

MORTEN RASMUS PUCK
OG JEPPE BUNDSGAARD

**UNDERSØGELSE AF
REPRÆSENTATIVITETEN
AF ICILS 2013**



AARHUS UNIVERSITET
DPU

Morten Rasmus Puck og Jeppe Bundsgaard

Undersøgelse af repræsentativiteten af ICILS 2013

DPU, Aarhus Universitet, 2016

Titel:

Undersøgelse af repræsentativiteten af ICILS 2013

Forfattere:

Morten Rasmus Puck og Jeppe Bundsgaard

Udgivet af:

DPU, Aarhus Universitet, 2016

© 2016, forfatterne

1. udgave

Kopiering tilladt med tydelig kildeangivelse

Omslag og grafisk tilrettelæggelse:

Knud Holt Nielsen

Forsidefoto:

Colourbox

Indhold

ICILS-UNDERSØGELSENS REPRÆSENTATIVITET	7
EN KVANTITATIV UNDERSØGELSE UDFORDRINGER	8
METODISKE OVERVEJELSER	9
UNDERSØGELSESMETODEN.....	10
TEST FOR HVOR REPRÆSENTATIV ICILS ER.....	11
ANALYSE AF KØNSFORDELINGEN	12
ANALYSE AF ETNICITET: EFTERKOMMERE	12
ANALYSE AF ETNICITET: INDVANDRERE	13
ANALYSE AF ETNICITET: DANSKERE	14
ANALYSE AF KØNSFORDELINGEN HOS LÆRERNE	15
FORDELINGEN AF SKOLER PÅ REGIONER	15
ANALYSEN AF FORDELINGEN AF SKOLER PÅ KOMMUNENIVEAU	17
KONKLUSION PÅ REPRÆSENTATIVITETSTESTEN.....	20

ICILS-undersøgelsens repræsentativitet

ICILS 2013 er den nyeste internationale sammenligning af elevers kompetencer fra IEA (International Association for the Evaluation of Educational Achievement). IEA er erfarne med denne form for tests, da det er dem, der står for blandt andet TIMSS (Trends in International Mathematics and Science Survey) og PIRLS (Programme of International Reading Literacy Survey). IEA har meget strikse krav til omfanget og kvaliteten af data, som skal være opfyldt før man kommer med i hovedtabellerne i den internationale rapport. Danmark nåede ikke disse krav ved dataindsamlingen i 2013.

Selve udvælgelsen (samplingen) foregår på baggrund af inddeling af danske skoler i en række grupper, bl.a. på baggrund af størrelse, lokalitet osv. Dette kaldes stratificering. Fra disse grupper udtrækkes 150 skoler som benævnes samplede skoler. Da det kan forventes at ikke alle skoler ønsker at deltage i undersøgelsen, udtrækkes desuden for hver samplede skole *First replacement* og *Second replacement*.

Når en skole indvilliger i at deltage, sender skolekoordinatoren en liste over alle elever på 8. klassetrin. Fra denne liste udtrækker IEA's software tilfældigt 20 elever som derefter deltager i testen.

Da listen over skoler tog udgangspunkt i de senest tilgængelige tal ved udtrækningen stammede fra 2009, så en række skoler var enten lukket eller fusioneret ind i en anden skole. Konkret var 9 af de 150 samplede skoler lukket, hvilket betød at de 9 skoler og deres erstatningsskoler forsvinder fra vores stikprøve (de repræsenterer skoler der er lukket, og de tilbageværende skoler indeholder skoler der repræsenterer skoler der har fået ekstra elever fra lukkede skoler). Det betød at vi kunne få deltagelse fra maksimalt 141 skoler.

IEA's strenge krav til data indebærer at for at kunne være repræsenteret uden anmærkninger i hovedtabellerne, skal 85% af de *samplede* skoler deltage. Et land kan optræde i hovedtabellerne med en anmærkning hvis 85% af samtlige 141 skoler eller deres erstatningsskoler deltager, blot 50% af disse er *samplede* skoler. Hvis disse krav ikke opfyldes, vil data fra dette land ikke optræde i hovedtabellerne, men blive rapporteret i særskilte tabeller i tilknytning til hovedtabellerne.

I Danmarks ICILS-undersøgelse deltog 103 skoler, som fordeler sig på denne måde:

Tabel 1: Fordelingen af skoler i ICILS			
Antallet af skoler, der har deltaget og deres status i testen	Antal	Andel i stikprøven	ud af de 141 skoler
<i>Samplede skoler</i>	54	0,52	0,383
<i>First replacement</i>	25	0,24	0,177
<i>Second replacement</i>	24	0,23	0,170
I alt	103		0,730
Nationale skole (First eller Second replacement hvor samplet skole deltager) og skoler med mistede eller for få data	9		
Kilde: ICILS			

Som det fremgår, har Danmark kun haft 73,05 % deltagende skoler ud af de 141 udvalgte skoler, lever Danmark ikke op til IEA's grænse på 85%, for at kunne optræde i hovedtabellerne med anmærkninger.

