidet er nul, $E(v_i | E_i, X_{1i}, X_{2i}, E_i X_{2i}) = 0$.

For at modellen har den mindst mulige varians, skal modellen have homoskedastiske fejllid. Fejlledet skal derved have samme varians uanset værdien af de uafhængige variable, $Var(v_i | E_i, X_{1i}, X_{2i}, E_i X_{2i}) = \sigma^2$. Dette sikre, at de sædvanlige t-test, F-test og konfidensintervaller er valide. Hvis modellen har heteroskedastiske fejllid, kan det medføre en misvisende fortolkning af inferens.⁵

3.2 IV-estimation

På baggrund af [Crespo et al., 2014] forventes det, at et individs valg af uddannelse er endogent bestemt af uobserverbare faktorer som genetik og barndomsmiljø. Det er derved

⁴Gauss Markov betingelserne er beskrevet i [Wooldridge, 2016] side 73-77.

⁵Heteroskedastiske fejllid beskrevet i [Wooldridge, 2016] side 81-83.

en bestemt type af individer, der på baggrund af socioøkonomisk karakteristika, 'self-selector' ind i en lang uddannelse. Den fjerde af Gauss-Markov antagelserne er derved ikke overholdt, da $E(E_i, v_i) \neq 0$. Hvis Second Stage estimeres ved brug af OLS vil estimerterne derfor blive biased.

IV anvendes i de situationer, hvor en eller flere af de forklarende variable er endogene. Ved at introducere et validt instrument (z_i), som i denne analyse udgøres af uddannelsesreformer, kan IV producere unbiased estimerter. Uddannelsesreformerne kan anvendes som et validt instrument, da de (1) er relevante, det vil sige, at reformerne forudsiger en del af variationen i den endogene variabel (uddannelseslængden). $cov(z, E_i) \neq 0$. Og (2) reformerne er eksogene, det vil sige, at reformerne kun påvirker kognitive evner igennem uddannelseslængden. Instrumentet er derved ukorreleret med Second Stage fejllede, $cov(z, v_i) = 0$.

Følgende First Stage ligning anvendes til at undersøge uddannelsesreformernes effekt på uddannelseslængden.

$$E_i = \delta_1 + \delta_2 z_i + \delta'_{31} X_{1i} + \delta'_{32} X_{2i} + \delta'_4 X_{2i} z_i + \epsilon_i \quad (\text{First Stage})$$

Instrumentet (z_i) angiver individ i 's obligatoriske uddannelseslængde. ϵ_i er et fejlede der er ukorreleret med uddannelseslængde og kontrolvariable.

I First Stage regresseres uddannelseslængden på den obligatoriske uddannelseslængde og eksogene baggrundsvariable. Da den obligatoriske uddannelseslængde indgår eksogent, opdeles variation af uddannelseslængden (E_i) i to dele. En del der ikke er korreleret med Second Stage fejllede, $cov(\hat{E}_i, v_i) = 0$, og en del der kan være korreleret med Second Stage fejllede, $cov(\epsilon_i, v_i) \neq 0$. Den estimerede værdi af uddannelseslængden (\hat{E}_i) fra First Stage modellen er derved eksogen, og kan indsættes i stedet for den faktiske uddannelseslængde (E_i) i Second Stage. Denne fremgangsmetode kaldes for Two-Stage least squares (2SLS).

⁶ Ved brug af IV kan der opnås konsistente og unbiased estimerter af γ_2 og ψ

4 Data

I den empiriske analyse anvendes den tværsnittede paneldatabase: Survey of Health, Aging and Retirement in Europe (SHARE), der består af 7 spørgerunder (Wave 1-7) som er indsamlet i perioden 2004-2019. Det anvendte datasæt sammenkobler barndomsvariable fra Wave 3 og 7⁷ med grunddata fra easySHARE⁸. easySHARE indeholder, udover

⁶Afsnittet er baseret på [Hove., 2016] Forelæsningsnotat: "Instrumental Variables Estimation"

⁷Wave 3 og 7 adskiller sig fra de øvrige spørgerunder (wave's) ved at indeholde mikrodata om individets barndom. Kilde: [SHARE release guide 7.0.0, 2019]

⁸easySHARE er et simplificeret datasæt der samler mikrodata for alle adspurgte individer i de syv spørgerunder. Kilde: [SHARE release guide 7.0.0, 2019]

grunddata, også eksperimentelle data for respondenternes kognitive evner. Ved at gøre brug af et unikt respondentnummer sammenkodes kognitive evner, barndomsvariable og grunddata på individniveau. Barndomsvariable anvendes som proxy for individets socioøkonomiske status. Datasættet er konstrueret således, at hvert individ kun indgår en gang med deres baseline interview. Observationerne i stikprøven er således uafhængige af hinanden.

4.1 Uddannelsesreformer som naturligt instrument

En uddannelsesreform, der forlænger den obligatoriske uddannelseslængde, kan anses som et naturligt instrument. Reformen skaber en eksogen variation i uddannelseslængden på begge sider af den første berørte kohorte (elevårgang). Det antages, at kohorterne før og efter en reformen er relativt ens i forhold til kognitive evner og socioøkonomiske karakteristika. Det vil sige, at der ikke er en signifikant forskel på de elevårgange, der henholdsvis berøres og ikke berøres af en uddannelsesreform. En uddannelsesreform sætter en nedre grænse for uddannelseslængden. Givet at der ikke er forskel på kohorterne vil nogle individer, efter reformens implementering, blive tvunget til at tage en længere uddannelse end de ellers havde tiltænkt. Det må antages, at et individs socioøkonomisk status ikke er bestemt af uddannelsesreformerne, og at uddannelsesreformerne derfor alene påvirker kognitive evner gennem uddannelseslængde. Uddannelsesreformerne kan derfor anvendes som et validt instrument i en IV estimation.

I det anvendte datasæt har jeg udvalgt de SHARE respondenter, som har været berørt af en uddannelsesreform og har besvaret enten Wave 3 eller 7 om barndomsmiljø. Derfor indgår kun de lande, hvor der er registreret en uddannelsesreform, der har forlænget den obligatoriske uddannelseslængde, for de kohorter der er repræsenteret i SHARE-populationen. Det betyder at analysen anvender uddannelsesreformer implementeret i perioden 1950-1969 for landene Østrig, Sverige, Holland, Italien, Frankrig, Danmark og Tyskland⁹ [Brunello al., 2009], [Murtin & Viarengo, 2008], [Arendt, 2005]. For hvert individ er den obligatoriske uddannelseslængde beregnet, som den alder man minimum skal have for at droppe ud af skolen fratrasket den alder man har ved skolestart i individets hjemland. Det antages derved, at de adspurgte individer har studeret i samme land, som de er bosiddende i. For at minimere potentielle effekter af udeladte variable, der både korrelerer med uddannelsestrends og niveauet af kognitive evner, begrænser jeg stikprøven til kun at indeholde kohorter 7 år før og efter den første berørte kohorte. Tabel 1 lister de reformer og kohorter der er medtaget i den empiriske analyse.

⁹Tyskland er eneste land der indfører reformerne på regionsniveau.

