

Indhold

Figurer	iii
Tabeller	iv
1 Social netværksanalyse	1
1.0.1 Grundlæggende begreber og datastruktur i social netværksanalyse .	2
1.0.2 Relativ risiko og styrken af forbindelser	3
1.0.3 Segmentering og klyngeanalyse	6
1.0.4 Opsummering	9
2 Datamateriale	10
2.1 DISCO	10
2.2 Timeløn	11
2.3 Køn	11
2.4 Ledighed	11
3 Delanalyse 1: Opdeling af arbejdsmarkedet i delmarkeder	12
3.1 Opdeling af arbejdsmarkedet i klynger	13
3.2 Dannelsen af klynger	14
3.2.1 Niveau 2	16
3.2.2 Niveau 3	17
3.2.3 Niveau 4	18
3.2.4 Niveau 5	19
3.3 Den interne mobilitet i klyngerne	21
3.3.1 Klynge 2.78: Børnepasning i private hjem samt turist- og rejseledere .	23
3.3.2 Delmarked 2.66: Omsorgsarbejde i private hjem og plejearbejde på institutioner	25
3.3.3 Klynge 3.24: Transport af varer, ambulanceførere og vagtarbejde	25

3.4	Delkonklusion	26
4	Delanalyse 2: Forskellen på et delmarked og et segment er påvisningen af særegne (specifikke? partikulære) sociale processer indenfor delmarkedet	27
4.1	Indkomstfordeling for alle	28
4.2	indkomstfordelinger på delmarkederne	29
4.2.1	lala	32
4.3	Kønsmforskelle blandt delmarkederne	34
Bilag A	variabelbeskrivelser: løn	35
A.1	Timeløn	36
A.2	Centrale mål	38
A.3	sammenligning med DSTs rapport	39
A.4	relative lønmforskelle	41
A.5	sammenfatning	41
Bilag B	netværksskott	43
Litteraturliste		47
	Bøger og artikler	47
	Online kilder	48
	Danmarks Statistiks Manualer	48

Figurer

1.1	Eksempel på netværk	7
3.1		13
3.2		16
3.3		17
3.4		18
3.5		19
3.6	Intern mobilitet for klyngerne.	22
3.7		24
3.8		25
3.9		25
4.1	timeløn farvelagt efter udvalgte percentiler	28
4.2	Intern mobilitet for klyngerne.	30
4.3		33
4.4		33
B.1	Netværkskort: Intern mobilitet	44
B.2	Netværkskort: Timelønninger	45

B.3	Netværkskort: detaljerede forbindelser/edges	46
-----	--	----

Tabeller

1.1	Eksempel på en adjacency matrice	2
1.2	t	7
3.1	Karakteristika for klyngedannelsen	15
3.2	Sammenligning af 1. niveau og 5. niveau	20
4.1	t	28
A.1	centrale mål for TIMELON, 1996-2009	38
A.2	centrale mål for TIMELON, 1996-2009	39
A.3	Sammenligning med DST rapport	40

1 Social netværksanalyse

I dette speciale tilgår vi arbejdsmarkedet som et netværk af forskellige arbejdsstillinger. I netværket bevæger individer sig mellem forskellige typer af arbejdsstillinger. Det sker når en person går fra at være beskæftiget i en arbejdsstilling til at være beskæftiget i en anden arbejdsstilling.

Arbejdsstillingerne betragtes som noder i et netværk, og personernes bevægelser mellem forskellige arbejdsstillinger bestemmes som forbindelserne i netværket. Det er motiveret teoretisk i afsnit ??? og skal ikke uddybes yderligere her. Formålet med at anskue arbejdsmarkedet som et netværk er at kortlægge beskæftigelsesmønstre på arbejdsmarkedet for at se, hvilke arbejdsstillinger de ledige bevæger sig imellem, og hvilke arbejdsstillinger de ikke bevæger sig imellem. Vi har tidligere beskrevet arbejdssegmenteringsteorier kort, og vi vil her gå i dybden med vores centrale metodiske værktøj, social netværksanalyse, i en specifik udgave, som Toubøl og Larsen har udviklet til analyser af mobilitet. Denne metode har de kaldet Moneca - "*Mobility Network Clustering Algorithm*".

Dette afsnit uddyber og forklarer de centrale antagelser og beregninger i Toubøl og Larsens to artikler om deres nyopfundne metode, samt opridser fundamentale begreber i netværksanalysen efterhånden som de introduceres i forbindelse med gennemgangen af Moneca. Da der er tale om en ny anvendelse af social netværksanalyse, er der ikke meget andet litteratur at benytte sig af end de to forfatteres to artikler, da vi er de første udover omtalte skabere af metoden, der benytter sig af den. Et delmål for dette speciale er derfor at undersøge, hvad Moneca-algoritmen er i stand til, samt være opmærksom på dens eventuelle begrænsninger og mulige fejlbehæftninger. Det er derfor et delmål for dette speciale at evaluere hvorledes Moneca fungerer i relation til videnskabelige mål for reliabilitet og validitet. Dette vil løbende blive diskuteret i specialet, men først skal algoritmens grundlæggende funktioner beskrives.

1.0.1 Grundlæggende begreber og datastruktur i social netværksanalyse

Moneca er en overordnet set en kvantitativ, deskriptiv metode, hvis formål er:

1. at vise tilstedeværelsen, fraværet og styrken af relationer mellem forskellige grundkategorier af interesse.
2. at benytte et sæt kriterier baseret på centrale mål indenfor social netværksanalyse til at slå disse grundkategorier sammen i større grupper, herfra kaldet *segmenter*. Deraf navnet Moneca: *Mobility Network Clustering Algorithm*.

Grundkategorierne bestemmes ud fra det empiriske formål. Toubøl og Larsen har benyttet Moneca til at se på den sociale mobilitet mellem forskellige beskæftigelseskategorier indenfor hele det danske arbejdsmarked. Relationerne defineres som skift fra et arbejde indenfor én beskæftigelseskategori til arbejde indenfor en anden beskæftigelseskategori. Herefter benyttes Moneca-algoritmens 2. skridt til at undersøge, hvorledes denne datadrevne inddeling af arbejdsmarkedet i segmenter kan give ny indsigt i klasseinddelingen.

Det centrale er derfor hvilke grundkategorier, der benyttes, samt hvad der tæller som en relation mellem to grundkategorier. Toubøl og Larsen kommer selv med andre forslag til mulige anvendelser af Moneca, eksempelvis kortlægning af klassemobilitet gennem ægteskaber (Toubøl m.fl. 2013, s. 27). Her ville grundkategorierne blive bestemt ud fra en given klasseinddeling af interesse, og relationerne mellem disse klasser defineres som ægteskaber. I netværksterminologi betegnes kategorierne som *noder*, mens relationerne mellem noderne betegnes som *edges*, eller på dansk: forbindelser.¹ I vores speciale ligger vi derfor konceptuelt helt lig Toubøl & Larsen, når vi definerer vores noder som beskæftigelseskategorier, og edges som skift mellem disse kategorier². Vi har det til fælles, at vi ser på en social mobilitetstabel, der viser skift fra beskæftigelseskategorierne i rækkerne til beskæftigelseskategorierne i kolonnerne. Denne er illustreret i tabel 1.1 ved hjælp af empirisk data fra vores mobilitetstabel.

Tabel 1.1: Eksempel på en adjacency matrice

	Til: Tandlæge	Til: Folkeskolelærer	Til: Pædagog	Til: Automekaniker
Fra: Tandlæge	264	0	0	0
Fra: Folkeskolelærer	0	6148	288	6
Fra: Pædagog	0	454	9308	0
Fra: Automekaniker	0	13	13	1861

¹ Det er ofte en udfordring at oversætte de tekniske termer fra netværksanalyse til dansk. Hvor det ikke har været muligt at benytte et tilstrækkeligt unikt oversat dansk begreb, har vi derfor beholdt de engelske termer.

² Omend forskellen i populationsdefinition - ledige fremfor alle beskæftigede - har stor betydning for den konkrete operationalisering af begreberne, hvilket vil blive behandlet i de efterfølgende kapitler.

I netværkstermer kaldes sådan en tabel for en *adjacency matrice*, da den har samme udfald i både rækker og kolonner, og datamatricen derfor er kvadratisk (Scott 2000, s. 55). I vores tilfælde skal det tolkes sådan, at rækkeudfaldene er de beskæftigelseskategorier, man kom *fra*, og kolonneudfaldene er de beskæftigelseskategorier man er gået *til*, efter en ledighedsperiode. Det ses at 264 personer på et tidspunkt har skiftet fra tandlæge til tandlæge i løbet af de 14 år, og ikke har skiftet til nogle af de andre tre erhverv. Skift fra samme beskæftigelse til samme beskæftigelse som fra tandlæge til tandlæge, noteres grundet den kvadratiske form langs diagonalen i en adjacency matrice, og denne har derfor en særlig status. Vi har at gøre med en *retningsbestemt, vægtet adjacency matrice*, hvilket er den mest komplicerede form for data i social netværksanalyse (ibid., s. 61). Med vægtet skal forstås, at de enkelte celler ikke bare tilkendegiver en binær opdeling i tilstedeværelse eller fravær af en forbindelse, eksempelvis mellem folkeskolelærer og automekaniker, men at den også angiver en værdi for styrken af denne forbindelse. At den er retningsbestemt betyder at matricen ikke er symmetrisk langs diagonalen, da bestemmelsen af retningen sker ud fra hvorvidt man ser på et udfald rækkevis eller kolonnevis. Samt - i en vægtet matrice - kan styrken naturligvis også variere alt efter om man befinder sig over eller under diagonalen. I tilfældet automekaniker ses det, at hvis man i matricens nederste del kigger på Automekaniker, er relationen til tandlæge 0, det vil sige fraværende, mens den til folkeskolelærer og pædagog har en styrke på 13 i begge tilfælde. I selve diagonalen, det vil sige *den interne mobilitet*, er styrken forventeligt langt højere. Hvis man ser på automekaniker i matricens øverste del, det vil sige kolonnevis, er der et fravær af forbindelse til tandlæge, en styrke på 6 til folkeskolelærere, og et fravær af forbindelse til pædagoger. Det viser hvad der menes med en retningsbestemt, vægtet adjacency matrice: Styrken i mobiliteten fra automekaniker til folkeskolelærer er på 13, mens den fra pædagog til automekaniker er på 6, altså halvt så kraftig.

1.0.2 Relativ risiko og styrken af forbindelser

Indtil videre har vi kun talt om styrken af forbindelser som antallet af skift, altså en absolut enhed. Det er imidlertid ikke særlig retvisende, da styrken af forbindelsen dermed ikke står relativt til størrelsen af den kategori, den udspringer af. I 1996 var der eksempelvis 5.207 tandlæger, mens der var 53.676 beskæftigede pædagoger, altså omtrent ti gange så mange. Et skift fra tandlæge til en anden profession bør derfor vægtes højere end et skift fra pædagog til en anden profession. Her kommer konceptet relativ risiko ind i billedet, som en

ratio mellem to proportioner (Agresti og Finlay 1997, s. 244, 271):

$$\frac{\pi_A}{\pi_B} \quad (1.1)$$

Den relative risiko (RR) fortæller os hvad chancen er for at begivenhed B sker, relativt til begivenhed A. Begivenhed A kender vi. Det er den simultane sandsynlighed for udfald x og udfald y i de to stokastiske variable I og J () (Malchow-Møller og Würtz 2003, s. 41), givet ved sandsynlighedsfunktionen:

$$f(i, j) = P(I = i, J = j) \quad (1.2)$$

I eksemplet fra tabel 1.1 er den simultane sandsynlighed for at have skiftet fra folkeskolelærer til pædagog: $\frac{45}{45} = 0,45 = \pi_A$. Begivenhed B er til vores formål³ karakteriseret som *den forventede værdi*. Det er derfor en helt igennem teoretisk størrelse - dens værdi er udelukkende bestemt ud fra teoretiske forventninger til hvad chancen *bør* være for at begivenheden indtræffer.

Toubøl & Larsen foreslår at det forventede udfald π_B bør være uafhængighed mellem $f(I)$ og $f(J)$, hvilket betyder at de marginale sandsynligheder for I og J giver den simultane sandsynlighed (ibid., s. 44):

$$f(i, j) = f_I(i) * f_J(j) \quad (1.3)$$

Hvis vi ser arbejdsmarkedet som et frit marked, er der således tale om den perfekte markedstilstand; alle søger til og fra jobs som man bør forvente ud fra fordelingen af de givne jobs, uden at være betinget af hvilket job man kommer fra eller går til (ibid., s. 43). Lad π_{ij_B} være udtrykket for den simultane sandsynlighed under uafhængighed for den i de række og den j de kolonne, det vi i starten benævnte som begivenhed B. π_{ij_A} er den observerede simultane sandsynlighed vi rent faktisk observerer, det vil sige $f(i, j)$. Den relative risiko, beregnet ud fra de observerede marginale sandsynligheder, er derved givet ved udtrykket

$$\frac{\pi_{ij_A}}{\pi_{ij_B}} = \frac{f(i, j)}{f_I(i) * f_J(j)} = RR_{(i, j)_{observeret}} \quad (1.4)$$

Det betyder at hvis den relative risiko for celle (i, j) er lig 1, er formel 1.3 sand, og dermed eksisterer der uafhængighed mellem job i og job j . Hvis RR er under 1, er sandsynligheden

³Indenfor medicinsk forskning vil relativ risiko typisk blive brugt sammen med oddsratio-værdier til at bestemme risici mellem en gruppe patienter tildelt en ny medicin samt en kontrolgruppe. Relativ risiko bruges derfor typisk til at vurdere forskelle mellem to konkrete grupper. Toubøl og Larsen benytter det i stedet som det bruges indenfor hypotesetest som det kommer til udtryk i eksempelvis χ^2 -testen (Malchow-Møller og Würtz 2003, s. 205).

for jobskifte mellem de to jobtyper mindre end forventet givet uafhængighed, hvorimod en RR på over 1 betyder at det er mere sandsynligt end under uafhængighed. Toubøl & Larsen definerer derfor en forbindelse mellem to typer job som uafhængighed eller derover, altså $RR \geq 1$ (Toubøl m.fl. 2013, s. 9). Man kan tolke dette som udtryk for, at hvis $RR < 1$, eksisterer der barrierer i tilgangen fra det ene job til det andet, hvorimod $RR \geq 1$ betyder at en sådan barriere ikke eksisterer.

