

Om forfatterne



Nicoline Sørensen
nicoline.sorensen@gmail.com



Andreas Østed
andreas_osted@live.dk



Clara Wolff
clara.jespersen@outlook.dk



Rebekka Thomsen
rebekka.kaysen@gmail.com

Alle studerende på
Datavidenskab, Aarhus
Universitet.



Palle Villesen er lektor,
ph.d. ved Center for
Bioinformatik (BiRC),
Aarhus Universitet.
Han har været vejleder
på projektet omtalt i
artiklen.
palle@birc.au.dk



ET DATAANALYTISK PERSPEKTIV PÅ FÆDREBARSEL

Nybagte fædre holder i gennemsnit 34 dages barsel. En gruppe studerende i datavidenskab har undersøgt, hvad dette gennemsnit egentlig dækker over. De finder store forskelle i fædres barsel afhængig af uddannelsesniveau, typen af job, og hvor i landet man bor.

I alle samfundsrelevante problemstillinger – lige fra klimaets udvikling over pandemihåndtering til arbejdsmarkedsforhold – spiller data en afgørende rolle. Hvordan vi indsamler og bruger disse data er i høj grad med til at afgøre, hvad vi som et videnssamfund opnår af viden om problemstillingerne og beslutter af mulige løsninger.

Set fra et dataanalytisk perspektiv er det ikke afgørende, om det er en naturvidenskabelig eller samfundsaglig problemstilling, man kigger på. Det er grundlæggende de samme matematiske metoder, man tager i anvendelse for at finde ud af, hvad data kan fortælle os.

Vi er fire studerende på Datavidenskab ved Aarhus Universitet, som har lavet et projekt, hvor vi analyserer offentligt tilgængelige data om fædrebarsel. Vi valgte emnet, fordi fædrebarsel har været, og stadig er, et højaktuelt, omdiskuteret emne i medierne. Vi ønskede at bruge dataanalyse og matematisk modellering til at undersøge og udfordre

det generaliserende billede af fædrebarsel i medierne.

I medierne er udgangspunktet for argumentation om fædrebarsel ofte det gennemsnitlige antal barselsdage, som i 2019 (nyeste tilgængelige data) var cirka 34 arbejdsdage (knap 5 uger) for fædre. Men hvad dækker gennemsnittet egentlig over,

Ny barselsaftale

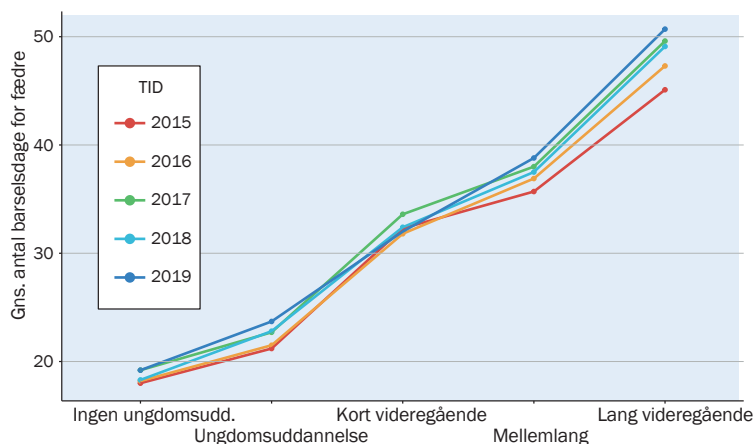
EU vedtog et nyt tiltag i 2019, som indebærer, at alle forældre i medlemslandene skal have to måneders forældreorlov, der ikke kan overdrages til deres partner – dette kaldes også for øremærket barsel. På den baggrund indgik folketinget 2. oktober 2021 en ny barselsaftale, som trådte i kraft 2. august 2022. Den nye aftale betyder, at fædre fremover vil have 11 ugers øremærket barsel i modsætning til den nuværende lovgivning, hvor de har to ugers øremærket barsel.

når vi dykker ned i data? Når vi med matematisk modellering undersøger fædrebarsel på flere parametre såsom fædrenes uddannelsesniveau, arbejdsbranche og geografisk placering, kan vi se, at der er store forskelle i gennemsnittet grupperne imellem. Derfor er det ikke retvisende at bruge det samlede gennemsnit for alle fædre, når der skal debatteres fædrebarsel. Netop det vil vi gerne belyse yderligere ved at kigge nærmere på data på disse områder.