Tabel 1: Uddannelsesreformer der påvirker den obligatoriske uddannelseslængde

Land	Berørte kohorter	YCE^* (før reform)	YCE^* (efter reform)
Østrig ¹	1940 - 1953	8	9
Tyskland(<i>Schleswig – Holstein</i>) ¹	1934 - 1947	8	9
Tyskland(<i>Hamburg</i>) ¹	1927 - 1940	8	9
Tyskland(<i>Niedersachsen</i>) ¹	1940 - 1953	8	9
Tyskland(<i>Bremen</i>) ¹	1936 - 1949	8	9
Tyskland(<i>Nordrhein – Westphalia</i>) ¹	1946 - 1959	8	9
Tyskland(<i>Hessen</i>) ¹	1946 - 1959	8	9
Tyskland(<i>Rheinland – pfalz</i>) ¹	1946 - 1959	8	9
Tyskland(<i>Baden-Württemberg</i>) ¹	1946 - 1959	8	9
Tyskland(<i>Bayern</i>) ¹	1948 - 1961	8	9
Tyskland(<i>Saarland</i>) ¹	1942 - 1955	8	9
Sverige ¹	1943 - 1956	8	9
Frankrig ¹	1946 - 1959	8	10
Holland ²	1930 - 1943	6	8
Italien ²	1942 - 1955	5	8
Danmark ³	1938 - 1951	5	7

*YCE: Obligatorisk uddannelseslængde (år), ¹Kilde: [Brunello et al., 2009],

²Kilde: [Murtin et al., 2008], ³Kilde: [Arendt, 2005]

4.2 Mål for kognitive evner

De kognitive evner er målt som en samlet score af en hukommelsestest og en fluencytest. I hukommelsestesten blev respondenter bedt om at lytte til 10 almindelige ord, hvorefter han skulle gengive ordene lige med det samme og igen 5-10 minutter efter. I fluencytesten blev respondenter bedt om at nævne så mange forskellige dyr som muligt indenfor et minut.¹⁰ Den samlede score af testene anvendes som mål for kognitive evner. For at sikre, at skalaen er konsistent over tid, er scoren standardiseret ved at fratrække gennemsnittet og dividere med standardfejlen. Den kumulativ fordeling af kognitive evner er vist i appendiks A.1. Det ses heraf at niveauet af de kognitive evner er tæt på normalfordelt.

4.3 De forklarende variable

Uddannelseslængden: Alle respondenter opgiver deres uddannelseslængde (E_i) i deres baseline-interview. Da uddannelseslængden er selvrapporteret, er der stor sandsynlighed for, at der indgår målefejl i denne variabel. Hvis målefejlen er tilfældig, da vil de estimerede parametre γ_2 og γ_4 være biased mod nul når de estimeres med OLS. Ved brug af et validt instrument i IV-estimationen tages der forbehold for tilfældige målefejl. Hvis målefejlen derimod ikke er tilfældig, således at det er nogle bestemte individer der systematisk overdriver eller underdriver deres uddannelseslængde, da vil estimatet være bi-

¹⁰Kilde: [SHARE release guide 7.0.0, 2019]

ased i både OLS og IV. Fortolkningen af inferens kan i dette tilfælde være misvisende.¹¹ Fordelingen af uddannelseslængden i stikprøven kan ses i appendiks A.2. Det ses heraf, at der er sammenhobninger af observationer ved uddannelseslængden 6 år, 8 år og 16 år. Dette skyldes sandsynligvis den bunde længde af henholdsvis grundskole, overbygningen og gymnasial uddannelse, men kan også skyldes systematiske målefejl.

Baggrundvariable: Respondenternes køn, alder, land og landespecifikke kohortetrends indgår i vektoren (X_{1i}) . De landespecifikke kohortetrends er beregnet som en interaktion mellem respondentens kvadrerede alder og en lande-dummy. Denne er medtaget, da effekten af uddannelse på kognitive evner kan variere afhængig af kohorte og land.

Barndomsvariable: Da SHARE ikke indeholder data om forældrenes indkomst eller uddannelse, anvendes barndomsvariablene som proxy for individets socioøkonomisk status, helbred og relative evner. Dette er faktorer der kan påvirke uddannelsens effekt på kognitive evner. De barndomsvariable der indgår i vektoren (X_{2i}) er: (1) *udkantsområde*, som er en dummy for, at individets barndomshjem var placeret i et udkantsområde. (2) *værelser ≤ 2* , som er en dummy for, at der maksimalt var 2 værelser i barndomshjemmet. (3) *bøger ≤ 25* , som er en dummy for, at der maksimalt var 25 bøger i barndomshjemmet. (4) *Primære forsørgers job*, som er en dummy for, at den primære forsørger havde et lavtlønnet job¹². (5) *brugsgoder*, som er en variabel der angiver antallet af brugsgoder i barndomshjemmet¹³. (6) *Dårlig elev*, som er en dummy for, at individet klarede sig relativt dårligere i matematik og sprog end sin årgang som 10-årig. (7) *dårligt f. helbred*, som er en dummy for, at individet vurderede, at han havde et relativt dårligt fysisk helbred som barn. (8) *dårligt m. helbred*, som er en dummy for, at individet vurderede, at han havde et relativt dårligt mentalt helbred som barn. Barndomsvariablene er udvalgt på baggrund af [Crespo et al., 2014].

Tabel 2 angiver de gennemsnitlige værdier, for de variable der er anvendt i analysen.

¹¹[Wooldridge, 2016] side 293-295

¹²Den primære forsørger var beskæftiget som håndværker eller beslægtet fagarbejde, anlægs- eller maskinoperatør, ansat i hæren eller midlertidigt ansat.

¹³Brugsgoder: installeret bad, adgang til koldt vand, adgang til varmt vand, installeret toilet samt central varme.

Tabel 2: Beskrivende statistik

Variabel	(Gennemsnit)
Kognitive evner (ikke standardiseret)	31,99 point
Uddannelseslængde	12,20 år
Obligatorisk uddannelseslængde	7,51 år
Alder	59,97 år
Kvinde	53,99 %
Reistreret par	43,19%
Udkantsområde	31,08 %
Værelser ≤ 2	20,88 %
Bøger ≤ 25	54,25 %
Primære forsørgers job	42,12 %
Brugsgoder	2,88 stk.
Dårlig elev	49,85 %
Dårligt f. helbred	9,39 %
Dårligt m. helbred	1,74 %
Østrig	11,24 %
Tyskland	3,94 %
Sverige	18,60 %
Frankrig	18,42 %
Holland	10,70 %
Italien	20,83 %
Danmark	16,27%
Stikprøve størrelse	8.907

5 Empirisk analyse

I dette afsnit undersøges det, om effekten af uddannelseslængde på kognitive evner er betinget af barndomsmiljø og socioøkonomisk status. Først estimeres First Stage, for at undersøge om uddannelsesreformer kan anvendes som valide instrumenter i Second Stage. Derefter estimeres Second Stage, for at undersøge om effekten af uddannelse på kognitive evner varierer i populationen på baggrund af socioøkonomiske karakteristika.

5.1 Effekten af uddannelsesreformer på uddannelseslængde

For at finde effekten af uddannelsesreformer på uddannelseslængde estimeres First Stage med forskellige specifikationer. Resultatet er præsenteret i Tabel 3.