Indtil videre har vi beregnet den simultane sandsynlighed under uafhængighed ud fra de marginale fordelinger af de simultane sandsynligheder observeret i krydstabellens celler. I tilfældet med mobilitet blandt ledige afspejler det dermed hvor mange mobile, der skulle være i hver celle, ud fra hvor mange mobile der i det hele taget er. Det kan af sociologiske årsager være en problematisk antagelse. Hvis et bestemt job ikke fordrer en særlig stor mobilitet i det hele taget, bør det tages med i beregningen af den forventede værdi, hvilket ikke sker ved at benytte de marginale sandsynligheder for de mobile - fordi det netop kun beregnes ud fra dem, der rent faktisk er mobile. Hvis der blandt tandlæger eksempelvis er 5 % mobile, mens der blandt tjenere er 20 %, er det ikke hensigtsmæssigt at tandlægenes 5 % tæller for 100 %, ligesom tjenernes 20 % mobile nu også tæller som 100 %. Det er ikke desto mindre konsekvensen af at benytte de observerede marginale sandsynligheder fra formel 1.4. Vi vil i stedet have at tjenernes mobilitet på 20 % skal tælle fire gange så meget som tandlægenes 5 % i beregningen af deres respektive π_{ij_B} . Toubøl & Larsen benytter derfor de marginale sandsynligheder, ikke fra den observerede fordeling, men fra den oprindelige fordeling blandt alle beskæftigede, hvorfra de mobile "er trukket fra", for at blive i jargonen. Dermed bliver forventningen til indholdet i den enkelte celle, π_{ij_B} , udtrykt i form af af de marginale sandsynligheder for det samlede antal beskæftigede. Dermed forhindres det at tandlægenes 5 % og tjenernes 20 % begge tæller for 100 %, og i stedet tages der udgangspunkt i forventninger, der er informeret af vores viden om den samlede fordeling blandt alle tandlæger og tjenere. Vi laver derfor følgende omskriving af formel 1.4, hvor variablene K og L står for de marginale sandsynligheder for henholdsvis række- og kolonnefordelingerne for alle beskæftigede. Det vil sige de fordelinger, som fordelingerne af variablene I og J er trukket fra K og L , og vi kan derfor opskrive følgende formel:

$$\frac{\pi_{ij_A}}{\pi_{ij_B}} = \frac{f(i, j)}{f_K(k) * f_L(l)} = RR_{(i, j)_{population}} \quad (1.5)$$

Der er yderligere en fordel ved at benytte de marginale sandsynligheder for populationen, og det er at man kan tage højde for strukturelle ændringer i beskæftigelse over tid. Det vil sige, hvis givne fagkategorier skrumper eller vokser, vil π_{ij_B} i den modificerede form, baseret på $f_K(k) * f_L(l)$, tage højde for dette. For at forstå det skal man huske på, at selvom K og L er samme variabel, er der forskel i fordelingen hos de to, da de er forskudt med et

enkelt år. I vores tilfælde er række-variablen udtryk for tidsperioden 1996 til 2008, da perioden, vi beskriver, strækker sig fra 1996 til 2009. Man kan derfor ikke inkluderes med et "start-job" i 2009, og det omvendte gælder for kolonne variablen: Man kan ikke ende i et job i 1996, da vi ikke ved hvilket job man kom fra, i 1995. Det betyder at de to marginale sandsynligheder for henholdsvis K og L er forskudt med et enkelt år, og ændringer mellem disse to år medregnes i forventningerne til den enkelte celle π_{ijB} .

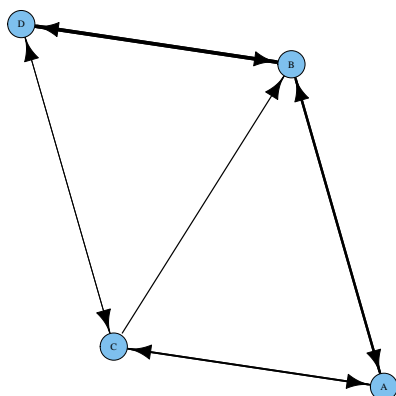
Vi har hermed redegjort for principperne bag styrken af forbindelser i den retningsbestemte, vægtede mobilitetstabel, som Moneca er baseret på. Det centrale er, at styrken af forbindelser er et teoretisk funderet valg. π_{ijB} bestemmes ud fra en betragtning om hvilket mønster, de observerede frekvenser i cellerne bør følge. Det vil sige, 1) fra hvilken population mener man at éns egen population er en subpopulation af, og 2) hvilket mønster i denne populations fordeling er det interessant at vurdere som udgangspunkt for om forbindelser eksisterer eller ej. Vi har her valgt, ligesom Toubøl & Larsen, at benytte de marginale sandsynligheder i populationen af alle beskæftigede til at vurdere hvilket mønster, der burde eksistere under uafhængighed. Det gør vi, da vi, jævnfør ovenstående diskussion af mobilitet blandt de hypotetiske tandlæger og tjenere, mener at den bedst viser det mønster der burde eksistere såfremt ingen mobilitetsbarrierer eksisterede, samt tager højde for erhvervsstrukturelle forhold, der bestemmer mønstrene også hos vores subpopulation af ledige.⁴.

1.0.3 Segmentering og klyngeanalyse

At kortlægge forbindelserne mellem beskæftigelses kategorier, hvori mobiliteten er så stor at der ikke kan siges at eksisterer barrierer, baseret på definitionen af barrierer fra foregående afsnit, er i sig selv interessant. Moneca har, som beskrevet på side 1.0.1, også til formål at benytte centrale mål fra social netværksanalyse til at vurdere hvilke beskæftigelses kategorier, hvori styrken af forbindelserne antager en sådan karakter, at de kan siges at tilhøre et segment, eller klynge. Vi vil komme nærmere ind på de sociologiske argumenter for at lave denne opdeling i afsnit ??, mens dette afsnit primært omhandler metoden der benyttes til denne segmentering.

⁴Vi kunne også have baseret vores forventninger til cellens indhold på eksempelvis mønstrene hos alle mobile, fremfor alle beskæftigede, ud fra en betragtning om at de forventede mønstre blandt ledige burde følge denne fordeling under uafhængighed. Det kunne vi dog ikke se nogle gode argumenter for. Ved at benytte denne alternative fordeling ville vi kun introducere en række usikkerheder omkring hvilke sociale strukturer man prøvede at beskrive. Man kan se det sådan at vi dermed ville benytte mønstrene hos mobile, der i forvejen er trukket fra populationen af alle beskæftigede. Det ville dermed introducere et tredje led af strukturelle fordelingsmekanismer, hvilket kunne være hensigtsmæssigt, såfremt vi mente at vores population af ledige rent faktisk var en subpopulation (af subpopulationen) alle mobile. Det har vi dog intet belæg for at mene, da populationen af alle mobile i langt de fleste tilfælde er kendetegnet ved at gå mere eller mindre direkte fra et job til et andet. Derfor virkede mønstret under uafhængighed blandt alle beskæftigede mere hensigtsmæssigt.

Figur 1.1: Eksempel på netværk



Tabel 1.2: Adjacency matrice til netværk

	A	B	C	D
A	-	-	-	-
B	4	-	-	-
C	3	2	-	-
D	0	6	1	-

At lede efter sammenhængende subgrupper i et netværk, er en af de helt fundamentale opgaver i social netværksanalyse (Touboel2015), og en række mål indenfor eksisterer for at afgrænse og identificerer sådanne subgrupper, såsom *k-cores*, *m-cores*, *klike* og *klan*. Problemet med disse mål er at de beregnede klynger ikke er diskrete, men overlapper. Det er naturligvis ikke et problem i sig selv. Men for at reducere kompleksitet i komplicerede og og tæt forbundne netværk, så forskelle og ligheder fremtræder tydeligt, kan det være ønskeligt. En anden fordel ved Moneca er muligheden for at håndtere vægtede netværk, hvilket ovenstående koncepter traditionelt er dårlige til. Fordelene ved at benytte social netværksanalyse til at bestemme klynger, er at denne tager udgangspunkt i klyngedannelsen ud fra et princip om *forbindelse*, fremfor *enshed* (Touboel2015). Det betyder kort sagt om man ser på en gruppe relationelt i forhold til andre grupper, eller om nodernes tilhørsforhold bestemmes ud fra karakteristika, de har til fælles. Det er en central forskel i videnskabsteoretisk grundlag, hvilket vi kommer nærmere ind på i afsnit ???. Monecas algoritme baserer sig på det netværksteoretiske mål kaldet en *klike*. En klike er i netværkstermer defineret som en subgraf, hvor alle noder er forbundet med alle andre noder i subgrafen, og som samtidig ikke er indeholdt i andre subgrafer (Scott 2000, s. 112). Moneca kan ses som en algoritme, hvis formål er at afgøre de tvivlsspørgsmål, hvor en node kan tilhøre flere forskellige klike, hvilket ofte er tilfældet i komplekse netværk med mange forbindelser mellem noderne. Figur og tabel ?? er en simpel illustration af et sådant tvivlsspørgsmål om tilhørsforhold. Moneca ville i dette tilfælde komme frem til at netværket i figur ?? består af klyngerne [A|C] og [B|D]. Det sker ud fra følgende procedure:

-
1. Først lægges de to kraftigst forbundne noder sammen. Det vil her sige $[B|D]$, hvor styrken af forbindelsen er seks.
 2. Derefter gør det den samme med de to næstmest forbundne noder, $[A|B]$, der har en styrke på 4. Eftersom $[B]$ allerede er en del af den foreløbige klynge $[B|D]$, spørger Moneca om det er muligt at indlemme $[A]$ i den allerede etablerede foreløbige klynge. Det kan ikke lade sig gøre, da $[A|B]$ ikke er forbundne.
 3. Moneca går derfor videre til den tredje stærkeste forbindelse, $[A|C]$. Hverken $[A]$ eller $[C]$ er en del af en foreløbig klynge, og de lægges derfor sammen.
 4. Den fjerde stærkeste forbindelse er $[B|C]$. $[B]$ og $[C]$ er allerede parret med henholdsvis $[D]$ og $[A]$, og Moneca beregner derfor om $[A|B|C|D]$ udgør en klike, altså alle er forbundne med hinanden. Da det ikke er tilfældet, stopper Moneca og har dermed etableret klyngerne $[A|C]$ og $[B|D]$.

Det vil sige at Moneca ikke etablerer nogen af de maksimale kliker, $[A|B|C]$ eller $[B|C|D]$. Den semistrenge stopregel om klike-tilhørsforhold indenfor klynger betyder desuden, at man med en vis sandsynlighed kan sige at en klynge rent faktisk *er* en samlet størrelse - da den kun kan dannes hvis alle noderne har forbindelse til hinanden, hvilket eksemplet illustrerer: De to maksimale kliker etableres netop ikke, og hvis de gjorde, ville troværdigheden af grænseadragningen mellem klyngerne være langt mere tvivlsom.

For at opsummere ovenstående i mere generelle termer: Moneca starter med at slå noderne med de to mest intense forbindelser sammen til en foreløbig klynge, og går derefter videre til den næstmest intense forbindelse. Hvis en eller begge af noderne i de efterfølgende forbindelser allerede har forbindelse til en tredje eller fjerde node, vurderer Moneca, om denne indgår i en klike med den allerede etablerede klynge. Det er her vigtigt at understrege, at denne vurdering ikke er baseret på styrken af forbindelserne, men udelukkende om der eksisterer en forbindelse⁵. Moneca fortsætter med denne procedure i prioriteret rækkefølge fra de mest intense forbindelser til de mindst intense, indtil alle noder er placeret i kliker med andre noder, der endnu ikke er "optaget" af en mere intens forbindelse. Kriteriet for, om Moneca tillader at slå de tre noder sammen, er om de tilsammen former en klike, altså alle er forbundet til hinanden. Hvis de ikke er det, går den videre til den næstmest intense forbindelse, og fortsætter med at forbinde noder indtil der ikke længere kan etableres flere kliker. Kriteriet om at allerede-etablerede klynger kun kan lægges sammen med nye noder, hvis disse indgår i en klike med alle medlemmer af klyngen, er den stop-regel, der gør at

⁵Man kan forestille sig en fremtidig version af Moneca foretage en mere avanceret vurdering i disse tvivls-spørgsmål, hvori styrken af relationen kunne indgå som vurderingsgrundlag.

Moneca ikke bare ender med at etablere det redundante stykke information, at hvert enkelt komponent⁶ er en klynge i sig selv (Touboel2015).

Efter denne første segmentering er Moneca beregnet til at gentage proceduren indtil det ikke længere er muligt at skabe større segmenter, fordi den førømtalte klike-regel forhindrer det. Det er vigtigt at fremhæve, at Moneca i de efterfølgende klyngeinddelinger baseret den på side ?? beskrevne procedure, *ikke* længere tager de oprindelige, interne forbindelser mellem grundkategorierne i betragtning, hvis disse er blevet lagt sammen med andre grundkategorier. I den efterfølgende klyngeinddeling vil de interne forbindelser mellem noderne på et lavere niveau ikke indgå i beregningerne i forbindelserne mellem de nyskabte noder. Det vil sige at klike-reglen kun tager højde for forbindelser til noder *på det niveau noderne befinder sig på, og ikke forbindelserne på de lavere niveauer*. Når man vurderer kvaliteten af klyngerne på de højere niveauer, er det derfor centralt at se på en række standardmål for klyngens interne forbindelser, for at vurdere rimeligheden af at vurdere klyngen som en samlet størrelse. Det er dette spørgsmål vi nu afslutter gennemgangen af Moneca med.

det handler om at specificere sit netværk således, at man kan finde frem til de relationer, som viser fælles standarder for aktører, og sandsynliggør gensidig forpligtigelse overfor hverandre Laumann, Marsden og Prensky 1983, s. 33).

1.0.4 Opsummering

⁶Et komponent betyder en subgraf, hvor alle noder er forbundne gennem stier. Et netværk kan således bestå af flere komponenter, der per definition ikke er forbundne (ellers ville de være en del af samme komponent), samt noder uden forbindelse til andre noder (*isolates*) (Scott 2000, s. 100).

2 Datamateriale

2.1 DISCO

Arbejdsmarkedssegmentering måles på jobskifte

Datamaterialet bygger på Danmarks Statistik

De primære variable som bruges er i den sociale netværksanalyse er DISCO-88 og ANXSFREM

Danmarks Statistik bruger DISCO-fagklassifikationskoder til at inddele arbejdsmarkedet i forskellige jobfunktioner, som afhænger forudsætningerne for udførelsen af et job og typen af job. Klassifikationen opdeler arbejdsmarkedet i 492 et fircifret niveau, som fx omfatter lægearbejde, sygeplejearbejde, pædagogisk arbejde og snedkerarbejde. Det er disse jobfunktioner, som ligger til grund for dette speciale.

DISCO har et væsentlig databrud for i 2010, hvor DISCO-88 ændres til DISCO-08. Derfor anvendes der udelukkende data fra perioden 1996 til 2009. Der bruges alene data fra DISCO-koder, som har mere end xxx ansatte i perioden 1996 til 2009. Baggrunden er, at der er for få personer til at lave en meningsfuld social netværksanalyse.

Analyseudvalget består af X.XXX.XXX observationer med X.XXX.XXX personer i alderen 16 til 69 år. som skifte job fra et job til et andet job i det efterfølgende år. Til at lokalisere dem der har lavet et reelt jobskifte er der anvendt ANXSFREM.

Anayseudvalget spænder fra XXX.XXX personer i 1997 til XXX.XXX i 2008. Med operationaliseringen sket der et bortfald i dem, som ikke har haft meningsfulde DISCO-koder. Frafaldet er på XX procent. Frafaldet har dog ikke betydning for løn.

2.2 Timeløn

Jeg benytter indkomstvariablen timeløn. I bilag A beskriver jeg variablen samt den datarensning jeg foretager, samt sammenligner den med de andre mulige indkomstvariable og DSTs egne mål for indkomst for forskellige erhvervsgrupper. De væsentlige pointer fra bilaget er:

- Timelønsvariablen er det bedste skøn på indkomst, og er ganske nøjagtigt selv på mit detaljerede DISCO-niveau.
- Medmindre andet siges, er timelønninger *et gennemsnit* af lønningerne i hele perioden 1996-2009. I beregningen af dette gennemsnit er perioden 1996-2008 inflationskorrigeret til 2009 priser.
- Månedslønninger beregnes ved at gange timelønningen med 160,33. Det er DSTs metode (DST 2009, s. 32).

Når man undersøger sit eget forslag til arbejdsmarkedsstruktur skal man være opmærksom på, at det ikke kan valideres med ved hvad Boje kalder "outcomes af sociale processer," såsom arbejdsløshed, hyppighed i jobskift og, ikke mindst: lønforskelle (Boje 1985, s. 28). Det er indkomst og indkomstforskelle mellem delmarkederne, jeg vil gennemgå nu.

I netværksmetodologiske termer er Laumann, Marsden & Prensky inde på det samme, når de advarer imod at validere et socialt netværk baseret på de selvsamme attributter, der er brugt til at konstruere det (Laumann m.fl. 1983, s. 29). Løn er ikke brugt til at konstruere mit netværk, men må siges at være så tæt et outcome af beskæftigelse, at det ikke fremstår eksternt i Lauman et als (skriver man det sådan? #todo) forstand. Lønniveauerne på de af Moneca skabte *delmarkeder* er derimod interessante for validiteten, hvis vi accepterer det som et *kriterie-relateret validitetsmål*, der gør det muligt at validere den interne struktur og konsistens i et klasseskema (Oesch 2006, s. 94)¹. Udover denne mere metodologisk nødvendige validering af mit bud på en segmenteringsstruktur, er lønninger *i sig selv* interessante, da det er direkte relaterbart til arbejdsmarkedets struktur, og utvivlsomt er den mest benyttede indikator for social status og magtposition i den sociale struktur Oesch s. 95, (Weber, Andrade, Boje, Marx, Wright, Goldthorpe, Harrits, Scott - find dem og skriv dem ind #todo.)