Det fremgår desuden at størstedelen af de deltagende skoler er *samplede* skoler, men at der også er erstattet skoler i vidt omfang. Halvdelen af skolerne er af første eller andet udskiftningslag. Danmark lever således ikke op til IEA's meget strenge krav (det samme var i øvrigt tilfældet for yderligere tre af de 21 deltagende lande). Kravene overgår dog markant hvad man i andre sammenhænge stiller til statistiske undersøgelser, og vi vil derfor med god grund kunne hævde at de indsamlede data kan bruges som grundlag for pålidelige analyser og konklusioner.

Men for at sikre os at de skoler der har deltaget, ikke er kendetegnet ved særlige træk, har vi gennemført en grundig analyse af hvorvidt de deltagende skoler (stikprøven) kan betragtes som repræsentative for alle danske skoler (populationen). Før vi undersøger den stikprøve vi har anvendt, vil vi kort angive begrundelser for at det ikke lykkedes at opnå de ønskede deltagelsesprocenter.

En kvantitativ undersøgelses udfordringer

ICILS består af en prøveundersøgelse (field trial) som gennemførtes i 2012, et år før hovedundersøgelsen (main study). I prøveundersøgelsen anvendes et mindre antal skoler (25 skoler udtrækkes), men i øvrigt afprøves alle faser af undersøgelsen med første version af instrumenterne. Ved prøveundersøgelsen stod det hurtigt klart for det danske ICILS-team, at det er meget svært at få skoler til at deltage i større kvantitative

undersøgelser. Derfor gjorde vi meget ud af at forberede kontakten til skolerne, vi formulerede grundigt hvilke fordele skolerne selv ville få af at deltage, og vi var klar på at der skulle gøres et meget stort benarbejde for at få skolerne til at deltage. Dette arbejde lykkedes i den forstand at vi ved påskeferiens start i 2013 havde tilsagn fra 120 skoler om at de ville deltage og gennemføre testen fra påske og frem mod sommerferien. Dette var lige netop nok til at vi ville nå 85%-kravet. Vi planlagde derfor at fortsætte arbejdet med at få skoler til at deltage, Men i påskeferien gik forhandlingerne mellem KL og Danmarks Lærerforening i hårdknude og KL varslede en lockout. Denne lockout trak ud i mere end en måned, og da den var slut, var en del af de planlagte testsessioner og lærerspørgeskemabesvarelser allerede passeret. De resterende aftaler blev i stort omfang aflyst af skolerne. Årsagerne hertil var mange og forståelige, og på trods af vores forsøg på at overtale skolerne til at genfastsætte et testtidspunkt, måtte vi sande at vi ved starten af sommeren var nede på under 60 deltagende skoler.

Vi fik efter forhandlinger med IEA lov til at fortsætte dataindsamlingen i efteråret, så eleverne nu gik i 9. i stedet for i 8. klasse. Vi tilbød alle skoler at vi kunne komme med et sæt computere og selv stå for testen af eleverne (vi ansatte fire studentermedhjælpere til dette arbejde), og vi gjorde i øvrigt hvad vi kunne for at lette arbejdet på skolen så meget som muligt. Da efterårsferien nærmede sig, kunne vi se at antallet af skoler også nærmede sig de 120 skoler som ville svare til 85%, så vi fortsatte kontakten til skolerne, nu med henvendelse til både samplede skoler og deres første- og anden-erstatning, og ind imellem med tredje, fjerde og femte henvendelse på telefon, mail og brev til de samme skoler. Vi overtrådte mange skolers grænse for hvor mange gange vi med rimelighed kunne henvende os. Og på trods af alt endte vi med at have 112 skoler der testede deres elever. Heraf havde 5 skoler en skole før sig "i rækken" (samplet skole eller first replacement). To skoler testede for få elever (under 50% af de udtrukne) og kunne derfor ikke tælle med, og for at føje spot til skade blev usb-stik med data fra to skoler stjålet fra en bil før de var uploadet til testserveren. I det endelige datasæt har vi således 103 deltagende skoler.

Metodiske overvejelser

Den danske del af ICILS-projektet kunne således ikke leve op til de fastsatte krav fra IEA, og derfor har vi gennemført en repræsentativitetsanalyse af stikprøven for at finde ud af om vi med rimelighed kan konkludere på baggrund af de danske ICILS-resultater.