Tabel 3: Effekten af reformer på uddannelseslængde (First Stage)

	OLS (1) E_i	OLS (2) E_i	OLS (3) E_i
Oobligatorisk uddannelseslængde (z_i)	0.759*** (0.048)	0.556*** (0.043)	0.520*** (0.085)
Kvinde	-0.561*** (0.074)	-0.627*** (0.067)	-0.616*** (0.067)
Alder	0.054** (0.021)	0.074*** (0.018)	0.059*** (0.018)
Udkantsområde		-0.592*** (0.073)	-1.320*** (0.375)
Værelser ≤ 2		-0.272*** (0.085)	-0.412 (0.445)
Bøger ≤ 25		-1.538*** (0.079)	-3.150*** (0.440)
Primære forsørgers job		-0.574*** (0.067)	-0.211 (0.359)
Brugsgoder		0.357*** (0.023)	0.823*** (0.124)
Dårlig elev		-1.423*** (0.068)	-2.039*** (0.363)
Dårligt f. helbred		-0.094 (0.110)	0.650 (0.598)
dårligt m. helbred		-0.069 (0.254)	0.420 (1.293)
z_i x udkantsområde			0.098** (0.048)
z_i x værelser ≤ 2			0.020 (0.058)
z_i x bøger ≤ 25			0.209*** (0.056)
z_i x primære forsørgers job			-0.048 (0.046)
z_i x brugsgoder			-0.063*** (0.016)
z_i x dårlig elev			0.083* (0.046)
z_i x dårligt f. helbred			-0.098 (0.076)
z_i x dårligt m. helbred			-0.059 (0.165)
Konstant	3.409** (1.415)	5.295*** (1.223)	6.482*** (1.336)

Note: Alle regressioner er kontrolleret for land og landespecifikke kohortetrends.

Robuste standardfejl er vist i parenteser. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Kolonne (1) indeholder resultaterne af First stage baseline-specifikationen, hvor der er kontrolleret for køn, land, alder og landespecifikke kohortetrends. Koefficienten for obligatorisk uddannelseslængde er signifikant og viser, at hvis den obligatorisk uddan-

uddannelseslængde stiger med et år, da stiger den faktiske uddannelseslængde med 0,76 år i gennemsnit. Den obligatoriske uddannelseslængde kan derved forklare en relativ stor del af variationen i uddannelseslængden. Med en t -værdi på -0,3, hvilket er signifikant på et 5 pct. signifikansniveau, bekræfter en endogenitetstest at uddannelseslængden er endogen bestemt.¹⁴ For at teste om baseline modellens varians er konstant udføres en Breusch-Pagan test.¹⁵ Testen for baseline modellen giver en teststørrelse på 163. Når denne holdes op imod den kritiske værdi på 26 forkastes nulhypotesen om homoskedastiske fejl. Da stikprøven er tilpas stor vil de robuste standardfejl nærme sig en t -fordeling.¹⁶ Derfor anvendes Whites robuste standardfejl i resten af analysen.

Vektoren X_{2i} , der indeholder barndomsvariable, tilføjes til baseline modellen, for at undersøge om effekten af uddannelsesreformer på uddannelseslængde er betinget af socioøkonomiske karakteristika. Hvis koefficienterne for barndomsvariablene er signifikante, da er effekten af uddannelsesreformer på uddannelseslængde heterogen. Det vil sige, at individers uddannelseslængde bliver påvirket forskelligt af en uddannelsesreform. Resultatet er vist i kolonne (2). Det ses heraf, at alle barndomsvariablene, på nær barnets fysiske og mentale helbred, indgår på et 1 pct. signifikansniveau. Fortegnene på koefficienterne er som forventet for alle barndomsvariable. Det ses at individer der kommer fra den lavere socioøkonomiske klasse (målt ved antal værelser, antal bøger og primære forsørgers job), samt de individer der klarer sig relativt dårligere i skolen, generelt tager en kortere uddannelse end gennemsnittet. Den gennemsnitlige effekt af uddannelsesreformer på uddannelseslængde er med disse specifikationer 0,56 år. Hvilket er lavere end ved baseline modellen i kolonne (1). Barndomsvariablene kan derved forklare en del af variationen i uddannelseslængden.

For at teste om uddannelsesreformerne påvirker uddannelseslængden for bestemte grupper af individer tilføjes interaktionsled mellem obligatorisk uddannelseslængde og barndomsvariablene. Resultatet ses af kolonne (3). Det ses heraf, at flere af interaktionsledende indgår signifikant. Særligt interaktionsleddet mellem instrumentet og færre end 25 bøger i barndomshjemmet, samt interaktionsleddet mellem instrumentet og antallet af brugsgoder i barndomshjemmet, har en signifikant betydning. Fortegnende på disse parametre indikerer, at individer, der er opvokset med færre end 25 bøger i barndomshjemmet, i større grad får deres uddannelseslængde forøget som følge af en reform. Derimod får individer, som er opvokset i et hjem med mange brugsgoder, i mindre grad forlænget deres uddannelseslængde som følge af en uddannelsesreform. Fortegnene på interaktionsledende i kolonne (3) viser derved, at der formentlig eksisterer selv-selektion i uddannelseslængde. Individer, der er opvokset i en lavere socioøkonomisk klasse, vælger at tage en kortere uddannelse og responderer derfor i højere grad på en uddannelsesre-

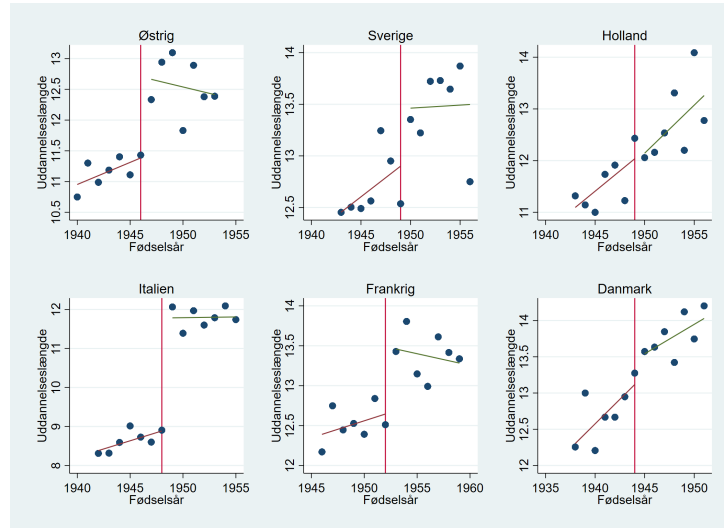
¹⁴Testen er beskrevet i appendiks B.1

¹⁵Testen er beskrevet i appendiks B.2

¹⁶[Wooldridge, 2016], side 244-247.

form. En f-test viser, at interaktionsleddene mellem obligatorisk uddannelseslængde og barndomsvariablene ($z_i X_{2i}$) tilsammen er signifikante på et 1 pct. niveau og har en F-teststørrelse på 9,42¹⁷. Derved er interaktionsledende gode instrumenter til at forudsige uddannelseslængden for grupper med specifikke socioøkonomiske karakteristika.

Figur 2: Uddannelsesreformernes betydning for uddannelseslængden.



Note: Den lodrette røde streg indikere det tidspunkt uddannelsesreformen trådte i kraft. De lineære approksimationer illustrer den gennemsnitlige uddannelseslængde før og efter reformen.

Figur 2 viser den gennemsnitlige uddannelseslængde fordelt på individers fødselsår for landene Østrig, Frankrig, Italien, Holland og Sverige. Den lodrette røde streg indikerer det tidspunkt uddannelsesreformen trådte i kraft. Der ses heraf, at den gennemsnitlige uddannelseslængde blev forøget som følge af en uddannelsesreform, og at der eksisterer landespecifikke trends. Figuren bekræfter derved, at den obligatoriske uddannelseslængde kan anvendes som instrument til at forudsige en eksogen variation i uddannelseslængden i de forskellige lande. Tyskland er ikke medtaget i denne figur, da uddannelsesreformerne i Tyskland blev implementeret på regionsniveau. Da der indgår relativt få tyske respondenter i stikprøven (3,9%) er der meget få observationer i flere af de tyske regioner. Det er derfor svært at finde lineære regionstendenser for de tyske regioner.¹⁸

5.2 Effekten af uddannelseslængde på kognitive evner

Tabel 4 indeholder estimationsresultaterne for Second Stage. Kolonne (1) og (2) er estimeret med OLS, hvorimod kolonne (3) og (4) er estimeret med IV.