2.3 Køn

2.4 Ledighed

¹ Jeg laver ikke et klasseskema, men er tæt nok på til jeg synes det giver mening at benytte Oeschs validitetskriterie.

3 Delanalyse 1: Opdeling af arbejdsmarkedet i delmarkeder

Dette kapitel har til formål at svare på det første forskningsspørgsmål:

Forskningsspørgsmål
1.
Er der en opdeling af arbejdsmarkedet for arbejdstagere i delmarkeder, hvor mobilitet indenfor delmarkederne er hyppig, og mellem delmarkederne sjælden?

Dette forskningsspørgsmål vil jeg besvare ved at vise resultatet af min analyse - en opdeling af arbejdsmarkedet for arbejdstagere i delmarkeder. Spørgsmålet er såvidt der er en opdeling, om opdelingen afspejler de få delmarkeder, der er omdrejningspunkt i amerikanske regi (Piore 1980; Gordon m.fl. 1982) eller flere små delmarkeder, som andre danske undersøgelser enten har teoretiseret eller empirisk påvist (Boje 1985; Toubøl m.fl. 2013).

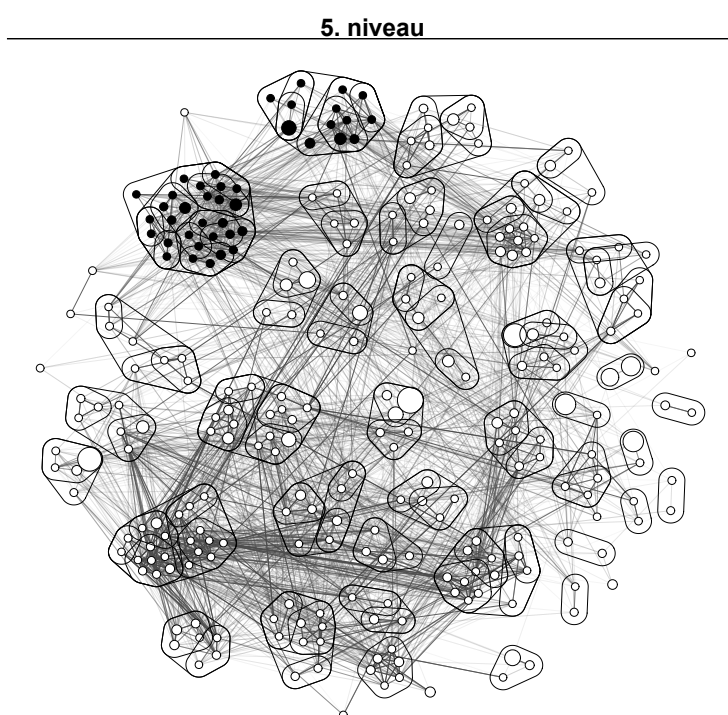
Jeg vil derfor først vise dannelsen af delmarkederne med Moneca-algoritmens klynge-dannelse.

Det er i sig selv er ikke svar på forskningsspørgsmålet. For at vurdere, om der er en reel opdeling af det danske arbejdsmarked i delmarkeder, hvor mobilitet indenfor delmarkederne er hyppig, og mellem delmarkederne sjælden, så er det nødvendigt at vurdere kvaliteten af denne opdeling - er opdelingen stærk eller svag?

Derfor vil jeg se på jobmobiliteten mellem delmarkederne, hvor for at vurdere hvor stærk denne opdeling er.

3.1 Opdeling af arbejdsmarkedet i klynger

Min analyse viser 54 klynger på det danske arbejdsmarked, og er det endelige netværksskott, portrætteret i figur 3.1 som et netværk af 273 noder opdelt i 54 klynger. De 273 noder dækker over de i kapitel ?? beskrevne erhvervsgrupper. Disse erhvervsgrupper bindes sammen af forbindelser, som angiver jobskift mellem forskellige erhvervsgrupper i perioden 1996 til 2009. Jo flere jobskifte, der er mellem erhvervsgrupper, jo kraftigere er forbindelsen. Eksempelvis er forbindelserne kraftige fra 9312: Jord- og kloakarbejde samt andet anlægsarbejde til 9313: Medhjælp ved bygningsarbejde, da 1264 personer skiftede fra jord- og kloakarbejde til medhjælp ved bygningsarbejde i perioden.



Figur 3.1

Klyngerne er sammensat med Moneca-algoritmen. Klyngedannelsen sker ved at aggregerer erhvervsgrupperne i klynger, baseret på deres mobilitetsmønstre. Aggreringen af klynger sker ud fra beslutningsreglerne beskrevet i kapitel 1, hvor sammenlægningen af klynger foregår på stadigt højere niveauer - det vil sige, klynger slås sammen med andre klynger og så fremdeles - indtil der ikke længere er basis for yderligere sammenlægninger.

Om klyngerne også er *delmarkeder*, kommer i min definition an på jobmobiliteten, for-

stået i netværkstermer. For at der er tale om et delmarked, skal der ifølge Boje været meget mobilitet indenfor delmarkedet, og barrierer for jobmobilitet mellem delmarkeder. Derfor bliver jobmobilitet tolket sådan, at hvis jobmobiliteten er "høj" internt i klyngerne - fra nu af kaldet *den interne mobilitet* - har vi at gøre med et delmarked.

Udover det, tilføjer jeg to yderligere krav, der er centrale mål indenfor social netværksanalyse: *Densitet* og *den maksimale stilængde* (Scott 2000, 68f). Disse to mål siger noget om hvor godt de enkelte klynger "hænger sammen" internt. Disse to mål skal undgå den situation, at den interne mobilitet kan være høj grundet høj intern mobilitet i noderne selv, og ikke i klyngen som sådan. Samt tage højde for inklusionen af noder i klynger, hvor de er blevet inkluderet ved en fejl, det vil sige: Deres inklusion i en given klynge skyldes forbindelse til enkelte erhvervsgrupper, og ikke fordi de er velintegrerede i klyngen som sådan. Densitet og maksimal stilængde for klyngerne har primært en teknisk karakter. Fremlæggelsen og diskussionen af dem findes i bilag ?? . På baggrund af disse to mål tages 9 klynger til revision grundet disse kvalitetsmål. Ud af disse 9 splittes 4 klynger op. Dette uddybkes også i bilag bilag ?? .

Tilsammen er jobmobilitet, *densiteten* og *den maksimale stilængde* for en række noder med til at sige, om vi har at gøre med et delmarked eller ej. For at parafrasere Weber, så består et delmarked af de arbejdsmarkedssituationer, hvor mobilitet er nem og typisk ¹(Weber 1978, s. 302).

Til forskel fra densitet og maksimal stilængde har jobmobilitet en klar teoretisk baggrund, såvel som den er baggrund for klyngedannelsen i det hele taget. Det er denne klyngedannelse, jeg nu vil gennemgå. Det skulle gerne lette fortolkningen af det endelige netværksskott for læseren, samt give svar på om klyngerne kan betegnes som egentlige delmarkeder.

3.2 Dannelsen af klynger

Moneca-algoritmens klyngedannelse slutter på det 5. niveau. I første niveau starter vi med de 273 noder. På femte og sidste niveau er de 273 noder blevet aggreret til 54 klynger.

Aggreringen fra niveau 1 til niveau 5 fremgår af tabel 3.1. Her vises ændringerne i jobmobiliteten fra niveau til niveau, som den foregår internt i klyngerne og *eksternt* mellem klyngerne.

¹Parafrasen er over citatet "A »social class« makes up the totality of those class situations within which individual and generational mobility is easy and typical." (Weber 1978, s. 302).

Tabel 3.1: Karakteristika for klyngedannelsen

Niveau	1. niveau	2. niveau	3. niveau	4. niveau	5. niveau
Antal segmenter	273	114	69	56	54
Reduktion i antal segmenter	-	139%	65%	23%	4%
Intern mobilitet (gns.)	68%	75%	77%	79%	79%
Ekstern mobilitet (gns.)	32%	25%	23%	21%	21%
Mobilitet i alt	100%	100%	100%	100%	100%
Forøgelse i intern mobilitet i procent	-	10%	4%	2%	0.2%
Forøgelse i intern mobilitet i procentpoint	-	7%	3%	1%	0.2%

Det første iøjnefaldende er at den interne mobilitet på det (uaggregerede) niveau 1 er højt i sig selv: 68 %. Det indikerer at en væsentlig mængde skift foregår på et meget detaljeret DISCO-niveau. Langt de fleste bliver indenfor deres eget job

². Det understøtter i høj grad Gruskys antagelse om, at blah blah læs det han skriver om mikroklasser #todo).

Fra niveau 1 til niveau 2 er Moneca er i stand til at forhøje den interne mobilitet med 7 procentpoint, således at $\frac{3}{4}$ af jobmobiliteten foregår indenfor disse. Det er i god overensstemmelse med tidligere undersøgelser, der viser at hovedparten af skift i job sker indenfor relativt afgrænsede delmarkeder (Boje og Toft 1989, s. 124). Der er, så vidt jeg ved, ikke foretaget en jobmobilitetsundersøgelse på så detaljeret et Disco-niveau og med så omfangsrigt et datamateriale. Vi kan altså allerede konstatere, at det danske arbejdsmarked er voldsomt specialiseret, blot ved at se på de første to Moneca-niveauer.

Jeg vil nu gennemgå udviklingen i niveauerne visuelt, og beskrive udviklingen ud fra tabel 3.1.

Figurerne der afbilleder processen skal tolkes således: Forskellen på hvide og sorte noder i figurerne på side 16 til side 19 er simpel. Sorte node indikerer at denne node er blevet inkluderet i en ny klynge siden det foregående niveau. Hvid node indikerer at den *ikke* er blevet inkluderet i en ny klynge siden niveauet før. Jeg har ikke afbilledet niveau 1, da der ikke findes et foregående niveau at vise et skift *fra*.

²Eller ihvertfald et job, der ligger så tæt på, at man selv på mit usædvanligt lave 4-cifrede ISCO-niveau ikke kan skelne dem fra hinanden.

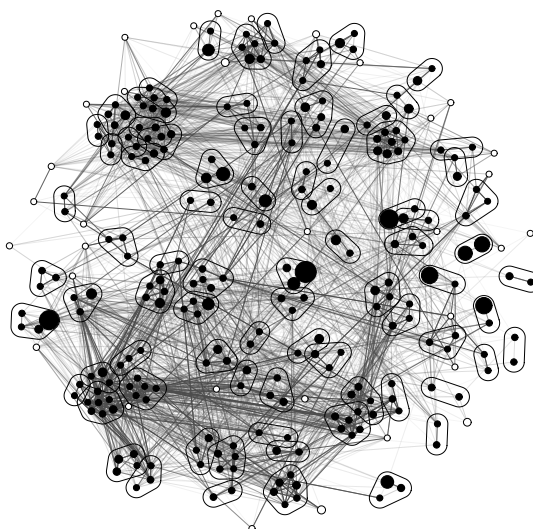
3.2.1 Niveau 2

Den mest markante aggregering i klynger og stigning i intern jobmobilitet fra finder sted i skiftet fra 1. niveau til 2. niveau, og herefter falder effekten af klyngedannelsen på den interne jobmobilitet støt. Af figur 3.2 fremgår aggrereringen.

På 2. niveau går vi fra de oprindelige 273 fritstående noder til 81 klynger, samt 33 endnu fritstående noder. Det vil sige 114 i alt³. Antallet af klynger er reduceret med 139 %, og den gennemsnitlige interne mobilitet i klyngerne er steget med 7 procentpoint.

Figur 3.2

2. niveau



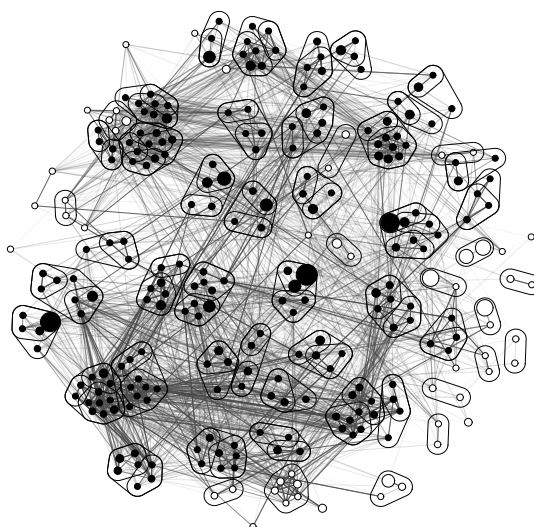
³ Der skelnes i tabel 3.1 ikke mellem enkelte noder og klynger, da klyngernes interne mobilitet jo netop skal erstatte nodernes, og en direkte sammenligning derfor er ønskelig.

3.2.2 Niveau 3

På 3. niveau, afbilledet i figur 3.3, inkluderes 20 enkeltstående noder i klynger, samt en række niveau 2 klynger aggrereres yderligere. Det reducerer antallet af grupper til 69. Stigningen i den gennemsnitlige interne mobilitet er på 3 procentpoint, hvilket er beskedent sammenlignet med stigningen fra niveau 1 til niveau 2. Tilgængæld er reduktionen i *antallet* af grupper på 65 %. Det vil sige at kompleksiteten i netværket reduceres, samtidig med at den interne mobilitet øges. Det er tendensen fra niveau 3 og frem: Der foretages stadig væsentlige sammenlægninger af klynger og noder til større klynger, men betydning af sammenlægningen på den gennemsnitlige interne mobilitet er lille. Det tolker jeg som udtryk for, at jo længere væk fra den sociale kontekst af den oprindelige erhvervsgruppe, vi kommer, desto mindre systematik er der i jobskiftet.

Figur 3.3

3. niveau



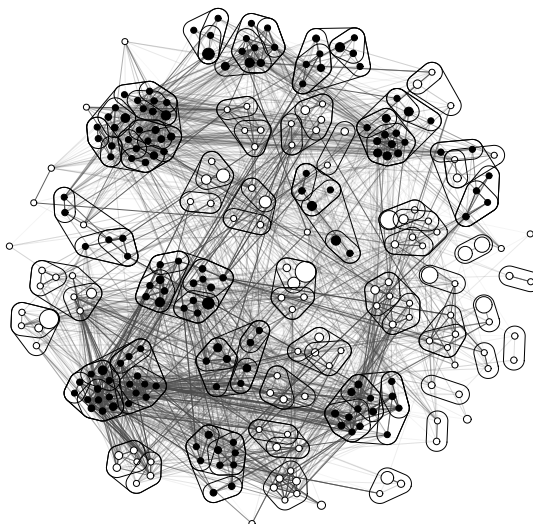
3.2.3 Niveau 4

På 4. niveau skabes 12 nye klynger. Kun en enkelt af disse sker i en sammenlægning med en enkeltstående node. Resten er tidligere klyngedannelser. Der er nu 56 grupper, hvoraf 14 er noder. Den interne mobilitet er steget med to procentpoint. ⁴Antallet af grupper er tilgængæld reduceret med 23 %, lidt under $\frac{1}{4}$. Det er på dette niveau, at sammenlægningerne binder allerede store klynger sammen, og vi får udvidede klynger med en betragtelig mængde forskellige erhvervsgrupper samlet under ét.