Repræsentativ betyder, at den stikprøve der udtages, skal have samme egenskaber som hele populationen. Fordi alle ikke undersøges, skal dem, der spørges, tale på

vegne af andre. Hver person repræsenterer altså også andre personer i populationen. Derfor er det vigtigt at de to populationer er sammenlignelige.

Et vigtigt element i ICILS-undersøgelsen er udtrækningen til stikprøven. Optimalt set vil vi gerne have en simpel tilfældig udtrækning, således at sandsynligheden for at en given skole/elev udtrækkes, er lige så stor som de øvrige skoler/elevs sandsynlighed. Problemet med simpel tilfældig udtrækning er at man kunne komme i en situation hvori skoler i bestemte geografiske områder ikke vil blive repræsenteret. Det vil eksempelvis ikke være repræsentativt, hvis skoler fra Region Nordjylland ikke var repræsenteret, således at undersøgelsens resultater ikke vil gælde for hele landet. Denne problemstilling kan løses ved at bruge en stikprøveudvælgelse man kalder stratificering, hvilket vil sige at man sørger for at udtrække nogle tilfældige skoler i samme geografiske område. På denne måde indeholder vores stikprøve fordelinger af skoler i alle regioner i Danmark. Dog har stratificering det problem at vores udtrækning ikke længere er helt tilfældig. Vores fordeling af udtrykket skoler i regionerne er nu planlagt, mens de enkelte udtrukne skoler er tilfældigt valgt. Dette løses i behandlingen af data ved at angive vægte for den enkelte skole. En vægt angiver hvor meget den enkelte skole skal veje i analysen af data.

Når skolerne er valgt, foretages en udvælgelse af 20 elever i 8. klasse på hver skole. Disse elever er tilfældigt udvalgt.

Undersøgelsesmetoden

Fra de nationale datacentraler (primært Danmarks Statistik (DST)) kan vi finde den sande fordeling af forskellige grupper i populationen (køn, alder, skolestørrelse osv.). Ved at sammenligne populationens fordeling med fordelingen i den stikprøve der er anvendt i ICILS, kan vi sandsynliggøre at stikprøven også på de undersøgte parametre ligner populationen.

Vi gennemfører undersøgelsen med brug af dummy-variable. En dummy-variabel reducerer alt til enten eller: Enten er man med, ellers er man ikke. Denne metode vil vi bruge på alle kriterier ligegyldigt om der er to valgmuligheder (køn), eller flere valgmuligheder (regioner). Med dummy-variable kan vi udregne andelen, for eksempel, af drenge ud af hele vores stikprøve. Gennemsnittet er bare udregnet som

$$\mu_{ICILS} = \frac{m_1 * 1 + m_0 * 0}{m_1 + m_0} = \frac{m_1}{m_1 + m_0}$$

hvor m_1 er antallet af deltagere der har 1 som observation, og m_0 er antallet af deltagere, der har 0 som observation. Observationen 1 betyder at de er i gruppen vi undersøger, og observationen 0 betyder at de ikke er i gruppen vi undersøger.

Ud fra dette gennemsnit udregner vi stikprøvens varians ud fra formelen

$$S^2 = \frac{1}{m_1 + m_0 - 1} (m_1(1 - \mu)^2 + m_0(0 - \mu)^2)$$

, som siger hvor meget variation der er i vores stikprøve af den variabel som vi undersøger. Dette vil vi bruge til at udregne standardfejlen for variablen.

$$s.e. = \sqrt{\frac{S^2}{m_1 + m_0}}$$

Denne standardfejl bruger vi til at udregne en t-værdi, som udtrækker om vi kan befæste eller forkaste vores hypotese om at vores stikprøvegennemsnit er lig med landsgennemsnittet.

Når man skal teste noget statistisk, er man nødt til at have en hypotese om den værdi der estimeres. Alle vores h_0 -hypoteser er at estimerede gennemsnit er lig med det observerede gennemsnit fra de nationale statistikker. Alternativ-hypotesen er at den estimerede værdi kan være både større og mindre end den faktiske værdi fra de nationale statistikker. For at opsummere vores hypoteser: $h_0: \mu_{ICILS} = \mu_{DST}$; $h_A: \mu_{ICILS} \neq \mu_{DST}$.