¹⁷[Staiger & Stock, 1997] anbefaler at F-teststørrelsen er større end 5,0.

¹⁸Se figur for de tyske reformers betydning for uddannelseslængden i appendix A.3.

Tabel 4: Effekten af uddannelseslængden på kognitive evner (Second Stage)

	OLS (1)	OLS (2)	IV (3)	IV (4)
	kognitive evner ¹	kognitive evner ¹	kognitive evner ¹	kognitive evner ¹
Uddannelseslængde (E_i)	0.062*** (0.003)	0.044*** (0.007)	0.103*** (0.013)	0.242** (0.101)
Alder	-0.018*** (0.004)	-0.015*** (0.004)	-0.016*** (0.004)	-0.017*** (0.006)
Køn	0.097*** (0.017)	0.080*** (0.016)	0.121*** (0.019)	0.190*** (0.049)
E_i x Udkantsområde		0.002 (0.005)		0.000 (0.023)
E_i x værelser ≤ 2		0.007 (0.006)		0.061** (0.030)
E_i x bøger ≤ 25		0.008 (0.006)		-0.221** (0.105)
E_i x primære forsørgers job		-0.004 (0.005)		0.011 (0.027)
E_i x brugsgoder		-0.002 (0.002)		0.036** (0.015)
E_i x dårlig elev		-0.005 (0.005)		0.004 (0.027)
E_i x dårligt f. helbred		0.010 (0.007)		-0.031 (0.054)
E_i x dårligt m. helbred		0.059*** (0.023)		0.230 (0.316)
Konstant	0.479* (0.250)	0.673** (0.261)	-0.182 (0.327)	-2.353* (1.425)
Barndomsvariable	Nej	Ja	Nej	Ja
Instrumenter	-	-	z_i	$z_i, z_i X_{i2}$
F-teststørrelse ²	-	-	164,13	28,27

Note: Alle regressioner er kontrolleret for land og landespecifikke kohortetrends.

Robuste standardfejl er vist i parenteser. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

¹ Kognitive evner er standardiseret. ²F-test af instrumenternes samlede signifikans.

Kolonne (1) viser estimationsresultatet for baseline specifikationen, der alene kontrollerer for køn, land, alder og landespecifikke kohortetrends. Den gennemsnitlige effekt af et ekstra års uddannelse på kognitive evner estimeres til at være 0,062 standard afvigelse, hvilket er signifikant på et 1 pct. signifikansniveau. Dette er en relativt lille effekt af uddannelse på kognitive evner.

Da First Stage regressionen viste, at heterogenitet og selektion spiller en rolle i valget af uddannelseslængde, tilføjes vektoren X_{i2} , der indeholder barndomsvariable, som kontrolvariable i kolonne (2). For at isolere effekten af uddannelseslængde på kognitive evner for bestemte grupper, med specifikke socioøkonomisk karakteristika, tilføjes interaktionssleddene mellem uddannelseslængde og barndomsvariable. Af kolonne (2) ses det, at den gennemsnitlige marginale effekt af uddannelseslængde på kognitive evner er faldet sammenlignet med kolonne (1). Dette er som forventet, da udeladte relevante variable kan give en opadgående bias, hvis barndomsvariablene korrelerer med uddannelseslængden. Udeladte

variable bias medfører en systematiske afvigelse i fejleddet, hvorved estimaterne i kolonne (1) er positivt biased.¹⁹ Af kolonne (2) ses det ligeledes, at det kun er interaktionsleddet mellem uddannelseslængde og barnets mentale helbred som indgår signifikant.

I kolonne (3) er IV estimationsresultaterne af baseline modellen præsenteret. Den obligatoriske uddannelseslængde (z_i) anvendes som eneste instrument for uddannelseslængden. En t-test viser, at instrumentet er signifikant på et 1 pct. niveau. Effekten, af uddannelse på kognitive evner, estimeres til 0,103 standard afvigelse per års uddannelse, hvilket er signifikant på et 1 pct. signifikansniveau. Effekten fra IV baselinemodellen er derved væsentligt større end effekterne estimeret med OLS (kolonne (1) og (2)), men har tilsvarende større robuste standardfejl. Da uddannelseslængde er endogent bestemt, foretrækkes IV estimationsmetoden fremfor OLS.

I kolonne (4) er barndomsvariablene (X_{i2}) og interaktionsleddene ($E_i X_{i2}$) tilføjet til IV estimationen. De anvendte instrumenter er den obligatoriske uddannelseslængde (z_i), samt interaktionsleddene mellem denne og barndomsvariablene ($z_i X_{i2}$). En F-test bekræfter, at instrumenterne tilsammen er signifikante på et 1 pct. niveau. Det estimeres, at et års ekstra uddannelse i gennemsnit forøger kognitive evner med 0,242 standard afvigelse. Dette resultat er signifikant på et 5 pct. signifikansniveau. Flere af interaktionsleddene indgår ligeledes signifikant. Dette bekræfter, at effekten af uddannelseslængde på kognitive evner varierer på baggrund af socioøkonomiske karakteristika.

Den empiriske analyse finder, at IV-estimerne, for effekten af uddannelseslængde på kognitive evner, er større men mindre præcise end estimaterne fra OLS. Dette stemmer overens med resultaterne i [Crespo et al., 2014]. Ifølge [Glymour et al., 2008] skyldes forskellen i estimaterne, at IV angiver 'the local average treatment effect (LATE)', det vil sige effekten for de individer der påvirkes af uddannelsesreformerne. Hvis de individer der påvirkes af reformerne har en større marginal effekt end gennemsnittet, da driver de niveauet af IV estimatet op.

First Stage estimationen viste, at det hovedsageligt er individer fra den lavere socioøkonomiske klasse, der fik forøget deres uddannelseslængde som følge af en uddannelsesreform. Hvis Glymour's hypotese skal holde, skal den lavere socioøkonomiske klasse modtage en større effekt af uddannelse på kognitive evner. Fortegnene på interaktionsleddenes parametre i Second Stage modsiger denne hypotese. Det ses derimod, at individer, der tilhører den lavere socioøkonomiske klasse, har en mindre effekt af uddannelse på kognitive evner.