Disse klynger af erhvervsgrupper er ikke baseret på den *teknistiske vision*, som Grusky advarer imod, og heller ikke en a priori teoretisk opdeling, som hos Goldthorpe og Oesch. Inddelingen er istedet baseret på den sociale nærhed, der antages at findes mellem jobs, hvor mobilitet er "let og typisk", som Weber siger det i sin definition af hvad der udgør en social klasse. ⁵

Figur 3.4

4. niveau

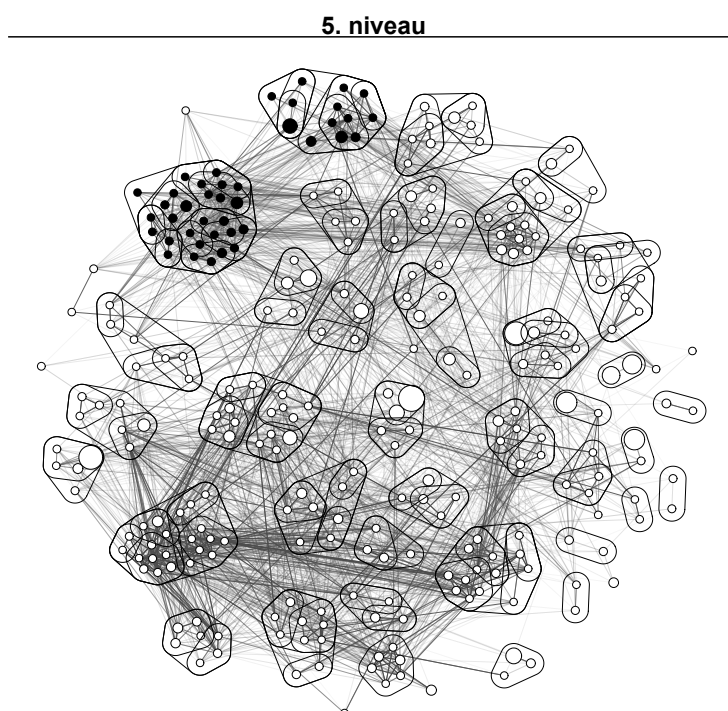


⁴Selvom to procentpoints stigning kan virke af lidt, viser effektforskninger, fx Andrade, at bare den 1 1/2 procent han kan finde i indkomstforskelle meget. De forklarer ofte kun 5-10 procent af den totale variation. skriv det her bedre #todo

⁵Der er naturligvis flere elementer end individuel jobmobilitet på spil. Det fulde citat lyder: "A »social class« makes up the totality of those class situations within which individual and generational mobility is easy and typical." (Weber 1978, s. 302).

3.2.4 Niveau 5

Det 5. niveau er det sidste niveau, da Moneca-algoritmen ikke kan aggregere klynger på et højere niveau end det. Vi mindes fra gennemgangen af Moneca-proceduren på side 7, at det betyder, at Moneca ikke kan finde nogle noder eller tidligere klynger, hvor der er intern mobilitet mellem *alle* klynger fra niveauet tidligere. I skiftet fra niveau 4 sker der tre aggregeringer: En inklusion af en enkelt node i en niveau 4- klynge, samt en større sammenlægning af en niveau 4- og en niveau 3-klynge. Det bringer os ned på 13 enkeltstående noder, samt 41 klynger.



Figur 3.5

Her er forbedringen i den interne mobilitet uanselig. Vi kommer ikke højere op end en gennemsnitlig intern mobilitet på 79 %. Dette forekommer mig rimeligt og acceptabelt, og en stigning på 11 procentpoint i den interne mobilitet fra niveau 1 til niveau 5 vurderer jeg som ganske tilfredsstillende, når det endelige resultat forklarer omtrent $\frac{4}{5}$ af al mobilitet på arbejdsmarkedet. Også selvom udgangspunktet var højt i forvejen, er det tilfredsstillende. Dette er i god overensstemmelse med tidligere undersøgelser, der viser at hovedparten af skift i job sker indenfor relativt afgrænsede klynger (Boje og Toft 1989, s. 124).

Den oprindelige 273 x 273 mobilitetstabel nu er reduceret til en mere overkommelig 54 x 54 mobilitetstabel. Ud af disse 54 kategorier er 13 enkelstående noter, der ikke er sammenlagt med andre noter, mens 42 er klynger. Tabel 3.2 viser yderligere centrale mål. Disse viser tydeligt at aggregeringen har skabt færre kategorier med højere intern mobilitet.

Tabel 3.2: Sammenligning af 1. niveau og 5. niveau

Niveau	grupper	gns. intern mobilitet	sd afvigelse (i procentpoint)	median	min	max
1. niveau	273	68%	12	68%	43%	97%
5. niveau	54	79%	10	80%	52%	96%

Medianen på det 1. niveau er på 68 %, mens det på 5. niveau er steget til 80 %. Standardafvigelsen for den interne mobilitet på det 1. niveau er på 12 %, mens den på det 5. niveau er på 10 %. At standardafvigelsen er faldet og medianen er steget, er vigtigt, fordi det fortæller os at gennemsnittet ikke er trukket op via enkelte klynger. Det er en generel forbedring af mobilitetsstrukturen. Det ses også ud fra minimumværdierne. Hvor de laveste værdier tidligere lå på lidt over 40 %, ligger de nu på lidt over 50 %. På de 3 centrale mål for kontinuerte variable (Find henvisning Malchow-Muller #todo) er der sket en forbedring. 5. niveau forklarer *mere* mobilitet, og kategoriernes interne mobilitet svinger *mindre* om gennemsnitsmobiliteten end i kategorierne på 1. niveau. Det er meget tilfredsstillende, da det indikerer at et højere aggregeringsniveau ud fra Monecas beslutningsprocedure forklarer *hvor* mobiliteten løber hen, når den bevæger sig udenfor de oprindelige kategorier.

Fra 3. niveau og frem er klyngedannelsens primære funktion at danne større klynger, det vil sige reducere kompleksiteten i jobstrukturen. Visse klynger indeholder langt flere beskæftigede end andre, og kan derfor anses som sammenlægninger, der er mere væsentlige end andre. Vi får dermed nogle klynger, der antalsmæssigt beskæftiger store dele af den danske befolkning. Dette vil jeg mene er et andet vigtigt parameter i klyngedannelsen, så længe at det ikke bryder med kravet om delmarkedsstatus, hvilket min gennemgang viser at det ikke gør: Når folk søger andet arbejde end indenfor deres erhvervsgruppe, vil langt de fleste finde job indenfor de klynger, som kortet viser.

I dette afsnit er aggregeringen fra niveau til niveau blevet beskrevet med henblik på at vurdere om vi kan tale om delmarkeder fremfor blot klynger. Det er klart, at der ikke er tale om dual eller tredeling af arbejdsmarkedet, men en langt mere fraktioneret struktur. Det skulle gerne stå klart, at vi kan tale om delmarkeder på det danske arbejdsmarked i langt de fleste tilfælde,

omend graden af intern mobilitet naturligvis varierer. Det følgende afsnit vil præsentere

det første "rigtige" netværksskort, hvor jeg vil undersøge nogle af de klynger, hvor den interne mobilitet er lav, og derfor næppe fortjener status af delmarkeder, samt nogle af dem, hvor den er meget høj, og vi derfor har et skoleeksempel på et delmarked. Dette afsnit har desuden tjent som en vejledning i at forstå hvordan netværksskortet er lavet, så de nemmere kan tolkes i resten af analysen.

3.3 Den interne mobilitet i klyngerne

Den interne mobilitet for hvert klynge er præsenteret i netværksskortet i figur 3.6 på side side 22. I analysen af spørgsmålet om delmarkeder, er logikken i kortets struktur blevet præsenteret. For at kunne forstå det endelige netværksskort, kræves der imidlertid yderligere forklaring, da både *størrelsen på noderne*, *nodernes størrelse* samt *forbindelsernes farve* indeholder information om strukturen på det danske arbejdsmarked.

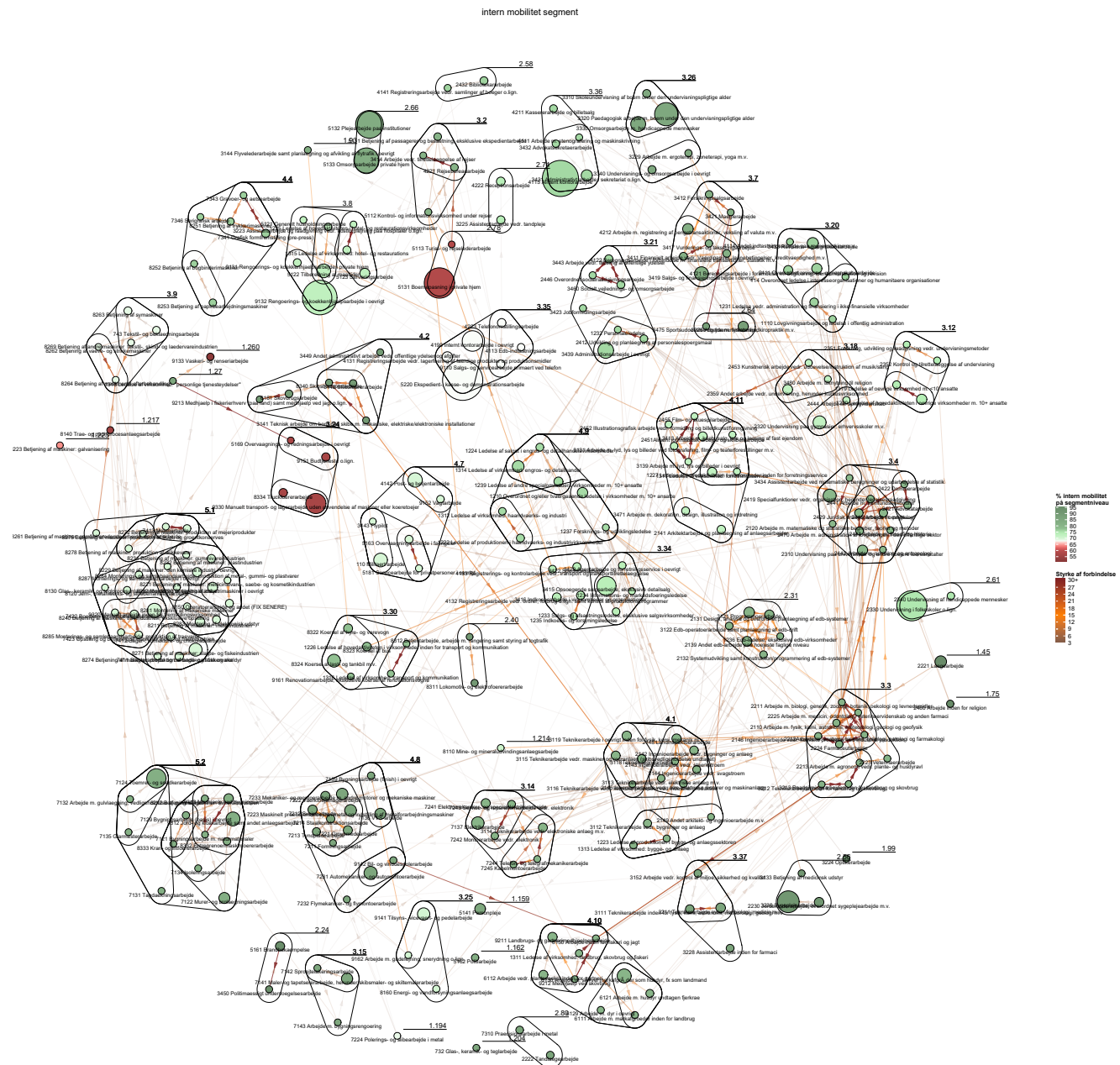
Størrelsen på en node repræsenterer ganske simpelt hvor mange personer, der i gennemsnit er i beskæftigelse indenfor erhvervsgruppen i årrækken 1996-2009⁶ Den største erhvervsgruppe, 5220: Ekspedient-, kasse- og demonstrationsarbejde indeholder 114.869 personer i gennemsnit, svarende til 4,9 % af det totale antal beskæftigede. Mens den mindste, 7346: Serigrafisk arbejde indeholder 504 personer, svarende til 0,02 % af det totale antal beskæftigede ⁷

⁶Når jeg fra nu af taler om hvor mange der er beskæftigede, er det underforstået at der er tale om det gennemsnitlige antal *indenfor årrækken 1996-2009*, med mindre det eksplicit fremhæves at der er tale om noget andet.

⁷Der er ikke tale om et 1:1 størrelsesforhold mellem antallet af personer i de forskellige erhvervsgrupper. Det er for komplekst til visuel afkodning i et allerede informationstungt netværksskort, da de mindste noder ville blive meget små og de store noder meget store. Derfor har jeg valgt et størrelsesforhold mellem noderne, der ikke er 1:1 med deres reelle forskelle i størrelse, men som alligevel giver en fornuftig fornemmelse for disse størrelsesforhold. Man skal dermed se både farve af forbindelserne og nodernes størrelse som en grov målestok af den variabel, den repræsenterer. Med vægten lagt på nem visuel afkodning, fremfor korrekt gengivelse af datakompleksiteten.

3.3. Den interne mobilitet i klyngerne

Figur 3.6: Intern mobilitet for klyngerne.



graden af den interne mobilitet ses ved nodernes farve. Farven går fra mørkerød, til hvid, til mørkegrøn, bestemt af graden af intern mobilitet i delmarkedet. Røde markerer en intern mobilitet på mellem 50 og 67 %, mens hvid er en intern mobilitet på 68 %, det vil sige: den gennemsnitlige interne mobilitet for noderne uden moneca aggregering. Det forekom mig som et passende sted at markedere en form for "neutralt" niveau af intern mobilitet. Grøn markerer en intern mobilitet på mellem 69 % og 96 %. Da der er tale om en gradient, er skiftet i farve en kontinuer overgang, hvor intensiteten af henholdsvis rød og grøn viser hvor langt ude på skalaen vi er, i den ene eller anden retning (husk at vær sikker på at det du skriver her er korrekt for skalaen i sidste instans #todo)⁸forbindelsens farve markerer som beskrevet i afsnit 1.0.2 forbindelsernes styrke, målt ved den relative risiko. Et forholdsmål, læseren kan genkalde sig udtrykker mobiliteten fra en erhvervsgruppe til en anden, givet den relative størrelse af antallet af beskæftigede i erhvervsgruppen. Det er styrken af denne relative risiko (RR), der afgør farven på forbindelserne, der kan antage tre nuancer: Helt lysegrå angiver en RR på 3⁹. Helt orange angiver en RR på 15. Helt mørkerød angiver en RR på 30 *eller derover*¹⁰

Forbindelserne er desuden retningsbestemte, hvor pilene på kortet angiver retningen af mobiliteten mellem noderne.

Ovenstående er de generelle retningslinjer for netværkshortet. Bortset fra nodernes farve, der vil variere, alt efter hvilken variabel, der farvelægges efter.

Så vidt beskrivelsen af hvad kortet viser. Fremfor at kommentere kortets indhold overordnet, vil jeg for overblikkets skyld zoome ind på nogle bestemte klynger eller delmarkeder, forklare hvad vi ser, og først derefter kommentere kortet i sin helhed.

3.3.1 Klynge 2.78: Børnepasning i private hjem samt turist- og rejseledere

Hvis vi kigger på klynge 2.78, kan vi se at den består af børnepasning i private hjem, samt turist- og rejseledere. Erhvervsgrupperne har en intern mobilitet på henholdsvis 59 % og 47 %, altså ganske lavt sammenlignet med den gennemsnitlige interne mobilitet på 68 %. Beg-

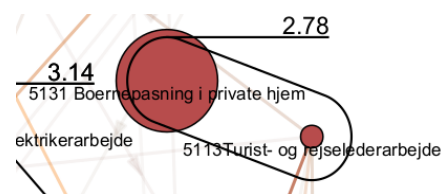
⁸ Skiftet i farve er, som nævnt, min egen vurdering, baseret på de deskriptive værdier for den interne mobilitet gennemgået i det foregående afsnit. Der findes ingen tidligere brug af Moneca algoritmen, hvori dette bliver analyseret.

⁹Husk på den præcise definition af relativ risiko: det vil sige, at der ved en lysegrå forbindelse er 3 gange så stor sandsynlighed for mobilitet fra den forladte beskæftigelse til den nye beskæftigelse, i forhold til hvis der var tale om et arbejdsmarked med helt fri bevægelse.