Vores hypoteser testes som en t-test:

$$t = \frac{\mu_{ICILS} - \mu_{DST}}{s.e.}$$

hvor μ_{ICILS} er denne gennemsnitlige værdi fra ICILS undersøgelsen. μ_{DST} er den gennemsnitlige værdi i variablen fra Danmarks Statistik. På baggrund af vurderinger af normalfordelingen sættes den kritiske værdi af t-værdien typisk til 1,96 (som betyder at 95% af de mulige stikprøver vil resultere i at hypotesen ikke kan afvises) (denne t-værdi svarer til en p-værdi på 5%). Hvis vores t-værdi er numerisk større end 1,96, vil vi sige at der er signifikant forskel imellem vores estimerede værdi og den faktiske værdi. Det vil sige at der er forskel på de to andele.

Test for hvor repræsentativ ICILS er

I det følgende gennemføres en analyse af hvor repræsentativ stikprøven i ICILS er i forhold til hele populationen, som er 8.-klasseelever i skoleåret 2012-2013. Vi betragter repræsentativitet som at stikprøven har de samme karakteristika, som hele populationen har. Konkret undersøger vi om kønsandelene, etniske tilhørsforhold, skolerne geografiske placering og lærernes kønsfordeling stemmer med de nationale statistikker.

Analyse af kønsfordelingen

I dette delafsnit undersøger vi om andelen af drenge i ICILS-undersøgelsen passer med andelen fra DST.

Tabel 2: Test af om drengenes andel passer	
Gennemsnittet fra ICILS: μ_{ICILS}	0,51
Gennemsnittet fra DST: μ_{DST}	0,51
Stikprøvevariansen S^2	0,25
Standardfejlen $s.e$	0,01
t-testværdien: $H_0: \mu_{ICILS} = \mu_{DST}$	0,02
Kilde: ICILS, DST tabel U1907. Egne beregninger	

Fra tabel 2 kan vi se at t-testværdien er 0,02, hvilket ikke er større end 1,96, så derfor kan vi ikke afvise vores H_0 -hypotese, og vi kan ikke sige at der er statistisk forskel på gennemsnittet fra ICILS og gennemsnittet fra DST. Sagt med andre ord kan vi gå ud fra at der ikke er forskel på ICILS og den generelle sammensætning af elever i de 8.-klasser.

Analyse af etnicitet: Efterkommere

Eleverne deles både i Danmarks Statistik og ICILS op i 3 forskellige kategorier: Efterkommere, indvandrere, og danskere. Forskellen imellem de 3 grupper er ifølge ICILS, at en indvandrer er født i udlandet og siden indvandret til landet, en efterkommer er en person, der er født i landet, men hvis forældre begge er indvandrere, og en dansker er en person der er født i landet, og mindst en af forældrene er født i Danmark.

Danmarks Statistik bruger en lidt anderledes definition som lyder "En person har dansk oprindelse, hvis han eller hun har mindst én forælder, som både er dansk statsborger og født i Danmark".

For at teste fordelingen af efterkommere opretter vi igen en dummy-variabel for om eleven er efterkommer eller ej. Dummyen er 1, hvis eleven er efterkommer, ellers 0.

Tabel 3: Test af andel af efterkommere i ICILS stemmer med andelen i Danmarks Statistik	
Gennemsnittet fra ICILS: μ_{ICILS}	0,06
Gennemsnittet fra DST: μ_{DST}	0,08
Stikprøvevariansen S^2	0,05
Standardfejlen $s.e$	0,01
t-testværdien: $H_0: \mu_{ICILS} = \mu_{DST}$	-3,95
Kilde: ICILS, DST tabel U1907. Egne beregninger	

Som det fremgår, er værdien fra t-testen numerisk større end 1,96, hvilket betyder at ICILS-testens andel af efterkommere og Dansk Statistiks andel er statistisk forskellige fra hinanden, og derfor må vi sige at de to andele ikke er ens. Dermed vil vi skulle konkludere at vores stikprøve ikke er repræsentativ. Men spørgsmålet er, hvad der har skabt denne forskel, og derfor vil vi undersøge om det er elever der er defineret som indvandrere eller danskere, der har en større andel end hvad de skulle have haft.

Analyse af etnicitet: Indvandrere

For at teste fordelingen af indvandrere laver vi igen en dummy-variabel som angiver om en elev er indvandrer eller ej.

Tabel 4: Test om indvandreres andel i ICILS stemmer med andelen i Danmarks Statistik	
Gennemsnittet fra ICILS: μ_{ICILS}	0,03
Gennemsnittet fra DST: μ_{DST}	0,03
Stikprøvevariansen S^2	0,03
Standardfejlen $s.e$	0,003
t-testværdien: $H_0: \mu_{ICILS} = \mu_{DST}$	-0,05
Kilde: ICILS, DST tabel U1907. Egne beregninger	

Her ser vi at t-testværdien er tæt på 0, og altså mindre end 1,96, således at vi ikke kan afvise vores hypotese om at de to andele er ens. Dermed er det ikke indvandrerne der udgør forskellen.