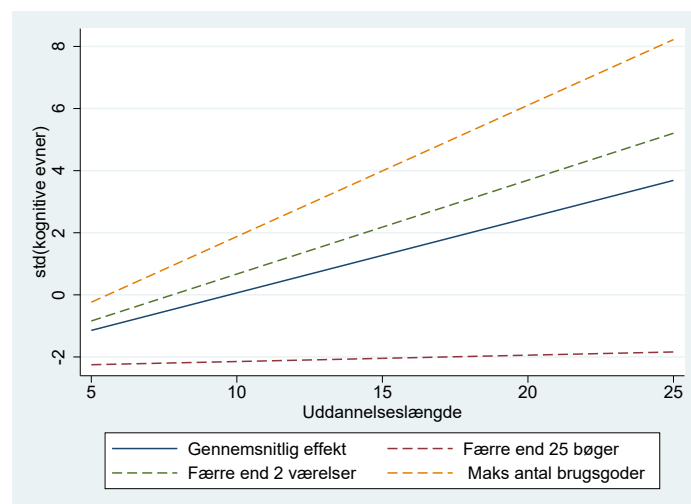
Effekten for de individer, der er opvokset med færre end 25 bøger i barndomshjemmet, er $\psi_{\leq 25 \text{ bøger}} = 0,242 - 0,221 = 0,021$ standard afvigelse per års uddannelse. Dette er væsentligt lavere end den gennemsnitlige effekt $\gamma_2 = 0,242$ standard afvigelse per års

¹⁹[Wooldridge, 2016] side 78-79

uddannelse. Individer, der er opvokset i hjem med det maksimale antal brugsgoder, har derimod en større effekt af uddannelseslængde på kognitive evner der er større end gennemsnittet: $\psi_{brugsgoder} = 0,242 + 0,036 = 0,278$ standard afvigelse per års uddannelse. At IV estimaterne er større end OLS estimaterne kan derved ikke forklares med Glymour's hypotese. Forskellen kan derimod skyldes, at der er målefejl i uddannelseslængden, hvilket giver en bias mod nul i OLS. Ligeledes kan endogeniteten af uddannelseslængden medføre, at OLS underestimerer effekten af uddannelseslængden på kognitive evner.

Den gennemsnitlige effekt af uddannelseslængde på kognitive evner (γ_2) stiger i kolonne (4), når interaktionsleddene og barndomsvariablen tilføjes til modellen. Dette tyder på, at estimatet af effekten i kolonne (3) drives ned af individer der tilhører den lavere socioøkonomiske klasse. Kontrolleres der ikke for socioøkonomisk klasse, da underestimeres den gennemsnitlige effekt af uddannelse på kognitive evner. Modellen anvendt i kolonne (4) er den foretrukne, da socioøkonomisk klasse er medbestemmende for niveauet af effekten.

Figur 3: Uddannelseslængdens effekt på kognitive evner betinget af baggrundsvariable



Figur 3 illustrerer den estimerede marginale effekt af uddannelse på kognitive evner betinget af specifikke barndomsvariable. Estimaterne er fra den foretrukne model (Tabel 4, kolonne (4)). Figuren illustrerer kun effekten for individer opvokset i barndomshjem med det maksimale antal brugsgoder, færre end 25 bøger samt færre end 2 værelser, da disse interaktionsled indgår signifikant i modellen på et 5 pct. signifikansniveau. Det ses heraf, at effekten af uddannelse på kognitive evner er betinget af individets socioøkonomiske karakteristika. Dog afviger barndomsvariablen $værelser \leq 2$ fra denne teori, da færre end 2 værelser i barndomshjemmet giver en positiv effekt af uddannelseslængde på kognitive evner. Effekten fra mindre end 2 værelser på kognitive evner er dog relativt beskednen.

6 Robusthedsanalyse

I dette afsnit vil jeg foretage tre robusthedstjek af den foretrukne Second Stage model (Tabel 4, kolonne (4)). Først undersøges det, om metodeændringer eller tidsvarierende faktorer driver forklaringen, dernæst udfordres den økonometriske antagelse om tilfældig stikprøve, og til sidst undersøges det, om effekten af uddannelseslængde på kognitive evner kan beskrives som en kvadreret funktion. Resultaterne af de tre robusthedstjek ses i tabel 5 og beskrives i følgende afsnit.

6.1 Metodeændringer og tidsvarierende faktorer

For at teste om metodeændringer, eller andre tidsvarierende faktorer, driver en del af effekten af uddannelseslængden på kognitive evner tilføjes wave-dummies i kolonne (1). Wave-dummierne indikerer hvilket år respondenterne foretog sit baseline interview. Det ses af kolonne (1), at estimationsresultaterne ikke ændres betydeligt ved at tilføje wave-dummies. Derimod forøges standardfejlen. Dette skyldes at tilføjelsen af irrelevante variable ikke påvirker estimatet, men derimod forøger variansen af modellen.²⁰ En fælles F-test for Wave-dummierne viser ligeledes, at de ikke indgår signifikant i modellen. Derved er modellen robust overfor eventuelle metodeændringer og tidvarierende faktorer.

6.2 Tilfældig stikprøve

Respondenterne i easySHARE er som udgangspunkt udvalgt tilfældigt, men for hvert individ der er tilfældigt udvalgt, er ægtefællen ligeledes spurgt om de ønsker at deltage i spørgeskemaet. Denne udvælgelsesmetode betyder, at 43 pct. af respondenterne er ægtefæller, hvorved antagelsen om tilfældig stikprøve kan være brudt. [Blossfeld et al., 2003] argumenterer for, at ægtefæller besidder flere af samme karakteristika så som uddannelseslængde og socioøkonomisk klasse. Hvis dette er tilfældet, da kan der i stikprøven være en overrepræsentation af bestemte grupper, hvorved Gauss-Markov's anden antagelse ikke er overholdt. Dette vil betyde, at estimaterne er inkonsistente og biased. I kolonne (2) er stikprøven restrikeret til kun at indeholde den ene ægtefælle. Igen ses det, at estimationsresultaterne ikke ændres betydeligt. En mindre stikprøve forøger modellens varians. Antagelsen om tilfældig stikprøve accepteres, da estimationsresultaterne ikke ændres betydeligt.

²⁰[Wooldridge, 2016] side 77-78

Tabel 5: Robustheds analyse

	IV (1) kognitive evner ¹	IV (2) kognitive evner ¹	IV (3) kognitive evner ¹
Uddannelseslængde (E_i)	0.233* (0.122)	0.232** (0.117)	-3.275 (9.838)
E_i x udkantsområde	-0.000 (0.023)	-0.012 (0.025)	2.785 (12.611)
E_i x værelser ≤ 2	0.059* (0.031)	0.050* (0.030)	5.793 (14.627)
E_i x bøger ≤ 25	-0.213* (0.118)	-0.183* (0.110)	-3.181 (7.468)
E_i x primære forsørgers job	0.010 (0.028)	-0.002 (0.027)	3.811 (9.188)
E_i x brugsgoder	0.036** (0.017)	0.037** (0.016)	0.454 (1.622)
E_i x dårlig elev	0.003 (0.027)	-0.017 (0.028)	1.132 (3.778)
E_i x dårligt f. helbred	-0.032 (0.053)	-0.008 (0.068)	2.457 (10.433)
E_i x dårligt m. helbred	0.222 (0.315)	0.202 (0.319)	-0.663 (1.702)
E_i^2			0.188 (0.522)
E_i^2 x udkantsområde			-0.139 (0.627)
E_i^2 x værelser ≤ 2			-0.302 (0.768)
E_i^2 x bøger ≤ 25			0.134 (0.332)
E_i^2 x primære forsørgers job			-0.202 (0.484)
E_i^2 x brugsgoder			-0.023 (0.090)
E_i^2 x dårlig elev			-0.059 (0.191)
E_i^2 x dårligt f. helbred			-0.139 (0.556)
E_i^2 x dårligt m. helbred			0.048 (0.122)
Konstant	-2.516 (1.971)	-2.629 (1.712)	11.391 (41.488)
N	8907	6988	8907
Barndomsvariable	Ja	Ja	Ja
Wave dummies	Ja	Nej	Nej
Instrumenter	$z_i, z_i X_{2i}$	$z_i, z_i X_{2i}$	$z_i, z_i X_{2i}, z_i^2, z_i^2 X_{2i}$
F-teststørrelse ²	17,61	26,10	15,29

Note: Alle regressioner er kontrolleret for land, køn, alder og landespecifikke kohortetrends.