¹⁰ I bestemmelsen af klyngerne, er alle forbindelser med en RR på over 1 medtaget som en forbindelse. På det visuelle kort er den informationsmængde upraktisk. Det bliver umuligt at aflæse den underliggende systematik i hvilke forbindelser, der værd at lægge mærke til. Enkelte forbindelser har RR-værdier på flere tusinde, men de fleste langt lavere: Medianen i er på 2,1. En relativ risiko på 3 svarer cirka til den 65. percentil, og 15 er cirka den 90 percentil. 30 svarer til cirka den 92,5te percentil, og er ikke valgt ud fra denne, men bestemt ud fra det punkt, hvor der sker en stigning i styrken af forbindelserne.

ge erhvervsgrupper er indenfor hovedgruppe 5 i Disco-nomenklaturet: *salgs-, service- og omsorgsarbejde*. Det er arbejde, der ifølge Danmarks Statistik er klassificeret som ISCEDs færdighedsniveau 2, hvilket betyder at det kræver uddannelse “på grundniveau” (DST-DISCO88 2016, tabel 1). Det er en klar indikator på at de formelle sociale lukningsmekanismer i disse to erhvervsgrupper er stærkt begrænsede.

En oplagt årsag er at beskæftigelserne her simpelthen er jobs, der af den ene eller anden årsag har en karakter, der gør at folk sjældent opholder sig længe i dem. Det kunne man kalde en socialt betinget lav intern mobilitet. Hvis arbejdet har denne karakter, må man forvente, at de sociale lukningsmekanismer næppe er særlig højtudviklede. Den mest



Figur 3.7

legitime af disse, i moderne samfund, må være adgangsgivende uddannelse. Vi vil derfor forvente at finde enten ufaglært fag, eller fag med lave uddannelsesmæssige krav. Det kan altså have en social årsag at den interne mobilitet er lav. Det er vigtigt at få afklaret, da disse klynger kan risikere ikke at kunne betegnes som delmarkeder i teoretisk forstand, hvis mobiliteten mellem erhvervsgrupperne ikke er hyppig. Det er ikke tegn på et metodologisk problem, men snarere et aspekt ved arbejdets karakter, og den sociale position, jobbet har på arbejdsmarkedet. I denne lille klynge er den gennemsnitlige alder 37 år¹¹, hvilket er ca. 5 år under populationsgennemsnittet på 42,3 år. Vi har derfor at gøre med to erhvervsgrupper, der tiltrækker en stor andel af unge mennesker, og for hvem dette arbejde ikke nødvendigvis siger noget om hvilket delmarked, de vil indtage i deres faste klasseposition (indsæt Goldthorpe bemærkning om hvornår folk har en fast klasseposition her og find citatet #todo). Der er naturligvis også dagplejemødre indenfor privat børnepasning, det vil sige voksne kvinder med meget tid i hjemme. Men det er rimeligt at antage at de også hopper ind og ud af andre positioner på arbejdsmarkedet.

Hvad angår forbindelsen til turist- og rejseleder, kunne man lidt sensationsvinklet sige, at der muligvis er tale om unge, fortrinsvis piger, der passer børn som teenagere, og derefter tager til Ibiza, sydamerika eller andre steder og oplever verden. For derefter at tage sig en uddannelse.

Der er andre klynger på kortet, hvor uddannelseskravene er lave, og som befinder sig på samme færdighedsniveau indenfor ISCED-klassifikationen, men hvor den interne mobilitet alligevel er høj. Det er et sådant delmarked, vi skal se på nu.

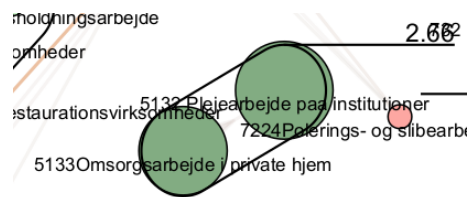
¹¹ Det også gælder de to jobs i klyngen hver især

3.3.2 Delmarked 2.66: Omsorgsarbejde i private hjem og plejearbejde på institutioner

beskæftigelsesgrupper i delmarked 2.66 befinder sig ligeledes i Discos hovedgruppe 5. Børnepasning i private hjem ligger ganske tæt på arbejdsfunktionerne i klynge 2.66. Så tæt at de på et 3-cifret Disco niveau er kategoriseret ens, det er kun på det 4. ciffer at de adskiller sig fra hinanden. delmarked 2.66 har en intern mobilitet på 83 %, væsentligt højere end klynge 2.78. Det kan ikke forklares ud fra formelle inddelinger såsom Discos hovedgrupper eller ISCEDs færdighedsniveauer.

Børnepasning i private hjem har tilsyneladende en noget anden social profil. Den interne mobilitet er meget lavere, og den ligger sammen med turist- og rejseleder, der også befinder sig i hovedgruppe 5, men ikke beskæftiger sig med omsorgsarbejde for folk, der har svært ved at klare sig selv. Forskellen i social sammensætning kan aflæses i den demokratiske sammensætning. Delmarked 2.66 har en gennemsnitsalder på 42 $\frac{3}{4}$ år, hvilket er er nærmest identisk med populationsgennemsnittet på 42,3 år.

Der er flere klynger i kortet, der viser lignende forskelle, trods samme formelle niveau i adgangskrav. Formelle færdighedsniveauer og uddannelseskrav er derfor ikke udslagsgivende for, hvorvidt der er tale om et klyng eller ej.

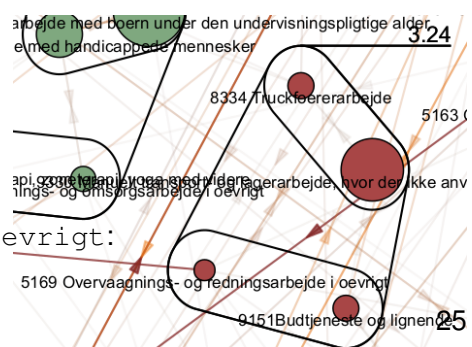


Figur 3.8

3.3.3 Klynge 3.24: Transport af varer, ambulanceførere og vagtarbjede

EN anden klynge 3.24 med lav intern mobilitet er en niveau 3 klynge, der indeholder to niveau 2 klynger. Den har en intern jobmobilitet på 57 %. Der er tale om tre typer beskæftigelse, der omhandler transport af varer. Den sidste, 5169: Overvaagnings- og redningsarbejde i øvrigt, indeholder beskæftigelser som ambulanceførere, livvagt og sikkerhedsvagt. Selvom det ikke direkte har samme omdrejningspunkt som de tre andre, fornemmer man alligevel en vis social mening i en jobcirkulation mellem de tre førnævnte, der også inkluderer denne sidste type job.

Set ud fra Goldthorpes syn på arbejdskontrakter, er det jobs, hvor specificiteten af kompetencer er lav, og muligheden for overvågning af arbejdet må antages at være høj. Det er desuden hårdt fysisk arbejde, eller, i tilfældet af 5169: Overvågnings- og redningsarbejde i øvrigt:



Figur 3.9

Arbejde, der kræver god fysisk form. Det er ikke overraskende, at denne type arbejde har en høj frafaldsprocent, da nedslidning og krav til fysik er høje.

3.4 Delkonklusion

For de fleste klynger er den interne mobilitet højere end 70 %, det vil sige, kan karakteriseres som delmarkeder. Få klynger som dem der bl.a. er blevet beskrevet ovenfor har en intern mobilitet på helt ned til 50 %.

Fælles for klyngerne er, at de vurderes til at være afgrænset i sådan et omfang, at der er væsentlige barrierer mellem klynget og resten af arbejdsmarkedet, der gør, at der inden for klynget er en nærhed, hvor mobilitet er "let og typisk".

Formålet med dette kapitel har været at besvare forskningsspørgsmålet: Er der en opdeling af arbejdsmarkedet for arbejdstagere i delmarkeder, hvor mobilitet indenfor delmarkederne er hyppig, og mellem delmarkederne sjælden?

Moneca-algoritmen har aggreret 273 jobfunktioner til 51 klynger. Denne opdeling afspejler Boje (1986) og Toubøl (2013). Klyngedannelsen er skabt via Moneca-algoritmen og kvalitetsvurderet fra niveau til niveau. Der er i alt blevet splittet fire klynger op, som ud fra mobilitet, densitet og antal stier ikke har kunne vurderes som meningsfulde delmarkeder. 79 % af al mobilitet foregår inden for delmarkeder, det vil sige, at mobilitet indenfor delmarkederne er hyppig, og mellem delmarkederne sjælden. Klyngerne skabt i aggreringen kan derfor betragtes som delmarkeder.

4 Delanalyse 2: Forskellen på et delmarked og et segment er påvisningen af særegne (specifikke? partikulære) sociale processer indenfor delmarkedet

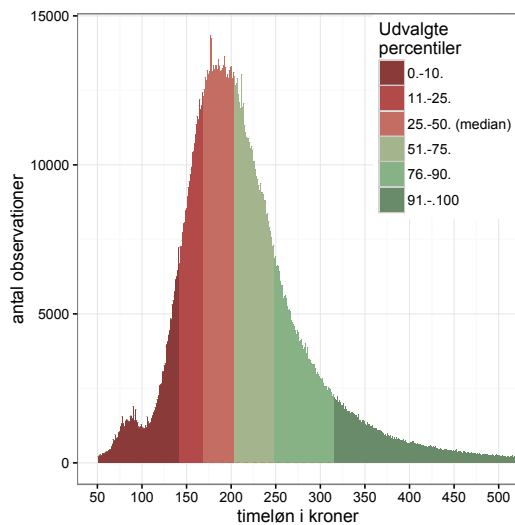
Efter at have konstateret at der er en opdeling af arbejdsmarkedet for arbejdstagere i delmarkeder, hvor mobilitet indenfor delmarkederne er hyppig, og mellem delmarkederne sjælden, har dette kapitel til formål at svare på det andet forskningsspørgsmål:

Forskningsspørgsmål
2.
Kan forskelle i de sociale processer vise, at der er tale om segmenter, og ikke blot delmarkeder?

Dette forskningsspørgsmål vil jeg besvare ved at kigge på de sociale processer i delmarkederne.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Figur 4.1: timeløn farvelagt efter udvalgte percentiler



Tabel 4.1: Udvalgte mål for fordelingen af timelønninger

Gennemsnit	227 kr.
Sd-afvigelse	82 kr.
Mindste værdi	25 kr.
Højeste værdi	4.492 kr.
1. percentil	130 kr.
10. percentil	157 kr.
25. percentil	178 kr.
Median	208 kr.
75. percentil	250 kr.
90. percentil	314 kr.
99. percentil	525 kr.
<i>n</i>	382.131

4.1 Indkomstfordeling for alle

Indkomstfordelingen i den arbejdende del af den danske befolkning kan ses i figur 4.1 og tabel 4.1.

Vi kan se at fordelingen er nogenlunde centreret omkring gennemsnittet og medianen, hvilket også kommer til udtryk i standardafvigelsen på 76 kr/t. stigningen i timeløn fra 100 kr/t til 175 kr/t er drastisk, mens faldet fra omtrent 200 kr/t og fremefter er noget mere støt i hældningen. Det kan man tolke sådan, at arbejdsmarkedsinstitutioner og overenskomster er ganske effektive til at regulere løndannelsen, således at lave timelønninger alligevel hindres i at ramme de niveauer, som fri prissætningen på arbejdskraften ville medføre. Det ses også, af den lange hale i timelønfordelingens øvre percentiler, at en lignende mekanisme ikke findes for de højere indkomster. Det danske skattesystems progressive beskatning af indkomst har selvfølgelig en hvis betydning, men det har ikke på samme måde en tydelig opdæmmende effekt, som vi ser, at der må findes for de lavere timelønninger. det er tydeligt, at timelønninger over den 90. percentil, tjener eksponentielt mere og mere frem mod den højeste timeløn på 4.492 kr.

Derudover findes der en lille pukkel af lønninger på mellem 25 og 75 kr¹. Eftersom vores population er alle i arbejde melle 16 og 70 år, er der i de fleste tilfælde sandsynligvis tale om

¹ Grundet en teknisk fejl, der ikke blev rettet inden min forbindelse til registerdata udløb, er lønninger mellem 25 og 50 kr/t ikke vist i grafen. Det betyder *ikke* at de ikke er tilstede i datamaterialet, blot at grafen ikke viser dem.

ungdoms- og studiearbejde.

4.2 indkomstfordelinger på delmarkederne

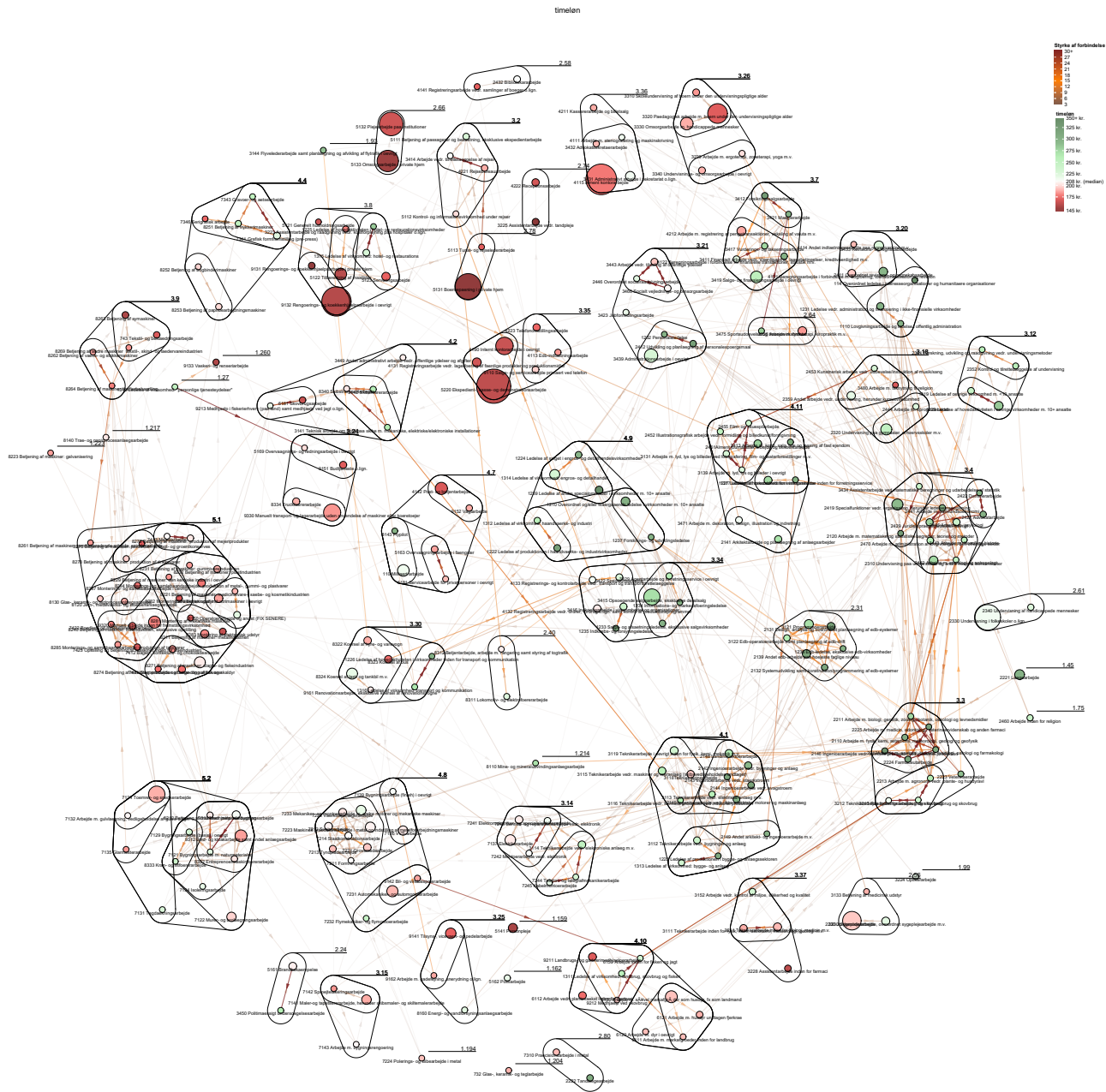
Efter at have set på fordelingen af timelønninger generelt, skal vi nu se hvordan det er differentieret ud fra arbejdsmarkedets delmarkeder og klynger.