Analyse af etnicitet: Danskere

Tabel 5: Test om danskeres andel i ICILS stemmer med andelen i Danmarks Statistik	
Gennemsnittet fra ICILS: μ_{ICILS}	0,91
Gennemsnittet fra DST: μ_{DST}	0,89
Stikprøvevariansen S^2	0,08
Standardfejlen $s.e$	0,01
t-testværdien: $H_0: \mu_{ICILS} = \mu_{DST}$	3,62
Kilde: ICILS, DST tabel U1907. Egne beregninger	

Her ser vi at t-testen er større end 1,96, hvilket gør at vi må afvise vores H_0 hypotese om at de to gennemsnit skulle være lige store. Og hermed har vi fundet ud af at den fordeling, som deltagerne i stikprøven er taget ud fra, er forskellig fra fordelingen af elevsammensætningen af 8.-klasserne i den danske folkeskole.

Dog bemærker vi at t-værdierne er "små" (-3,95 og 3,62) hvilket er "lidt" over 1,96. Årsagen til forskellen mellem ICILS og Danmarks Statistik kan ligge i de lettere forskellige definitioner af efterkommere og danskere i henholdsvis ICILS og DST. I ICILS bruges som sagt følgende definition af en dansker: "Students and/or at least one parent born in country of test", mens DST bruger definitionen: "En person har dansk oprindelse, hvis han eller hun har mindst én forælder, som både er dansk statsborger og født i Danmark."¹ Så forskellen kan ligge i om forælderen, der er født i landet, også har statsborgerskab i Danmark. Hvis forælderen ikke har dansk statsborgerskab, vil det i ICILS blive registreret som en dansker, mens det i DST vil blive registreret som en efterkommer. Da vi opererer med små andele, vil selv en lille ændring i en gruppes størrelse betyde at andelen vil blive kraftigt påvirket, og dermed kan vi gå fra en signifikant forskel til en insignifikant forskel (dvs. gå fra en t-værdi større end 1,96 til en t-værdi lavere end 1,96). Derfor vil vi i forhold til elevsammensætningen i ICILS godt kunne sige at den er tilnærmelsesvis repræsentativ. Dette kan vi sige da ICILS' andel af danskere er større end DST's andel. Andelen kan kun blive mindre når vi korrigerer for definitionsforskellen.

¹ <http://www.dst.dk/da/Statistik/emner/indvandrere-og-efterkommere/indvandrere-og-efterkommere.aspx#>