Robuste standardfejl er vist i parenteser. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

¹ kognitive evner er standardiseret.²F-test af instrumenternes samlede signifikans.

6.3 Effekten af uddannelseslængde på kognitive evner som kvadratisk funktion

I dette robusthedstjek undersøges det, om den marginale effekt af uddannelse på kognitive evner afhænger af, hvor lang en uddannelse man har i forvejen. Der tilføjes et kvadratisk led af uddannelseslængden (E_i^2) samt kvadratiske interaktionsled mellem uddannelseslængde og barndomsvariable ($E_i^2 X_{i2}$) til den foretrukne model.

Til instrumenterne tilføjes den kvadrerede obligatoriske uddannelseslængde (z_i^2) samt interaktionsled mellem denne og barndomsvariablene ($z_i^2 X_{i2}$). En F-test viser, at instrumenterne tilsammen indgår på et 1-pct. signifikansniveau. Af kolonne (3) ses det, at alle estimer ændres betydeligt, samt bliver insignifikante, når de kvadratiske led tilføjes. Modellen er derved ikke robust overfor tilføjelsen af de kvadratiske led.

I analysen inkludere jeg flere barndomsvariable som proxy for individets socioøkonomiske baggrund. Baggrundsvariablene er derved delvist korreleret. Ved at tilføje kvadratiske led, øges risikoen for multikollinearitet mellem de forklarende baggrundsvariable. For at teste om der er multikollinearitet i modellen, laver jeg en F-test for hver af instrumenternes lineære og kvadratiske form. Med en F-teststørrelse på 18,36, indgår den lineære og kvadratiske form af den obligatoriske uddannelseslængde (z_i) på et 1 pct. signifikansniveau. Af de lineære og kvadratiske interaktionsled, mellem instrument og barndomsvariable, indgår *udkantsområde*, *bøger* ≤ 25 og *brugsgoder* på et 1 pct. signifikansniveau. At flere af instrumenternes lineære og kvadratiske form indgår signifikant tyder på, at den marginale effekt af uddannelseslængde på kognitive evnerne afviger med så lidt, at det er svært at skelne den kvadratiske form fra den lineære form. Det er derfor sandsynligt, at der er multikollinearitet mellem de forklarende variable, hvorved flere af de forklarende variable er stærkt lineært afhængige. Dette vil få standard fejlen til at stige, hvorved effekten af uddannelseslængden på kognitive evner ikke længere vil indgå signifikant.

Det har i denne analyse ikke været muligt at finde signifikante værdier for den kvadratiske form. I Appendiks C.1 er de gennemsnitlige kognitive evner betinget af uddannelseslængde og barndomsvariable afbilledet. Det ses heraf, at trenden for de gennemsnitlige kognitive evner bedst kan beskrives som en aftagende funktion af uddannelseslængde, hvor barndomsvariablene *dårlig elev*, *bøger* ≤ 25 og *brugsgoder* giver et niveauskift i kognitive evner. Denne fortolkning er dog kun bygget på en grafisk bivariat analyse, og det kan derfor ikke bekræftes, at der eksisterer en sådan statistisk sammenhæng. Figuren ser dog ud til at bekræfte, at hældningen, der angiver det kvadratiske led, er relativt svag for flere af baggrundsvariablene. Med et større datasæt vil det måske være muligt at identificere den korrekte kvadratiske form.

På trods af at den lineære multipel model ikke er robust overfor kvadratiske led, er

den stadig at foretrække. Dette skyldes, at den kvadratiske form formeligt ikke afviger betydeligt fra den lineære form, samt at der med stor sandsynlighed er multikollinearitet mellem de forklarende variable.

7 Konklusion

I ovenstående analyse er det blevet undersøgt, hvorvidt effekten af uddannelseslængde på kognitive evner for ældre²¹ varierer på baggrund af socioøkonomiske karakteristika.

Dette undersøgte jeg ved at sammensætte et datasæt fra SHARE paneldatabase, der indeholder grunddata, barndomsvariable og eksperimentelle data for kognitive evner. Uddannelsesreformer, der forlænger den obligatoriske uddannelseslængde, er blevet anvendt som instrument i en IV estimation til at isolere den eksogene effekt af uddannelseslængde på kognitive evner. Interaktionsled, mellem obligatorisk uddannelseslængde og barndomsvariable, er blevet anvendt til at isolere den ekstra effekt som grupper, med specifikke socioøkonomiske karakteristika, har.

På baggrund af disse data har jeg fundet, at et ekstra års uddannelse har en signifikant og positiv effekt på kognitive evner uanset socioøkonomisk baggrund, men at niveauet af effekten varierer på baggrund af socioøkonomiske karakteristika. Individer, der er opvokset i et hjem med færre end 25 bøger, forbedrer deres kognitive evner med 0,021 standardafvigelse per års uddannelse. Individer, der er opvokset i hjem med det maksimale antal brugsgoder, forbedrer derimod deres kognitive evner med 0,278 standard afvigelse per års uddannelse. Dette tyder på, at den socioøkonomisk klasse påvirker effekten af uddannelse på kognitive evner.

8 Diskussion

Denne analyse er naturligvis ikke udtømmende. Enhver empirisk analyse er begrænset af kvaliteten af de tilgængelige data. I dette tilfælde anvendes paneldata fra SHARE, der er indsamlet med henblik på at lave empiriske analyser om det europæiske helbred. Derved er data ikke indsamlet med henblik på at undersøge om uddannelsens effekt på kognitive evner er betinget af socioøkonomiske karakteristika. Dette giver en række svagheder.

For det første er den sociale klasse angivet ved barndomsvariable, som ofte er angivet som en dummy. Derved omhandler analysen grupper med meget specifikke socioøkonomiske karakteristika, f.eks. færre end 25 bøger i barndomshjemmet. Ved at anvende flere barndomsvariable som proxy for socioøkonomisk klasse, vil barndomsvariablene i nogen grad værre korreleret. Dette øger risikoen for multikollinearitet. Hvis der

²¹Individer der er 50 år eller ældre.

i SHARE var data for f.eks. forældres indkomst, ville denne variabel sandsynligvis være en bedre proxy for socioøkonomisk klasse.

En anden svaghed er, at socioøkonomisk klasse også dækker over faktorer der er svære at måle som eksempelvis værdier, livsstil og sociale roller. Hvis disse faktorer påvirker effekten af uddannelseslængde på kognitive evner, da kan den socioøkonomiske klasse være af større betydning end estimeret i denne analyse. Det kunne derfor være interessant at undersøge, om effekten af uddannelseslængde på kognitive evner varierer på baggrund af andre mål for socioøkonomisk klasse.

At socioøkonomisk klasse har betydning for effekten af uddannelse på kognitive evner er forklaret på baggrund af [Becker & Toms, 1985]. Det antages, at individer, der er født ind i familier med fordelagtige familiespecifikke evner, modtager et større humankapitalafkast, når der investeres i dem. Det kunne derfor være interessant at undersøge, om forældre, der tilhører den højere socioøkonomiske klasse, investerer mere tid og ressourcer i deres børn end forældre fra en lavere socioøkonomisk klasse. Dette kunne eventuelt undersøges, ved at efterse hvilke socioøkonomiske klasser der tilkøber lektiehjælp på det private marked.