I figur 4.2 er netværksskortet farvelagt således, at hvid markerer medianen på 208 kr/t. De røde farver er timelønninger under 208 kr/t. Jo mørkere rød, desto lavere timeløn. De grønne farver er over 208 kr/t. Jo mørkere grøn, desto højere timeløn². Det ses at den grønne farves graduering er strukket længere end rød. Det er ikke overraskende, da vi netop ud fra figur 4.1 kan se, at timelønninger strækker sig lagnt højere op over medianen end under den, bl.a. grundet det danske arbejdsmarkedets indretning, som argumenteret for i beskrivelsen af figur 4.1.

² Delmarkeder med en gennemsnitlig timeløn på 350 kr/t er sat til 350, da skalen ville blive spredt for langt ud for at imødegå disse outliers, hvilket ville gøre det svært at se nuancerne i forskellen mellem de timelønninger, som de fleste erhvervsgrupper trods alt har.

4.2. indkomstfordelinger på delmarkederne

Figur 4.2: Intern mobilitet for klyngerne.



Det overordnede netværkskort præsenteret her, giver os allerede en ide om en vis systematik i lønforholdene i forskellige placeringer på arbejdsmarkedet. Det er velkendt at kortets delmarkeder er lavet ud fra mobilitet mellem erhvervsgrupperne. Selve placeringen af delmarkederne på kortet er imidlertid også sket ud hvilken mobilitet der findes *mellem* klyngerne. Derfor er det interessant, at vi ser en rødlig halvmåne forme sig i venstre side af kortet, mens en - lidt mindre - grøn formation findes i højre side. Ved at se på forbindelserne kan vi straks konstatere hvorfor, at kortet ser ud som det gør: Mellem de grønne, velbetalte delmarkeder ser vi meget kraftige forbindelser mellem delmarkederne.

De 18 delmarkeder, der her er tale om, befinder sig stort set alle i Disco-hovedgruppe 1, 2 eller 3. Hvilket vil sige erhverv med høje (boglige) uddannelseskrav, og/eller meget ledelsesansvar.

Der kan være to årsager til hvad vi ser på kortet: 1) Enten er udvekslingen af arbejdskraft på tværs af delmarkeder langt mere hyppig, eller 2) udvekslingen er mere specifik: Forbindelserne på kortet viser de kraftigste forbindelser, ikke alle forbindelser, og derfor vil mobilitet, der er mere jævnt fordelt, forblive usynlig på (denne udgave af) kortet. Ved at se på den interne mobilitet for sig for henholdsvis de 18 delmarkeder i den grønne formation for sig, og den røde halvmånes 36 delmarkeder for sig er det muligt at svare på. Den grønne formation har en gennemsnitlig intern mobilitet på 80 %, og den røde halvmånes er på 77 %³. Der er altså ikke tale om en væsentlig forskel i intern mobilitet mellem de to formationer. De kraftigere forbindelser i den grønne formation skyldes derfor, at den mobilitet, der findes, foregår langt mere struktureret langs bestemte linjer på tværs af delmarkederne. Den røde halvmånes mobilitet er "smurt tyndere ud", og derfor ses den ikke på kortet⁴. Sikft mellem den grønne formation af delmarkeder, med lange boglige uddannelse og/eller meget ledelsesansvar i højere grad sker via bestemte kanaler end tilfældet er i de resterende delmarkeder. Det er interessant, da det der netop karakteriserer den grønne formation, er at deres lønniveauer befinder sig i den bedst betalte halvdel af befolkningen.

Hvis man skal komme med en tentativ analyse af hvad denne forskel i mobilitetsstruktur kan betyde for den politiske orientering, erindrer læseren muligvis Claus Hjort Frederikssens udmelding som beskæftigelsesminister i 00'erne (find præcis årstal senere #todo), hvor han kom i vælten ved at udtale, at hvis man med en lang videregående uddannelse ikke kunne finde job, skulle man finde arbejde i netto eller blive postbud i stedet. Det udløste et ramskrig. Blandt dem med lang videregående uddannelse. Mens det for ikke-boglige klasser sandsynligvis var en ganske tilfredsstillende udmelding, og det er nok ikke usandsynligt,

³ Det samme gælder for medianen og standardafvigelsen: Den røde halvmåne har en noget højere standardafvigelse og lavere median, men kun et par procentpoints forskel.

⁴ I appendiks B, figur B.3 findes et kort der viser alle forbindelserne, hvor dette kan ses.

at udtalelser som denne, der gjorde Venstre til et bredt funderet parti i 00'ernes Danmark. Mit kort giver en troværdig forklaring på hvorfor, Da det ses, at netop erhvervsgrupper med lange videregående uddannelser, er vant til at søge beskæftigelse indenfor andre delmarkeder i langt mere strukturerede baner end resten af delmarkederne. Det er denne form for kulturkrig, sociologen Thomas Frank beskriver i sin bourdieu-inspirerede klasseanalyse i bogen "What's the Matter with Kansas? - how the conservatives won the heart of America" fra 2007. i Norge udkom en lignende bog i 2007 med en identisk Bourdieau-analyse, skrevet af journalisten Magnus Marsdahl. Analysen er kort sagt, at venstrefløjspartierne er blevet til de boglige klassers parti. Det lukrerer de borgerlige politikere på ved at markerer sig med forslag og udtalelser, der er på kant med de "dannede" sociale gruppers verdensbillede, hvilket de naturligvis reagerer på med deres lærde sprog. Disse udmeldinger viser venstrefløjens traditionelle arbejderklasse, at den verden, hvori de venstreorienterede værdier står stærkt, ikke er deres verden.

Det er selvfølgelig et noget stort brød at slå op, og jeg påstår ikke at den her fundne struktur kan bære en sådan bevisbyrde alene. Jeg vælger alligevel at præsentere den her, da jeg dog mener, at det netop er sådanne oplevelser af arbejdet, der giver et blik på verden, der kan formes til forskellige politiske perspektiver. Det er min overbevisning, at politiske værdiorienteringer ikke har samme kobling til arbejdslivet som tidligere, hvilket er et velfunderet empirisk standpunkt indenfor sociologisk og politologisk forskning (henvisning til Scott #todo). Det betyder imidlertid ikke, at politisk værdiorientering er afkoblet fra arbejdslivet, men at en række komplekse ændringer af parti, klasse og forbrugskulturen gør disse sammenhænge anderledes end tidligere. Som valget af Trump i november 2016 viste, er klasse stadig en relevant faktor, hvis man forstår hvad det betyder i dag (link til artikel fra Trumps datahold og så noget mere substantielt der viser at analytikerne tog fejl). Den her fremførte tolkning er et eksempel på hvorledes en oplevelse af arbejdslivet kan understøtte moderne sociologisk funderede analyse om politisk værdidannelse ud fra moderne klasseteori, såsom Thomas Franks og Magnus Marsdahls analyser.

4.2.1 lala

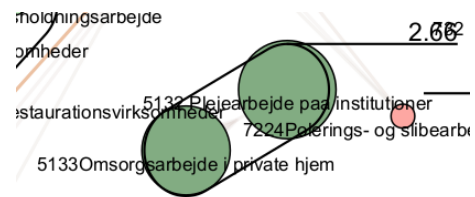
Hvilke delmarkeder har en høj timeløn og hvilke delmarkeder har en lav timeløn. For at illustrerer dette vil jeg zoomer ind på to delmarkeder. Det ene delmarked jeg vil fremhæve er delmarked 5.1, som består af jobfunktioner relateret til primært faglært maskin- og fabriksarbejde herunder produktioner af mad, træ og metal. Det andet delmarked jeg vil fremhæve er delmarked , som består af jobfunktioner relateret til akademisk arbejder herunder samfundsvidenskabeligt og humanistisk arbejde.

4.2. indkomstfordelinger på delmarkederne

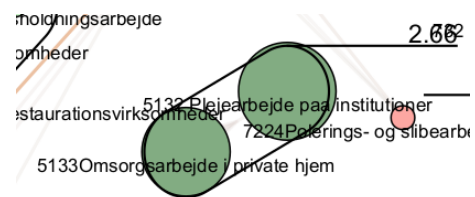
De personer som arbejder og skifte job inden for delmarked arbejder som advokater, dommere, sociologer, antropologer, statistikere, undervisere på universiteter, historikere, filosoffer, psykologer. De får en løn som i gennemsnit spænder fra XXX kr/time for de laveste (hvem er det) til XXX kr/time (hver er det). Udover en høj timeløn har de det til fælles at det typisk kræver en lang videregående uddannelse for at besidde disse job. Nogle jobs såsom advokat- og dommerarbejde har kræver udover en lang videregående uddannelse og en særlig efteruddannelse.

De personer som arbejder og skifte job inden for delmarked 5.1 arbejder som slagtere, bage-re, mejerister, arbejde med betjening af maskiner, monteringsarbejde, samlebåndsarbejde og manuelt arbejde. De får en løn som i gennemsnit spænder fra XXX kr/time for de laveste (hvem er det) til XXX kr/time (hver er det) med en enkelt afstikker på XXX kr/time. Udover en lavere timeløn har de det til fælles at det typisk kræver en erhvervsuddannelse eller et kortere kursus at besidde disse job.

Den første umiddelbare umiddelbare forklaring på lønsforskellen mellem de to segmenter er uddannelse. For at arbejde inden for de to delmarkeder kræver det en uddannelse. Forskellen er så længden på uddannelsen. Det er anerkendt, at der er en sammenhæng mellem uddannelse er løn. Human kapital er en økonomisk teori udviklet af Gary Becker, der beskriver, at arbejdstagere kan tage en uddannelse som en investering mod forventning om økonomisk compensation ved at vedkommendes produktivitet øges (fx Cahub og Zylberberg 2004, s. 69). Da de akademiske uddannelser tager længere tid end erhvervsuddannelser og kortere kursus, giver det også en højere økonomisk compensation.



Figur 4.3



Figur 4.4

4.3 Kønsforskelle blandt delmarkederne

A variabelbeskrivelser: løn

Jeg har valgt at beskrive indkomstbeskrivelse i realindkomst, da denne gør det muligt at lave sammenligninger i indkomst over tid. Alternativet er en nominel indkomstbeskrivelse, altså blot kroner og ører, som det ser ud på et givent tidspunkt (DST 2009, s. 6). En beskrivelse af realindkomsten tager udgangspunkt i købekraften, som den reelt ser ud, og korrigerer således for inflation mv. Tallene korrigeres ud fra forbrugerprisindekset (DST-Forbrugerprisindeks 2016).

Monecakortet er, som beskrevet i ???, et tværsnit lavet på baggrund af 1996 til 2009, altså 14 år. Et lignende tværsnit af indkomst virker derfor mest fornuftigt, da det på samme måde som selve netværkshortet viser strukturen, som den ser ud gennem alle de 14 år. Et alternativ ville være at vise lønninger for et enkelt år, eksempelvis det seneste, 2009. Det mener jeg ville være misvisende, da det ikke tager højde for de strukturelle udsving i indkomsten, der uværgeligt forekommer over en periode på 14 år. Derfor forekommer et gennemsnit af indkomsten over de 14 år som den fornuftigste måde at beskrive indkomst på, indenfor denne afhandlings rammer. Metoden til det er som nævnt at korrigere hvert af de tidligere år, således at alle indkomster i årene 1996-2008 bliver opdateres til 2009-priser, hvorefter gennemsnittet tages.

Danmarks Statistik har en række indkomstvariable til rådighed. Disse variable registrerer indkomst ud fra en række forskellige parametre. Selv tilsyneladende ens variable har alligevel diskrete, omend ofte betydningsfulde, forskelle i opgørelsesmetoden. Dette er ikke uskyldige forskelle, og i sammenligning af variablenes centralmål er det tydeligt, at disse forskellige opgørelsesmetoder har stor betydning for variablenes fordelinger.

Eksempelvis registrerer `DISPO_NY` den disponible indkomst efter skat og renter, mens `perindkialt` registrerer personindkomst i alt, hvilket defineres som summen af erhvervsindkomst, overførselsindkomst, formueindkomster samt anden ikke-klassificerbar indkomst, altså hvad der forekommer som rub og stub. Disse to variable beskriver et individs samlede

indkomst, men på forskellige måder, hvoraf den mest centrale og iøjnefaldende er, at den sidstnævnte måler før skat, og førstnævnte er efter skat. Derudover registreres den samlede indkomst ud fra forskellige kilder, som ikke altid er lige lette at gennemskue betydningen af. Dette ses også i deres fordelinger, samt i mængden af fejlværdier såsom negative indkomstværdier m.m.

Andre variable registrerer løn direkte, såsom `loen_mv`, der inkluderer hele den skattepligtige lønindkomst, inklusiv frynsegoder, skattefri løn, jubilæums- og fratrædelsesgodtgørelser samt værdi af aktieoptioner. Det gælder også `jobloen`, hvis opgørelsesmetode i beskrivelsen hos DST er mere centreret på ansættelses i registreringsmåneden november, omend også inkluderer andre ansættelser i ikke-novemberansættelser, som de kalder det.

Det har vist sig, at stort set alle indkomstvariablene, som jeg har til rådighed i registerdataet, ligger en del lavere end de tal, DST selv benytter til at beskrive løn indenfor forskellige DISCOkategorier i deres rapport *Indkomst* fra 2009 DST (2009).

Den eneste variabel, hvor det ikke er tilfældet, `TIMELON`, der beskriver timelønnen. Det er den variabel, der, udover sin kvalitet i form af let fortolkbarhed, kommer tættest DSTs egne tal i rapporten *Befolkningens løn 2012* fra 2013. Denne variabel vil derfor benyttes til at beskrive lønindkomst i denne afhandling, og selvom nedenstående vil inddrage de andre indkomstvariable som reference, er denne variabel, dens validitet og central omdrejningspunktet for dette appendiks.

A.1 Timeløn

variablen `TIMELON` har til formål at beskrive timelønnen, også for danskere, hvori lønninger normalt ikke opgøres i timeløn. "Variablen angiver den gennemsnitlige timeløn i novemberansættelsen for hoved- og bibeskæftigede lønmodtagere." (DST-TIMELON n.d.). Lønnen, der indgår i den gennemsnitlige timeløn, er det beløb, der årligt indberettes af arbejdsgiveren på oplysningssedlen til Skat. Dvs. at der indgår løn, feriegodtgørelse, løn under sygdom mv. Bidrag til pensionsordninger medregnes ikke.

Den centrale vanskelighed ved at beskrive timeløn, er, iflg. DST, at bestemme antallet af arbejdstimer for den enkelte, således at timelønsskønnet er pålideligt. Til at bestemme kvaliteten af timelønsskønnet i `TIMELON`, findes variablen `TLONKVAL`, der går fra 0 % til 100 %, hvor 0 % er maksimal sikkerhed af skønnet, og 100 % er maksimal usikkerhed. I denne vurdering indgår en række elementer til bestemmelsen af antal arbejdstimer, trukket af arbejdsgiverens oplysningssedler, ATP-bidrag mm. DST skriver, at værdier fra 0 til 50 % er af brugbar kvalitet, mens >50 % er af tvivlsom kvalitet. DST skriver om kvaliteten, at "*Set for alle ansættelsesforhold under ét er der en rimelig overensstemmelse mellem*

den beregnede timeløn i IDA og Danmarks Statistiks lønstatistik. For ansatte på fuld tid er der for 44 pct. af lønmodtagerne en forskel på 5 kr. eller mindre (Det økonomiske Råds sekretariat 2003). Der kendes ingen kvalitetsundersøgelser på et mere detaljeret niveau.” (DST-TIMELON n.d.). Det centrale her er, at der ingen kvalitetsundersøgelser findes på et mere detaljeret niveau.