Analyse af kønsfordelingen hos lærerne

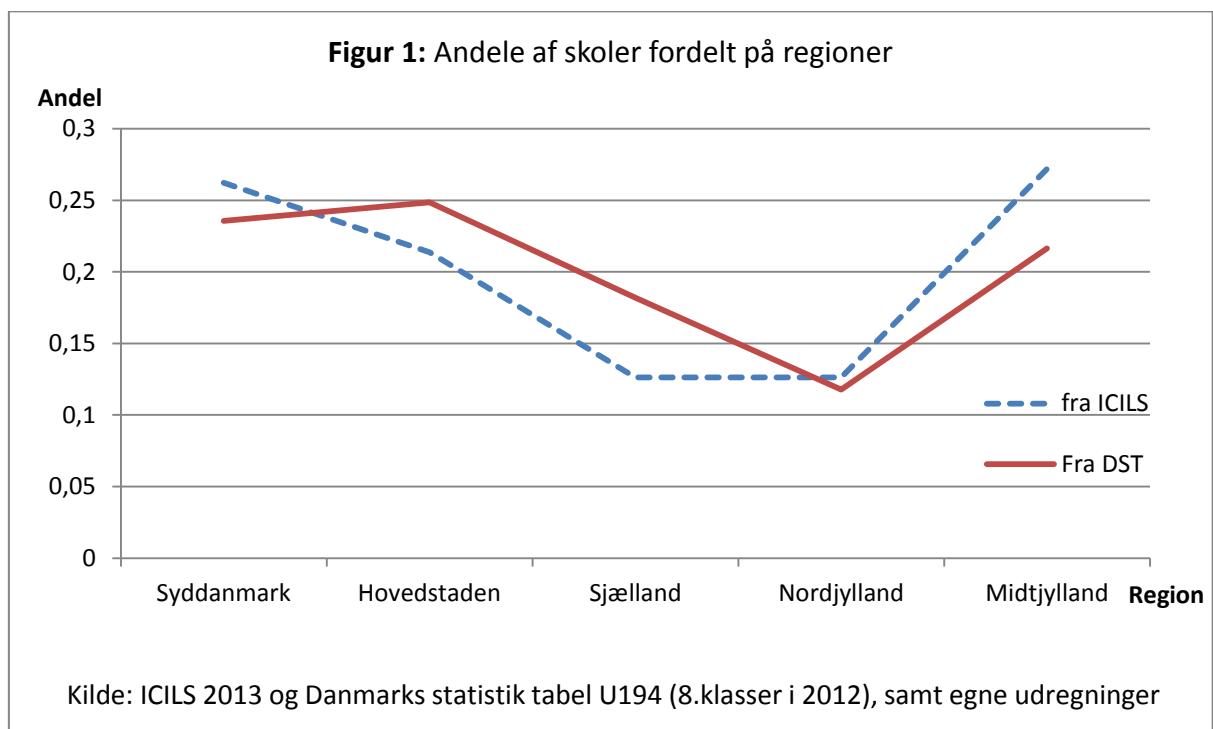
Vi vil nu undersøge kønsfordelingen af de lærere, der har besvaret spørgeskemaet i ICILS. I ICILS-dokumentationen kan man finde kønnet på dem, der har besvaret testen, og kønnet på alle de adspurgte. Vi vil antage at kønnet på alle de adspurgte er repræsentativt for alle lærere i 8. klasse. Denne antagelse bygger på udvælgelsesmetoden stratificering, som nøje udvælger de lærere der er repræsentative for lærere i 8. klasse. Grunden til at vi bruger kønnet fra de adspurgte i ICILS og ikke Undervisningsministeriets statistikker, er at Undervisningsministeriets statistikker ikke giver en klasseopdelt kønsfordeling. Sammenligner vi ICILS-andelen af kvindelige lærere med den samlede andel af kvindelige lærere, vil vi få en bias, da indskolingen har overvægt af kvindelige lærere i forhold til udskolingen. Her er dummyen lavet sådan at hvis observationen er en kvinde, har dummyen værdien 1. Hvis det er en mand, har dummyen værdien 0.

Tabel 6: Test om fordelingen af kvindelige lærere der har besvaret spørgeskemaet, stemmer med de adspurgte	
Gennemsnittet fra ICILS deltagerne: $\mu_{Deltager}$	0,59
Gennemsnittet fra ICILS adspurgte: $\mu_{Adspurgte}$	0,60
Stikprøvevariansen S^2	0,24
Standardfejlen $s.e$	0,02
t-testværdien: $H_0: \mu_{Deltager} = \mu_{Adspurgte}$	-0,19
Kilde: ICILS. Egne beregninger	

I tabel 6 ser vi at andelen af kvinder ud af alle lærerne, som har besvaret ICILS-spørgeskemaet, ikke er statistisk signifikant forskellig fra andelen af kvinder i den adspurgte gruppe. Dermed kan vi ikke afvise at fordelingen er repræsentativ.

Fordelingen af skoler på regioner

Vores stikprøve udgøres udover en stor gruppe elever og lærere af en række skoler, der er fordelt rundt om i landet. Vi er her interesseret i at skolerne er repræsentativt fordelt rundt om i landet. Det undersøger vi ved at sammenligne fordelingen af skoler på regioner i ICILS med tal fra Danmarks Statistik. Ved at se på figur 1 kan vi se at graferne følger hinanden meget godt, undtagen Region Syddanmark og Region Sjælland, som har en mærkbar afstand imellem de to andele.



Vi undersøger om der er signifikant forskel imellem ICILS-stikprøven og de faktiske tal for alle skoler fra DST. Vi anvender igen en dummy-variabel, hvor de enkelte regioner får tildelt en dummy, således at vi kan udregne stikprøvevariansen og t-testen.

Tabel 7: Test af om fordelingen af skoler på regioner i ICILS stemmer med DST					
	Syddanmark	Hovedstaden	Sjælland	Nordjylland	Midtjylland
Gennemsnittet fra ICILS: μ_{ICILS}	0,26	0,21	0,13	0,13	0,27
Gennemsnittet fra DST: μ_{DST}	0,24	0,25	0,18	0,12	0,22
Stikprøve-variansen S^2	0,20	0,17	0,11	0,11	0,20
Standardfejlen $s.e$	0,04	0,04	0,03	0,03	0,04
t-testværdien: $H_0: \mu_{ICILS} = \mu_{DST}$	0,61	-0,86	-1,69	0,26	1,26
Kilde: ICILS, DST tabel U194 (8.-klasser i 2012) og egne udregninger.					