Analysen viser, at uddannelseslængde forbedrer individers kognitive evner uanset socioøkonomisk klasse, men at individer fra en højere socioøkonomisk klasse forbedre deres kognitive evner relativt mere per års uddannelse. Det kunne derfor være relevant at teste, om forskellene i kognitive evner stiger med uddannelseslængden betinget af socioøkonomisk klasse.

Referencer

- [Nagoshi et al., 1993] Nagoshi, Craig T, Ronald C Johnson, and Kelly Ann M Honbo. "Family Background, Cognitive Abilities, and Personality as Predictors of Education and Occupational Attainment across Two Generations." *Journal of Biosocial Science* 25.2 (1993): 259-76. Web.
- [Hunter, 1986] Hunter, John E. "Cognitive Ability, Cognitive Aptitudes, Job Knowledge, and Job Performance." *Journal of Vocational Behavior* 29.3 (1986): 340-362. Web.
- [Ward et al., 1997] Scott Ward, Daniel Wackman, and Ellen Wartella (1977) ,"The Development of Consumer Information-Processing Skills: Integrating Cognitive Development and Family Interaction Theories", in *NA - Advances in Consumer Research* Volume 04, eds. William D. Perreault, Jr., Atlanta, GA : Association for Consumer Research, Pages: 166-171.
- [Burks et al., 2009] Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini. "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the National Academy of Sciences of the United States of America* 106.19 (2009): 7745-750. Web.
- [Falch et al., 2011] Falch, Torberg, and Sofia Sandgren Massih. "The Effect of Education on Cognitive Ability." *Economic Inquiry* 49.3 (2011): 838-56. Web.
- [Banks et al., 2012] Banks, James, and Fabrizio Mazzonna. "The Effect of Education on Old Age Cognitive Abilities: Evidence from a Regression Discontinuity Design*." *Economic Journal* 122.560 (2012): 418-48. Web.
- [Ermisch et al., 2012] Ermisch, John, et al., editors. *From Parents to Children: The Intergenerational Transmission of Advantage*. Russell Sage Foundation, 2012. Chapter 10. JSTOR, www.jstor.org/stable/10.7758/9781610447805.
- [Feinstein, 2003] Feinstein, Leon, *Inequality in the Early Cognitive Development of British Children in the 1970 Cohort*. 2003 *Economica* Volume 70, Issue 277. <https://doi.org/10.1111/1468-0335.t01-1-00272>
- [Decker et al., 2017] Falk, Armin, Fabian Kosse, Pia Pinger, Hannah Schildberg-Hörisch, and Thomas Deckers. *Socio-Economic Status and Inequalities in Children's IQ and Economic Preferences* (2019). Web.
- [Grant et al., 2010] Grant, Michael, D. Kremen, William Jacobson, S. Franz, Kristen Xian, C. Eisen, Carol Toomey, Hong Murray, and Seth Lyons. "Does Parental Education Have a Moderating Effect on the Genetic and Environmental Influences of General Cognitive Ability in Early Adulthood?" *Behavior Genetics* 40.4 (2010): 438-46. Web.

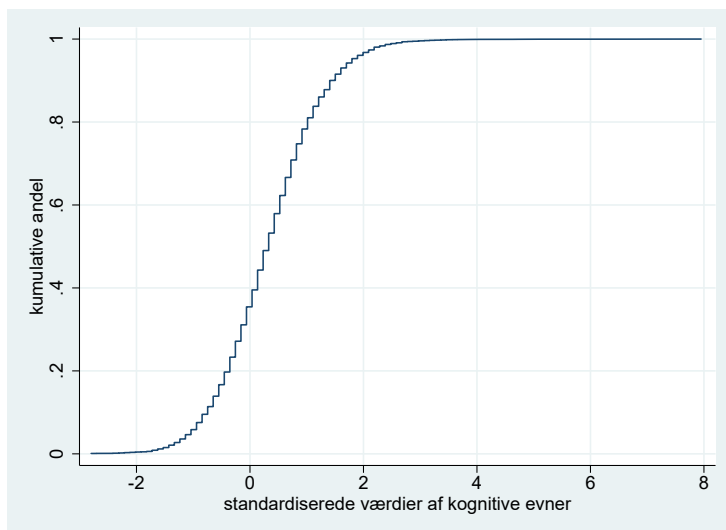
- [Ermisch et al., 2012] Ermisch, John, Jäntti, Markus, and Smeeding, Timothy M. From Parents to Children, the Intergenerational Transmission of Advantage. New York: Russell Sage Foundation, 2012. Print.
- [Glymour et al., 2008] Glymour, M M, I. Kawachi, C S Jencks, and L F Berkman. "Does Childhood Schooling Affect Old Age Memory or Mental Status? Using State Schooling Laws as Natural Experiments." *Journal of Epidemiology and*
- [Staiger & Stock, 1997] Staiger, Douglas, and James Stock. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65.3 (1997): 557-86. *Web.Community Health* 62.6 (2008): 532-537. Web.
- [Haveman & Wolfe, 1995] Haveman, Robert, Robert H, and Wolfe, Barbara, Barbara Schull. *Succeeding Generations, on the Effects of Investments in Children*. New York: Russell Sage Foundation, 1994. Print. Kapitel 2: "Toward understanding the determinants of childrens success & Kapitel 4: "Investments in Children: Some Simple Relationships".
- [Hansen et al., 2014] Naja Liv Hansen, Martin Lauritzen, Erik Lykke Mortensen, Merete Osler, Kirsten Avlund, Birgitte Fagerlund, Egil Rostrup *Subclinical Cognitive Decline in Middle-Age is Associated With Reduced Task-induced Deactivation of the Brain's Default Mode Network* . *Human Brain Mapping* (2014),DOI: 10.1002/hbm.22489
- [Tilvis et al., 2004] Reijo S. Tilvis, Mervi H. Kähönen-Väre, Juha Jolkkonen, Jaakko Valvanne, Kaisu H. Pitkala, Timo E. Strandberg *Predictors of cognitive decline and mortality of aged people a 10-year period*. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 59.3 (2004): 268-74.
- [Kommer et al., 2013] Tessa N. van Den Kommer, Hannie C. Comijs, Maria J Aartsen, Martijn Huisman, Dorly J.H. Deeg, Aatjan T.F. Beekman *Depression and Cognition: How Do They Interrelate in Old Age?* *American Journal Of Geriatric Psychiatry* 21.4 (2013): 398-410.
- [Clouston et al., 2015] Sean A.P Clouston, M. Maria Glymour, Graciela Muñoz Terrera *Educational inequalities in aging-related declines in fluid cognition and the onset of cognitive pathology* *Alzheimers Dement* (Amst) 1. (2005): 303–310.
- [Crespo et al., 2014] Laura Crespo, Borja López-Noval, Pedro Mira *Compulsory schooling, education, depression and memory: New evidence from SHARELIFE* *Economics of Education Review*, Volume 43, December 2014, Pages36-46, ISSN 0272-7757, <http://dx.doi.org/10.1016/j.econedurev.2014.09.003>
- [Heckman, 2006] James J. Heckman *Skill Formation and the Economics of Investing in Disadvantaged Children*. *Science* (New York, N.Y.) 312.5782 (2006): 1900-1902.