Der er derfor to problemer på spil i ovenstående: *Det første* er, at TIMELON beregnes for novemberansættelsen et givent år, hvorimod vores DISCO-variabel benytter den længste ansættelse indenfor et givent år. Der er, for mange mennesker, med ganske stor sandsynlighed fin overenstemmelse mellem længste ansættelse, og novemberansættelsen. Men vi har ingen mulighed for at tjekke det. *Det andet problem* er det, DST skriver i ovenstående paragraf: Der findes ingen kvalitetsundersøgelse af TIMELON på et mere detaljeret niveau. Et sådant detaljeret niveau kunne eksempelvis være inddelingen af befolkningen i 150 arbejdskategorier. Som jeg gør det, i dette speciale.

For at komme det første problem til livs, vælger jeg derfor at fjerne alle observationer med en TLONKVAL på >50 %. Det sker naturligvis ud fra DSTs anbefaling om at disse er “af tvivlsom kvalitet” generelt, men især fordi, at dette mål for tvivlsom kvalitet er baseret på deltidsansættelse. Det er mit skøn, at ved at fjerne personer, der arbejder under halv tid på et år, fjerner jeg mange af de fejl, der måtte forekomme, hvis en person, der 3 måneder af året arbejder som taxachauffør, men tilfældigvis i november har en ansættelse som jord- og betonarbejder. Det ville betyde, at jeg i min benyttelse af TIMELON ville få registreret lønnen som taxachauffør som en jord-og betonarbejder lønning. Ved at fjerne dem, hvor deres TIMELON ansættelse er under halvdelen af året, er faren for fejlkategoriserede lønninger reduceret væsentligt, da jeg nu kun har lønninger med for folk, der arbejder over halvdelen af året. Chancen for at denne halvdel af året inkluderer november, er derfor over halvdelen. Dette er naturligvis ikke en skudsikker gardering overhovedet, men er en nem måde at fjerne, så vidt jeg kan vurdere, den værste kilde til fejl i min lønbeskrivelse.

min bedste mulighed for at vurdere omfanget af disse problemer for validiteten af TIMELON, er i sidste ende at undersøge dens overenstemmelse med virkelighedens lønninger. Den overenstemmelse har jeg mulighed for at vurdere via følgende 3 metoder:

1. DSTs timelønsbeskrivelser i rapporten *Befolkningens Løn 2012*
2. common sense vurdering af lønniveauer for forskellige genkendelige faggrupper
3. Den relationelle forskel mellem de forskellige lønninger, og hvorvidt denne relationelle forskel er konsistent med de relationelle forskel jeg ser i andre af DSTS indkomstvariable

A.2 Centrale mål

Nedenstående tabel A.1 viser de centrale mål for hvert år af `TIMELON`. Disse er ikke korrigeret for inflation, da jeg har valgt ikke at omregne hele populations individuelle lønninger med forbrugerprisindekset, og i stedet kun valgt de relevante gennemsnitsværdier. Standardafvigelse og percentiler kan naturligvis ikke vises medmindre hele populations observationer er i samme enhed, derfor denne fremstillingsform.

Det bemærkes at ingen har en timeløn under 11 kr. Det skyldes at jeg har fjernet alle observationer der tjener mellem 0 og 10 kr, da dette forekom som en kunstigt lav timeløn, der kun kunne forekomme, grundet udregningsmetoden i `TIMELON`-variablen. Dette ville, ifølge mit skøn, give et misvisende billede af den reelle timeløn indenfor en given disco-kategori. I gennemsnit fjernes 369 personer pr. år på denne måde, altså et ubetydeligt bortfald.

Det ses desuden i tabel A.1 der i 2003 sker et lille fald i gennemsnitslønnen, fra 185,5 kr/t til 184,8 kr/t. Det efterfølgende år sker der en relativt kraftig stigning til 195,2 kr/t. Dette skyldes databrud hos DST. I denne henseende er det en fordel at min datastruktur kræver et gennemsnit over den 14-årige periode, da et databrud i en kortere periode indenfor tidsrammen derfor er af mindre betydning.

Tabel A.1: centrale mål for `TIMELON`, 1996-2009

År	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
N	1677704	1738875	1755089	1789319	1624672	1638380	1582448	1508256	1512068	1530139	1526006	1587350	1521759	1647924
Gennemsnit (2012 priser)	211,6	208,5	216,1	215,5	217,5	223,5	221,8	223,5	219,8	227,8	232,0	234,9	236,90	239,35
Gennemsnit (2009-priser)	196,8	193,9	201,0	200,4	202,2	207,8	206,3	207,9	204,4	211,9	215,7	218,4	220,31	222,58
Gennemsnit	150,6	151,6	159,8	163,9	169,8	177,7	181,3	185,6	184,8	195,2	202,3	209,8	217,4	222,6
Sdafvigelse	67	66	72	74	78	83	84	84	84	90	99	102	113	108
Mindste værdi	11	11	11	11	11	11	11	11	11	11	11	11	11	11
Højeste værdi	9974	5970	5952	8772	8019	9826	6850	7158	15165	8323	29429	11940	16977	28752
p25	115	116	121	125	129	135	137	141	141	149	154	159	164	169
p50	139	140	147	151	156	163	167	171	170	180	187	194	200	204
p75	170	171	181	185	192	201	205	210	208	220	228	237	245	250
p90	218	218	232	237	245	257	262	267	265	279	289	300	312	318
p99	379	380	408	418	433	456	467	476	477	499	521	548	580	580

Jeg har benyttet forbrugerindekset til to korrektioner af lønningerne. omregningen til 2012 priser sker for at kunne sammenligne direkte med DSTs rapport fra 2012, og omregningen til 2009 priser bliver benyttet i afhandlingen. Det kan måske forekomme omstændigt, men det virker mest hensigtsmæssigt at benytte et prisniveau fra et år i den tidsperiode, der rent faktisk er i analysen, fremfor et prisniveau 3 år længere fremme i tiden, og som udelukkende er med for at kunne teste validiteten af variablen op imod DSTs rapport.

Tabel A.2: centrale mål for TIMELON, 1996-2009

	2009 priser	2012 priser	DST 2012
gennemsnitsløn	207,8	223,5	243,0
N	1.617.142		1.353.000

Det ses af tabel A.2, at den gennemsnitlige timeløn i perioden 1996-2009, omregnet til 2012-priser, er 223,5 kr/t. i DST rapporten omhandlende samme år, er de kommet frem til en gennemsnitsløn på 243 kr/t, altså en forskel på 19,5 kr i timen. Det er naturligvis en klar forskel. Det er dog ikke så bekymrende, når man medtager det forhold, at der faktisk ikke er tale om to direkte sammenlignende tal: DSTs gennemsnitslønning er et forsøg på at ramme den reelle gennemsnitsløn for et bestemt år, her 2012. Mit formål er imidlertid at beskrive lønniveauet i en 14-årig periode. Det betyder at min gennemsnitslønning også indeholder strukturelle forskydninger og konjunktursvingninger i økonomien, som den har udviklet sig over de 14 år. Det er derfor forventeligt, at den ikke rammer præcis gennemsnitslønnen, som den så ud i et enkelt år, her 2012. I denne optik er forskellen på 19,5 kr på ingen måde bekymrende. Det ses også, at mit N har 264.142 flere personer i gennemsnit over årene, end DST benyttede til deres rapport i 2012. Datarensningen i DSTs rapport fremgår ikke tydeligt, så det er ikke muligt at klargøre hvorfor denne forskel eksisterer. Det er dog ikke bekymrende, når gennemsnittet alligevel ligger acceptabelt tæt.

A.3 sammenligning med DSTs rapport

Grundet omkodninger og xx (det ord Søren bruger) fra et 4-cifret DISCO-niveau til mit, omtalt i Bilag (#henvisning), er det ikke muligt at sammenligne direkte med DSTs rapport om timelønninger fra forskellige faggrupper. I 19 tilfælde svarer mit DISCO-variabel imidlertid direkte til DSTs oprindelige 4-cifrede niveau. Det er således muligt at lave en sammenligning mellem mit data og DSTs timelønssrapport, for at få en indikator på validiteten af mine timelønninger. Disse 19 tilfælde er opgjort i Tabel A.3. Det viser timelønnen i mit data, i DST rapporten, forskellen i timellen opgjort i absolutte værdier samt i procent. Rangeringen går fra højeste positive forskel til mellem DST rapportens timeløn for faggruppen og min timeløn for faggruppen, til højest negative forskel ditto. Det vil sige at at en positiv forskel i timeløn, rangeret øverst, betyder at mine tal ligger *under* x antal kroner under DSTs timelønssats, og en negativ forskel i timeløn, rangeret lavest, betyder at min timelønssats ligger x antal kroner *over* DSTs timelønssats. Endelig er den gennemsnitlige (numeriske) forskel opgjort nederst, i absolut værdi og i procent. Min timelønssats er som omtalt korrigeret til 2012-priser.

Tabel A.3: Sammenligning med DST rapport

Discokode	Eget data, sdaifvigelse 2009	Eget data, gns.	DST rapport, gns.	Forskel	Forskel i procent
110: Militært arbejde	78	218	312	94	29 %
7139: Bygningsarbejde (finish) Elektriker	75	192	220	28	12 %
3111: Teknikerarbejde inden for fysik, kemi, astronomi, meteorologi, geologi mv	55	193	219	26	11 %
7124: Bygningsarbejde (basis), toemrer- og snedkerarbejde	64	186	205	19	9 %
4110: Alment kontorarbejde	60	182	199	17	8 %
3118: Teknisk tegnearbejde	48	200	214	14	6 %
7122: Bygningsarbejde (basis), murer- og brolægningarbejde	65	197	210	13	6 %
3112: Teknikerarbejde vedr. bygninger og anlæg	72	246	259	12	4 %
2141: Ingeniører og arkitekter	108	306	317	11	3 %
2221: Læge	155	391	401	10	2 %
2411: Overordnet revisions- og regnskabsarbejde, herunder registeret revisor og statsautoriseret revisor	172	300	308	8	2 %
2331: Folkeskolelærere	49	219	224	5	2 %
9130: Rengørings- og køkkenhjælpsarbejde	48	158	162	4	2 %
3113: Teknikerarbejde vedr. elektriske anlæg mv	72	266	267	1	0 %
3114: Teknikerarbejde vedr. elektroniske anlæg mv	77	257	251	-6	-2 %
3115: Teknikerarbejde vedr. maskiner og roerianlæg, eksklusiv vedligeholdelse af maskiner om bord paa skibe	74	265	250	-14	-5 %
8322: Kørsel af hyre- og varevogn m.v.	90	196	170	-26	-15 %
2412: Udvikling og planlægning af personalespørgsmaal	160	327	296	-31	-10 %
6130: Arbejde med saavel markafgrøder som husdyr, fx som landmand	79	195	159	-36	-22 %
1211: Ledelse 10+ ansatte: overordnet og/eller tværgående ledelse i virksomheder, herunder administrerende direktør, bankdirektør, kreditforeningsdirektør, varehusdirektør	409	480	398	-82	-20 %
Gennemsnitlig forskel	-	-	-	24	9 %

Det ses af tabellen, at den gennemsnitlige timelønssforskel ligger på 24 kr/t, hvilket stemmer nogenlunde overens med den overordnede gennemsnitlige forskel på 19,5 kr/t. Det ses at tre faggrupper afviger ganske betragteligt fra DSTs rapport. I faggruppen 1211: Ledelse 10+ ansatte overvurderer jeg deres lønninger med 82 kr/t eller 20 %. Der er tale om en gruppe, hvis lønninger varierer ganske betragteligt, hvilket ses på standardafvigelsen fra år 2009¹. Her ses det, at 1211: Ledelse 10+ ansatte har en voldsomt høj standardvigelse, hvilket giver fin mening, indenfor en gruppe ledere af små virksomheder med eksempelvis 11 ansatte, samt chefer for en stribe af landets største virksomheder. Det giver derfor ikke anledning til bekymring, og må nok anses som et særtilfælde. De to andre kategorier med store udsving, 110: Militært arbejde, hvor jeg undervurderer deres lønninger med med 78 kr/t eller 29 %, samt 6130: Arbejde med såvel markafgrøder som husdyr, hvor forskellen er med 79 kr/t eller 22 % kan ikke umiddelbart forklares. Det eneste, der måske kan forklare det, er at der er tale om meget brede kategorier, der indeholder en række meget forskellige typer jobs, selv på dette nære DISCO-niveau. Det er værd at holde sig for øje om de brede kategorier, at de muligvis er mere upålidelige end de mere specifikke erhvervskategorier, omend det måske er meget at konkludere ud fra disse sparsomme sammenligninger. Jeg vil derfor på anden vis prøve at vurdere, om jeg kan stole på disse kategorier. Derfor vil jeg benytte den i indledningen omtalte 3. metode, nemlig den relative forskel i indtjening mellem faggrupperne, set over flere af DSTS indkomstvariable.

¹ Da jeg kun har korrigeret gennemsnittene for de enkelte faggrupper, og ikke hver enkelt observation i registerdataen, kan jeg ikke inflationskorrigerer standardafvigelsen. Jeg har derfor benyttet variansen for 2009, der derfor ikke kan sammenlignes direkte, men alligevel giver en ganske udemærket indikation af variansen omkring gennemsnittet

A.4 relative lønforskelle

Det er vigtigt, at de lønninger afspejler virkeligheden, men det allervigtigste må være, at den *relative forskel* mellem faggrupperne er pålidelig. Derfor har jeg sammenlignet med 4 andre af DSTs indkomstvariable, for at vurdere om den relative forskel i indtjening er nogenlunde ensartet over 5 variable. De fire andre indkomstvariable er `DISPON_NY`, `loen_mv`, `perindkialt` samt `joblon`, og er omtalt i indledning på s. side 36. En nærmere analyse ville være for omfattende at komme ind på her, men det skal kort nævnes, at den relative forskel er nogenlunde ensartet, taget i betragtning af de vidt forskellige opgørelsesmetoder. Det vigtige i denne sammenhæng er de tre førnævnte faggrupper ligger relativt pålideligt. 1211: Ledelse 10+ ansatte ligger som nr. 1 i alle indkomstvariablene, så det er meget tilfredsstillende. 110: Militært arbejde ligger ikke på præcis samme placering i alle de 5 variable, men ligger omtrent i midten i alle 5 variable, hvilket også må anses som tilfredsstillende. 6130: Arbejde med såvel markafgrøder som husdyr derimod ligger som nr. 93 i `timelon`, men placerer sig lige fra en bundkategori på 132. plads over indtjening i `loen_mv`, noget tilsvarende i `joblon`. Altså de andre to andre *lønnings* variable. Mens faggruppen i `DISPON_NY` og `perindkialt` placerer sig som henholdsvis nr. 36 og 17, altså ganske tæt på Danmarks bedst tjenende faggrupper i *overordnet indkomst*, som disse to variable måler. Hvad dette skyldes, kan være svært at sige, men i forhold til, om `timelon` er et pålideligt lønmål, er det ganske fint, at der tilsyneladende "bare" er tale om en DISCO-kategori, hvori deres måde at tjene penge på er sådan skruet sammen, at opgørelsesmetoden bliver vigtig. At arbejde med markafgrøder og husdyr er et specielt erhverv, i forhold til de mere almindelige lønarbejdsformer, vi ellers arbejder med, og som tabel A.3 ser ud til at dække ganske fint. Det gælder heldigvis et fåtal af faggrupperne, som set på s. #ref.