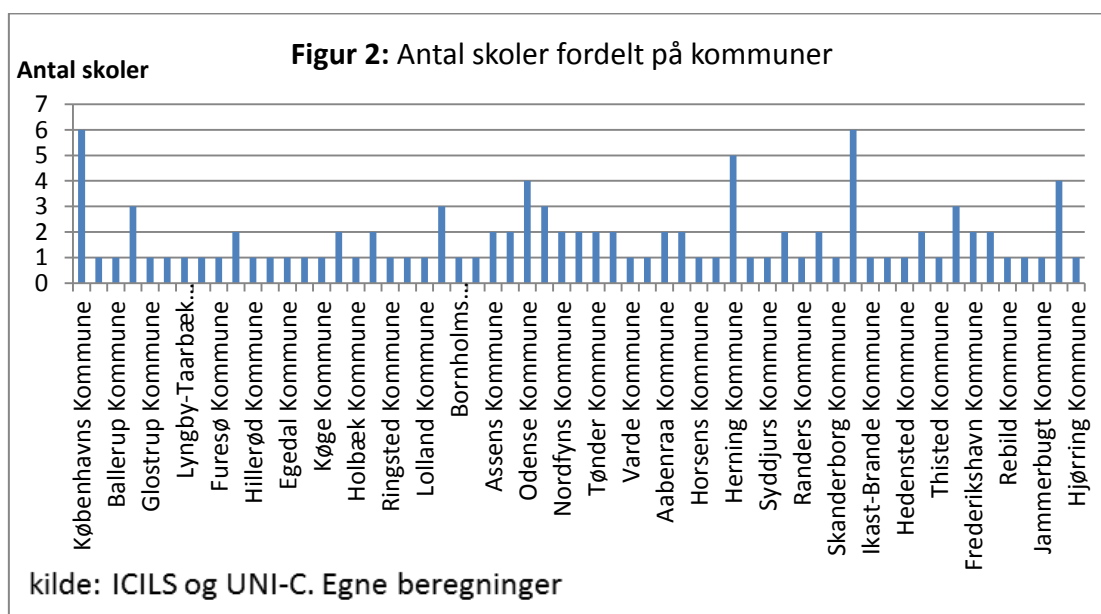
Her i tabel 7 kan vi se at alle t-test er mindre end numerisk 1,96, hvilket betyder at vi i ingen tilfælde kan afvise vores H_0 -hypoteser om at andelen fra ICILS og DST er ens. Vi kan således ikke afvise at stikprøven fra ICILS er repræsentativ.

Analysen af fordelingen af skoler på kommuneniveau

Når vi ser på den geografiske fordeling, er regioner et geografisk meget stort område, og for at sikre os at skolerne er fordelt rundt om i regionerne og ikke fx blot i de større byer, vil vi se på hvor mange kommuner, der har en skole med i ICILS-undersøgelsen. Som omtalt tidligere, har IEA udtrukket 141 *samplede* skoler (og 9 der var lukket på testtidspunktet), 141 *first replacement*-skoler og 141 *second replacement*-skoler, og disse skoler er naturligvis placeret rundt i de forskellige kommuner. Dog vil der være en sandsynlighed for at nogle kommuner ikke bliver udtrukket. Der er et begrænset antal af skoler i for eksempel Læsø, Fanø, Samsø og Ærø Kommune. Derfor har vi set på udtrækningen af skolerne og set på deres kommunale tilhørsforhold. Her fandt vi at ud af de 98 kommuner, der er i Danmark, har 94 kommuner en skole udtrukket, som enten er en *sampled* skole, en *first replacement*-skole eller *second replacement*-skole. Dermed skal vi sammenligne de deltagende skolars kommuner med 94 mulige kommuner.