- [Becker & Tomes, 1985] Gary S Becker & Nigel Tomes, 1986. *Human Capital and the Rise and Fall of Families* Journal of Labor Economics, University of Chicago Press, vol. 4(3), pages 1-39, July.
- [Mirowsky & Ross, 2003] John Mirowsky & Catherine E. Ross *Education, social status, and health* New York: Aldine De Gruyter; 2003,
- [Tucker-Drop et al., 2014] Elliot M. Tucker-Drob, Daniel A. Briley, K. Paige Harden *Genetic and Environmental Influences on Cognition Across Development and Context*. Curr Dir Psychol Sci. 2013;22(5):349–355. doi:10.1177/0963721413485087
- [Wooldridge, 2016] Jeffery M. Wooldridge *Introductory Econometrics - A modern approach*
- [Blossfeld et al., 2003] Hans-Peter Blossfeld & Andreas n.d. Trimm *Introduction In: Who marries whom - educational systems as marriage in modern societies: A conceptual framework*. 2003, Kluwer Academic publishers. Page 9-17
- [Brunello al., 2009] Giorgio Brunello, Margherita Fort, Guglielmo Weber *Change in compulsory schooling, Education and the Distribution of Wages in Europe* 2009, The Economic Journal, Volume 119, Issue 536, page 516-539
- [Murtin & Viarengo, 2008] Fabrice Murtin, Martina Viarengo *The expansion and convergence of compulsory schooling in western europe, 1950-2000* 2008, Economica (2011) 78, page 501-522
- [Arendt, 2005] Jacob Nielsen Arendt *Does Education cause better health? A panel data analysis using school reforms for identification* 2005, Economics of Education Review, vol. 24, issue 2, 149-160

A Appendiks: Udeladte figurer

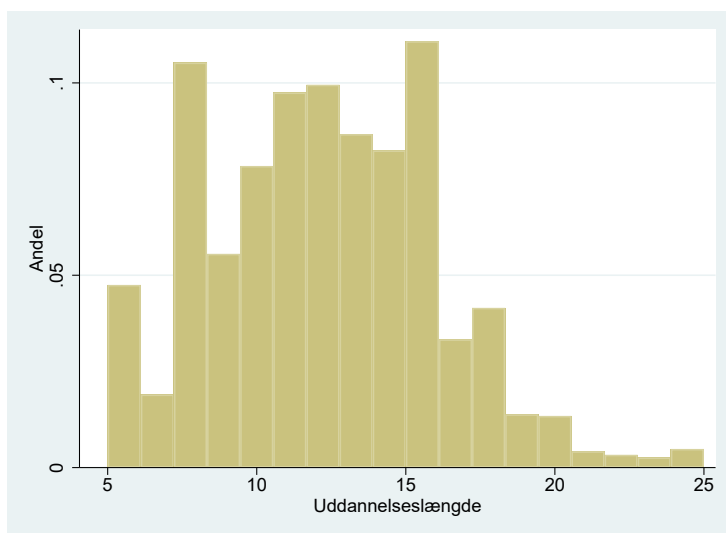
A.1 Den kumulative fordeling af kognitive evner i stikprøven

Figur 4: Den kumulative fordelingen af de standardiserede kognitive evner

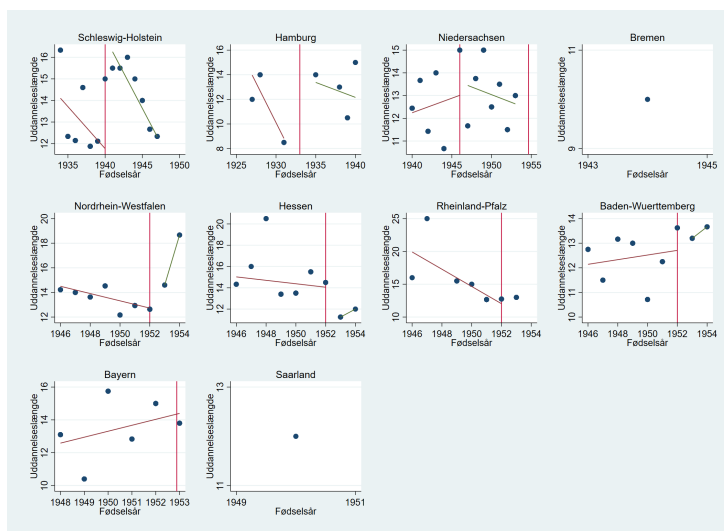


A.2 Histogram over uddannelseslængden

Figur 5: Fordelingen af uddannelseslængden i stikprøven



Figur 6: De tyske reformers betydning for uddannelseslængden på regionsniveau



A.3 De tyske reformers betydning for uddannelseslængden

B Appendiks: Anvendte tests

B.1 Endogenitets test

IV-estimatoren er kun at foretrække hvis uddannelseslængden er endogen. Hvis uddannelseslængden derimod er eksogen da er OLS-estimatoren at foretrække.²² Derfor bruges en endogenitetstest til at undersøge om uddannelseslængden rent faktisk er endogen bestemt. Først opskrives Second Stage for modellen med baseline specifikationer:

$$CA_i = \gamma_1 + \gamma_2 E_i + \gamma'_{31} X_{1i} + v_i \quad (\text{B.1})$$

Det forventes at uddannelseslængden er endogen, og derfor opskrives First Stage, hvor uddannelseslængde regresseres på instrument og uafhængige variabel:

$$E_i = \delta_1 + \delta_2 z_i + \delta'_{31} X_{1i} + \epsilon_i \quad (\text{B.2})$$

Fra denne regression estimeres fejllende $\hat{\epsilon}_i$. Derefter inkluderes disse fejllende i Second Stage regressionen (B.1):

$$CA_i = \gamma_1 + \gamma_2 E_i + \gamma'_{31} X_{1i} + \omega \hat{\epsilon}_i + v_i \quad (\text{B.3})$$

Da vi i baseline-modellen kun har en mistænkt endogen variabel kan en t-test anvendes. Nulhypotesen er, at variabelen er eksogen, og at fejllende ($\hat{\epsilon}$) derved er lig 0:

²²Kilde: Wooldridge 2016 side 481-482.

$$H_0 : \omega = 0 \tag{B.4}$$

Hvis $\omega \neq 0$ forkastes nulhypotesen. Uddannelseslængden er derved endogen og kan delvis forklares af instrumentet som i dette tilfælde er den obligatoriske uddannelseslængde.

B.2 Breusch-Pagan test

Formålet med Breusch-Pagan testen er at undersøge om modellen har heteroskedastiske fejllid. ²³. Fremgangsmetoden er således:

1. Second Stage modellen estimeres og modellens fejllid (\hat{v}_i) forudsiges.
2. Den kvadrerede værdi af fejllidende beregnes (\hat{v}_i^2). Denne antages som værende den sande værdi af (v_i^2).
3. De kvadrerede fejllid regresseres på de uafhængige variable. $\hat{v}_i^2 = z_i + \rho' X_{1i}$ Herfra gemmes R^2 -værdien.
4. En LM-test anvendes til at udregne p-værdien af en χ_k^2 -fordeling. såfremt p-værdien er lavere end det valgte signifikansniveau afvises hypotesen om homoskedastiske fejllid, og modelens varians er derved ikke konstant.

C Appendiks: Effekten af uddannelse på kognitive evner - fitted values.

C.1 De gennemsnitlige kognitive evner fordelt betinget af uddannelseslængde og barndomsvariable. Fitted values

²³Kilde: [Wooldridge, 2016], side 250-251

Figur 7: De gennemsnitlige kognitive evner betinget af uddannelseslængde og barn-
domsvariable. Fitted values