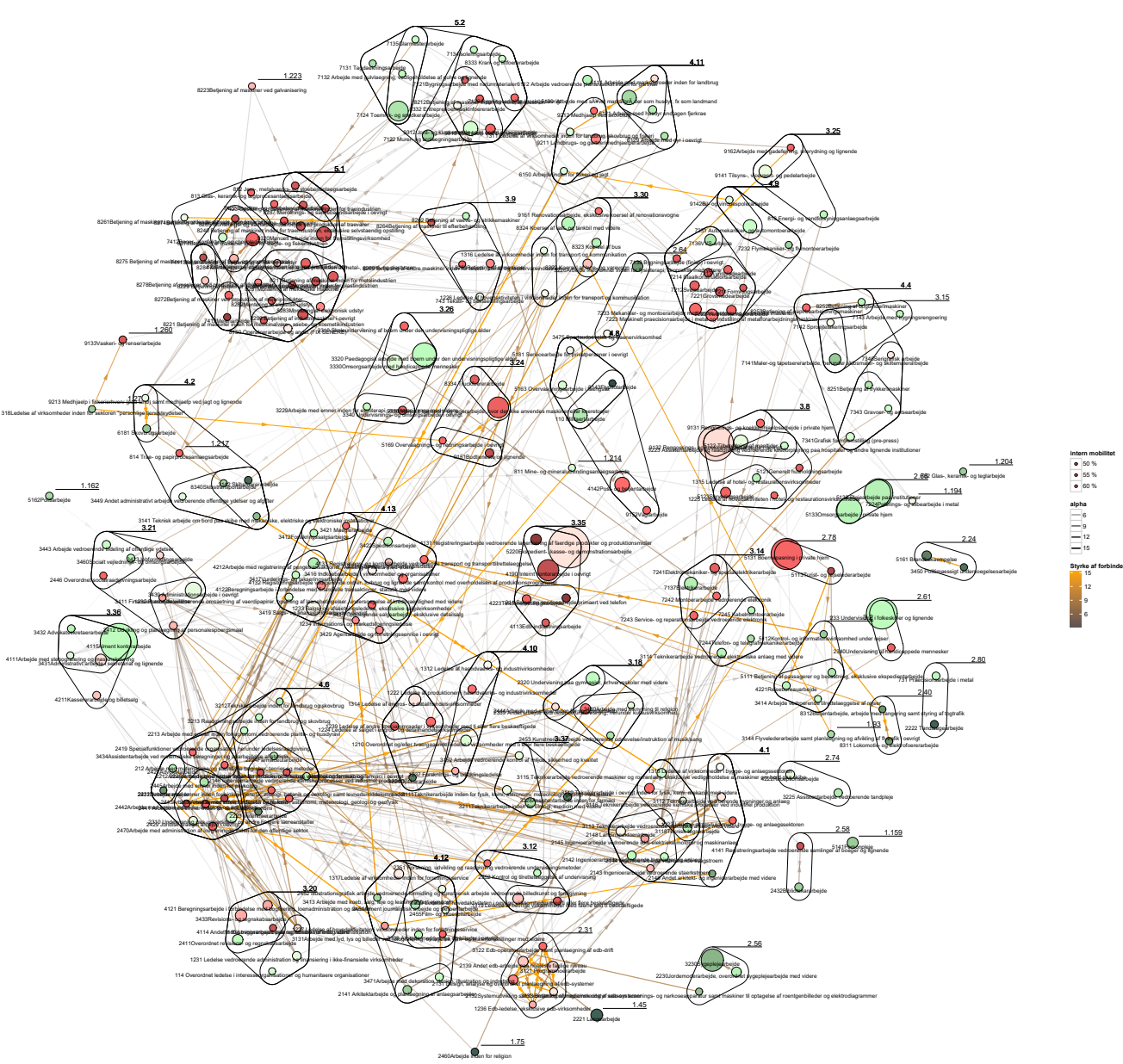
A.5 sammenfatning

Sammenfattende må man konkludere, at mit mål for timeløn er validt, selv indenfor så detaljeret et DISCO-niveau som dette. de tre faggrupper, der giver anledning til bekymring, kan i høj grad forklares ved deres særlige stilling i samfundet, og af disse to - ledere med over 10 ansatte samt militæret - viser det sig i sammenligning af relative indkomstforskelle at der ikke er anledning til bekymring. Den sidste faggruppes timelønsskøn må anses som netop dette, et skøn, omed forskellen mellem DSTs opgørelse og min er på 79 kr/t eller 22 %. 22 % er naturligvis en del, men ikke alarmenerende, og i betragtning af faggruppens særlige karakter giver det ikke anledning til tvivl om timelønsvariablens validitet i langt de fleste

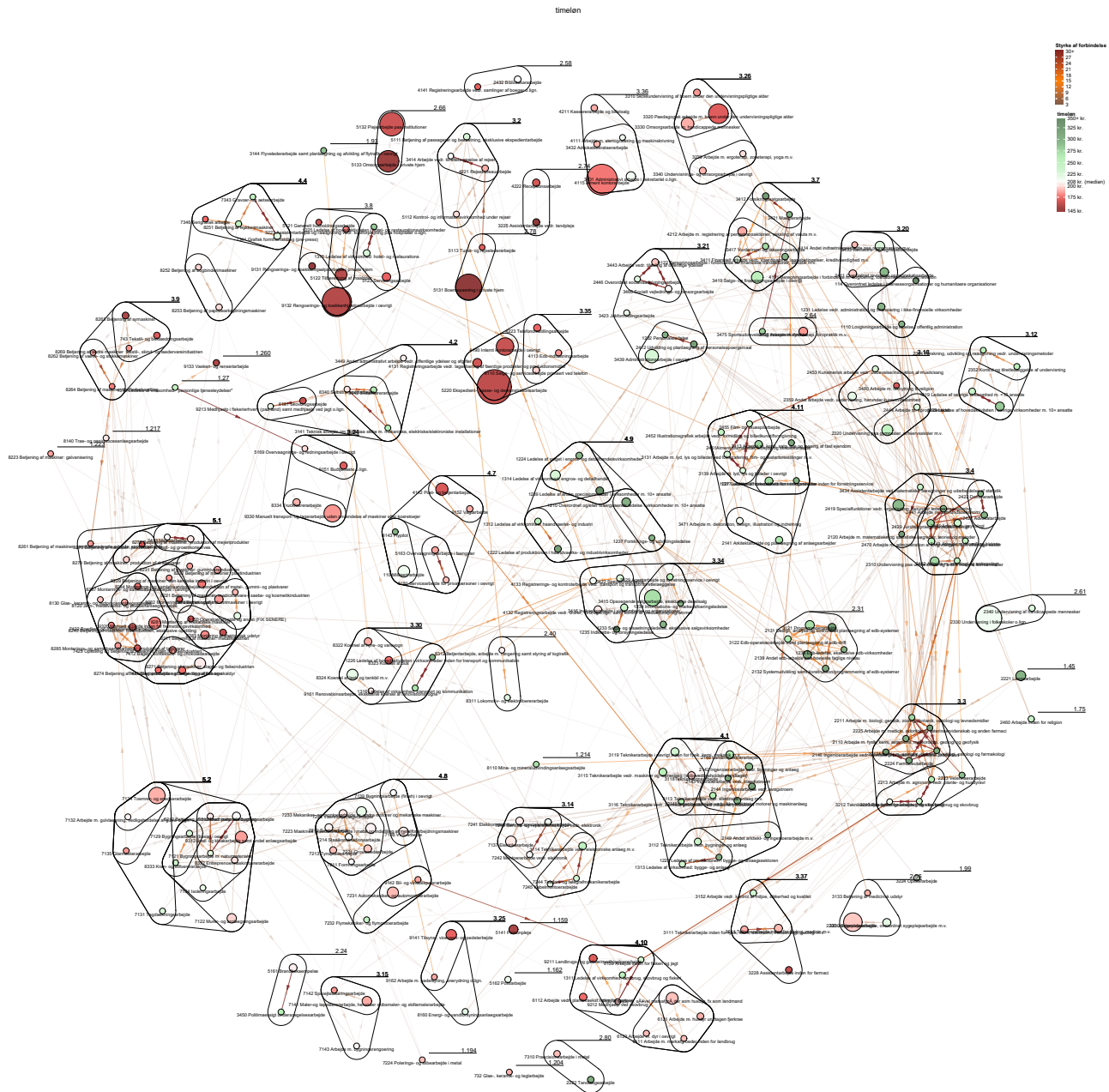
tilfælde. Det, der skal tages med forbehold i analysen, må være at for faggrupper, hvis arbejdsindhold har en ganske særlig karakter, kan timelønvariablen være unøjagtig, dog uden at være alarmende unøjagtig. Udover det er "stikprøven", der er mulig at sammenligne med DSTs rapport, i god overensstemmelse, og jeg vil derfor kun med det enkelte ovenstående forbehold i mente, benytte `timeløn` i analysen.

B netværkskort

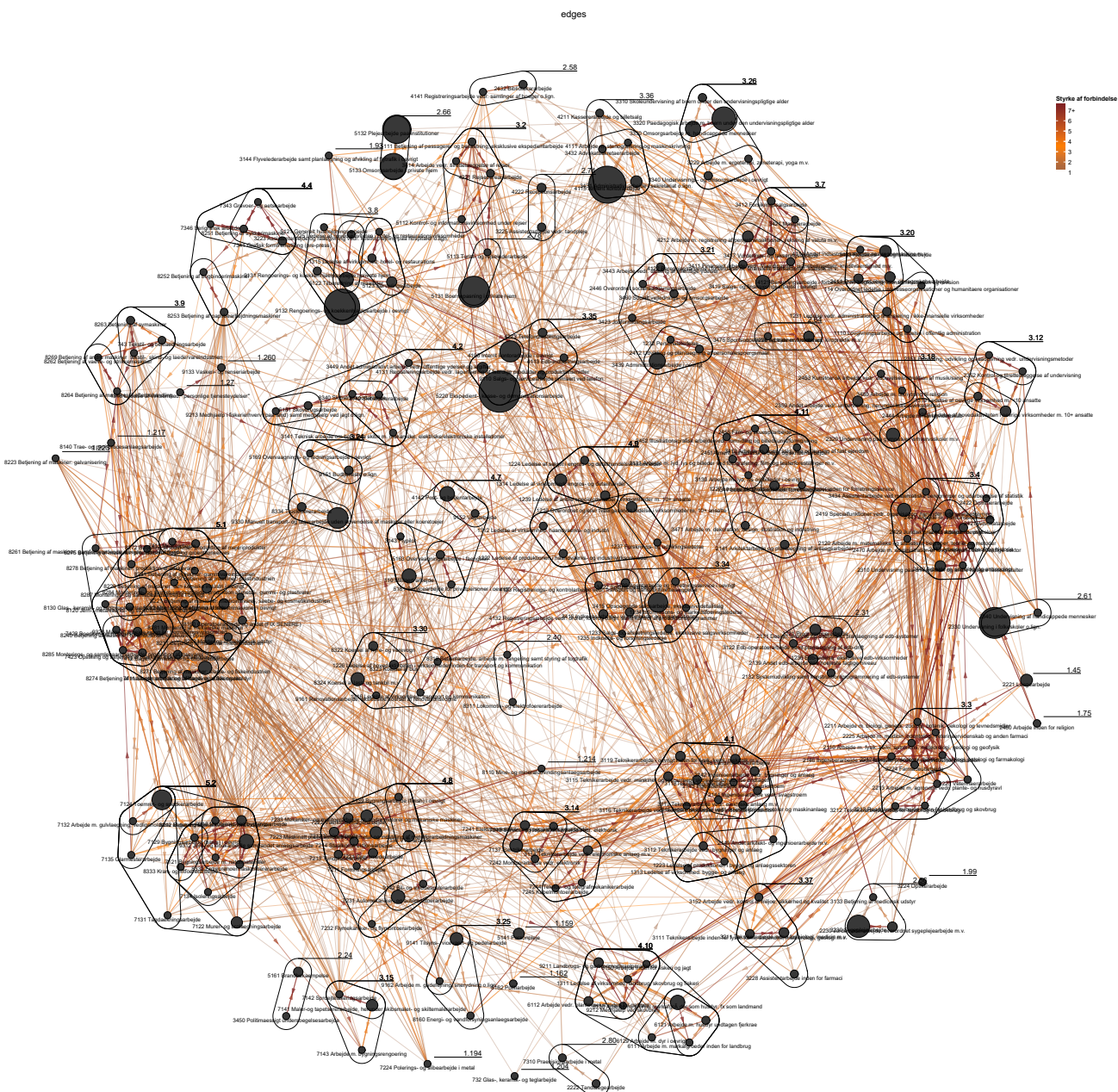
Figur B.1: Netværkskort: Intern mobilitet



Figur B.2: Netværksskørt: Timelønninger



Figur B.3: Netværkskort: detaljerede forbindelser/edges



Litteraturliste

Bøger og artikler

- [1] Alan Agresti og Barbara Finlay. *Statistical Methods for the Social Sciences*. New Jersey: Prentice Hall, 1997.
- [2] Thomas P. Boje. *Segmentering og beskæftigelse*. Roskilde: Roskilde Universitetscenter, 1985.
- [3] Thomas P. Boje og Christian Toft. *Arbejdsmarkeder og segmenteringsteorier, en introduktion til nyere segmenteringsteorier*. Aabenraa: Institut for grænseforskning, 1989.
- [4] Thomas Frank. *What's the Matter with Kansas?: How Conservatives Won the Heart of America*. Henry Holt og Company, 2007. ISBN: 9781429900324. URL: <https://books.google.dk/books?id=AJKrMcOyQ3wC>.
- [5] David M Gordon, Richard Edwards og Michael Reich. *Segmented work, divided workers: The historical transformation of labor in the United States*. New York: Cambridge University Press, 1982.
- [6] Edward Laumann, Peter V. Marsden og David Prensky. „Applied Network Analysis: A Methodological Introduction“. I: udg. af Ronald S. Burt og Michael J. Minor. Sage, 1983. Kap. The Boundary Specification Problem in Network Analysis, s. 18–34.
- [7] Nikolaj Malchow-Møller og Allan Würtz. *Indblik i Statistik*. København: Gyldendal, 2003.
- [8] Magnus E. Marsdal. *Frp-koden: hemmeligheten bak Fremskrittspartiets suksess*. Forlaget Manifest, 2007. URL: <https://books.google.dk/books?id=x0nbMQEACAAJ>.
- [9] Daniel Oesch. *Redrawing the Class Map: Stratification and Institutions in Britain, Germany, Sweden and Switzerland*. Palgrave Macmillan UK, 2006. ISBN: 9780230504592. URL: https://books.google.dk/books?id=6%5C_ZZCwAAQBAJ.
- [10] Michael J. Piore. „Economic Fluctuation, Job Security, and Labor-Market Duality in Italy, France, and the United States“. I: *Politics & Society* 9.4 (1980), s. 379–407. DOI: 10.1177/003232928000900401. eprint: <http://pas.sagepub.com/content/9/4/379.full.pdf+html>. URL: <http://pas.sagepub.com/content/9/4/379.short>.
- [11] John Scott. *Social Network Analysis: A Handbook*. Sage, 2000.

-
- [12] Jonas Toubøl, Anton Larsen og Carsten Strøby. „A network analytical approach to the study of labour market mobility“. I: *Paper presented at XXXIII Sunbelt Social Networks Conference of the International Network for Social Network Analysis (INSNA), Hamborg, Germany* (2013).
- [13] Max Weber. „Economy and society: an outline of interpretive Sociology“. I: Berkeley: University of California Press, 1978. Kap. IV Status Groups and Classes, s. 302–310.

Online kilder

- [14] DST-DISCO88. *Danmarks Statistiks beskrivelse af DISCO-88*. 2016. URL: <http://www.dst.dk/da/Statistik/dokumentation/Nomenklaturer/DISCO-88/Introduktion>.
- [15] DST-Forbrugerprisindeks. *Forbrugerprisindekset i Danmark*. 2016. URL: <http://www.dst.dk/da/Statistik/emner/forbrugerpriser/forbrugerprisindeks>.
- [16] DST-TIMELON. *Danmarks Statistiks beskrivelse af variabelen ANSXFREM*. URL: <http://www.dst.dk/da/Statistik/dokumentation/Times/ida-databasen/ida-ansaettelser/timelon>.

Danmarks Statistiks Manualer

- [17] DST. *Befolkningens løn 2012*. 2013. URL: <http://www.dst.dk/da/Statistik/Publikationer/VisPub?cid=19581>.
- [18] DST. *Indkomster 2009*. 2009. URL: <http://www.dst.dk/da/Statistik/Publikationer/VisPub?cid=015219>.