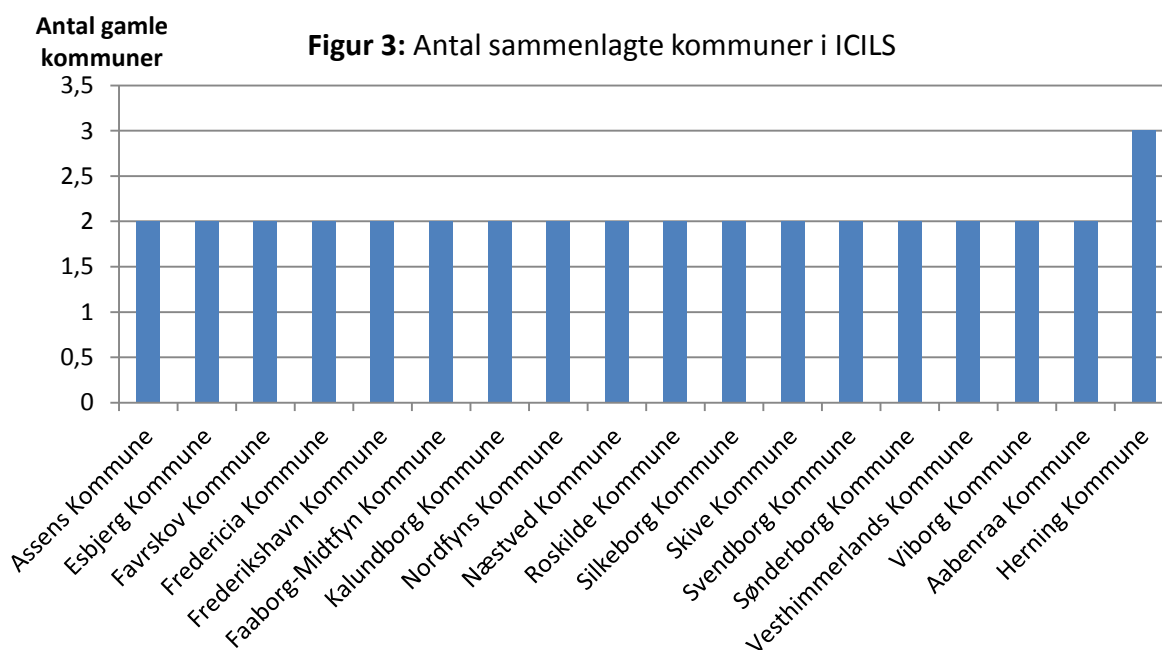
Vi har før nævnt at der deltog 103 skoler, som har en totalvægt, i ICILS-testen, og de fordeler sig på 59 forskellige kommuner. Det vil sige ud af 94 mulige har vi har haft en deltagelse målt på mulige kommuner på $\frac{59}{94} = 62\%$. Dermed er skolerne relativt

jævnt fordelt over landet, og alle skoler er ikke samlet i de store byer, som København, Aarhus, Aalborg og Odense. Antallet af skoler i hver kommune varierer dog, således at der er flest deltagende skoler i de store kommuner:



Et muligt problem ved de nye storkommuner er at skolerne i princippet kunne være placeret i den største by i kommunen, og vi dermed kunne mangle de mindre byer i kommunerne. Denne problemstilling undersøger vi ved at se på antallet af deltagende kommuner under den gamle struktur. Det vil sige at vi splitter de nye kommuner op i de gamle kommuner og ser hvad antallet af de gamle kommuner er. Dette giver os en ide om spredningen af skolerne i de nye kommuner.

Ud fra skoleid'et kan vi finde de gamle kommunenumre, hvorfra vi kan se hvor mange forskellige gamle kommuner, der deltager. Der er 83 gamle kommuner der deltager, hvilket vil sige at der er 19 kommuner, der er lagt sammen i strukturreformen 2007, hvilket giver os en ganske fin fordeling af skolerne inden for de nye kommuner. Hvis vi ser på hvilke gamle kommuner, der er lagt sammen til de nye kommuner, for-deler de sig således:



Kilde: ICILS, og UNI-C. Egne beregninger

Siden der er tre gamle kommuner slået sammen til én ny kommune i Herning, har vi 18 nye kommuner og to eller flere skoler fordelt over et større geografisk område fra de gamle kommuner. Det vil sige at de deltagende skoler i kommunen ligger spredt ud over den nye kommune. Dette er med til at sikre at vi får et mere repræsentativt udsnit af vores stikprøve, da vi har flere forskellige geografiske områder med i ICILS-undersøgelsen.

Konklusion på repræsentativitetstesten

I denne tekst har vi undersøgt om ICILS-undersøgelsen er repræsentativ, selvom ICILS-undersøgelsen for Danmarks vedkommende ikke opfylder IEA's krav. Vi har derfor undersøgt om fordelingen af deltagerne og skoler i ICILS-undersøgelsen følger fordelingen for eleverne i 8. klasse i grundskolen. Vores undersøgelser viser alle, med en enkelt forklarlig undtagelse gældende for andelen af danskere og efterkommere, at andelene af de forskellige grupper ikke adskiller sig signifikant fra hele Danmarks andele. Vi kan således konkludere at ICILS-undersøgelsen efter al sandsynlighed er repræsentativ for Danmark. Derfor vil vi kunne bruge resultaterne fra ICILS-undersøgelsen meningsfuldt, selvom Danmark ikke levede op til de internationale krav fra IEA.