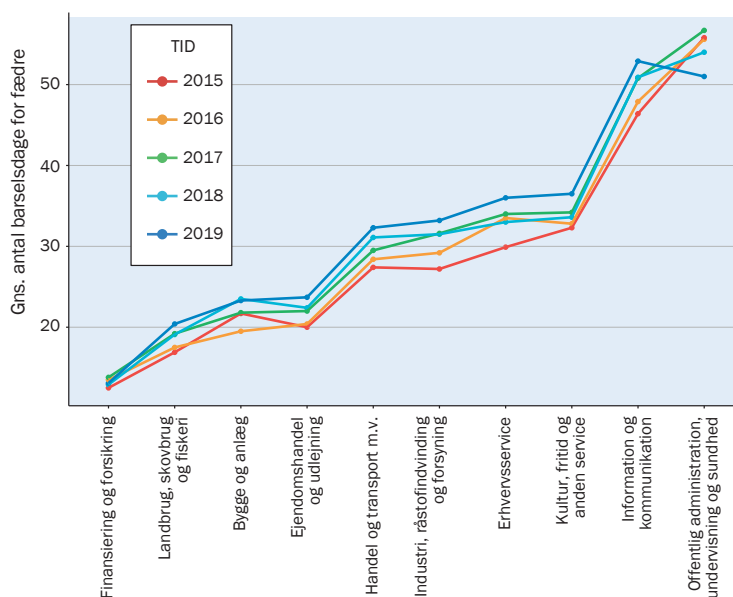
Uddannelsesniveau og branche påvirker barselslængden

Det gennemsnitlige antal barselsdage for fædre indenfor forskellige uddannelsesniveauer viser en voldsom forskel (figur 1). Fædre uden ungdomsuddannelse holder i gennemsnit under 20 dages barsel, mens fædre med en lang videregående uddannelse holder mere end 50 dages barsel i gennemsnit – altså over 2,5 gange så meget. Der er et tydeligt billede af, at højere uddannelsesniveau resulterer i mere fædrebarsel. En matematisk model, som tager højde for fædrenes uddannelsesniveau, forklarer hele 70% af variationen – altså kan cirka 70% af, hvor meget barsel mænd i gennemsnit holder, forklares alene ud fra deres uddannelsesniveau.

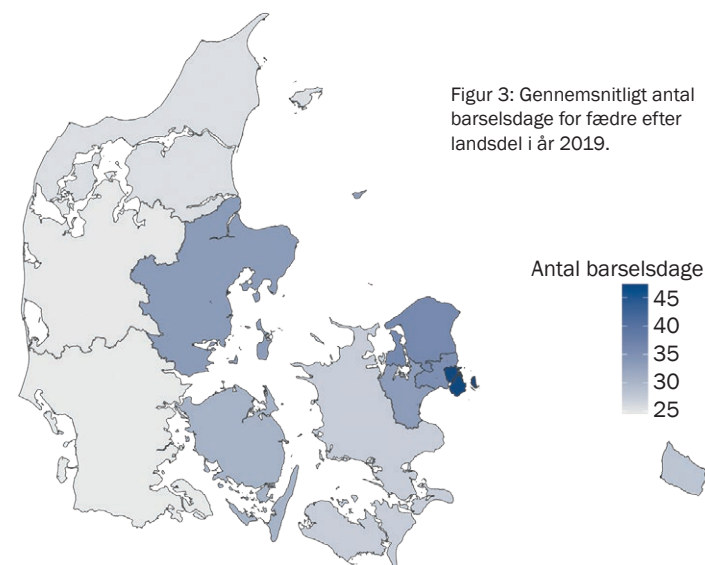
Der er også store forskelle i barselslængde indenfor forskellige brancher (figur 2). Fædre ansat indenfor brancherne “landbrug, skovbrug og fiskeri” og “byggeri og anlæg” holder under 20 dages barsel, mens fædre indenfor den offentlige sektor, samt “finansiering og forsikring” holder over 50 dages barselsorlov. Intuitivt vil man forvente en høj korrelation mellem branche og uddannelsesniveau, som man bør tage højde for. En branche som “finansiering og forsikring” vil formodentligt have en højere andel ansatte med højt uddannelsesniveau og derfor også et højere gennemsnitligt antal barselsdage. Derudover vil man formode, at der er en naturlig forskel brancherne imellem, da der er forskel på overenskomster og generelle forhold i hver branche. For eksempel kan forskellen skyldes,



Figur 1: Gennemsnitligt antal barselsdage for fædre mod uddannelsesniveau efter årstal.



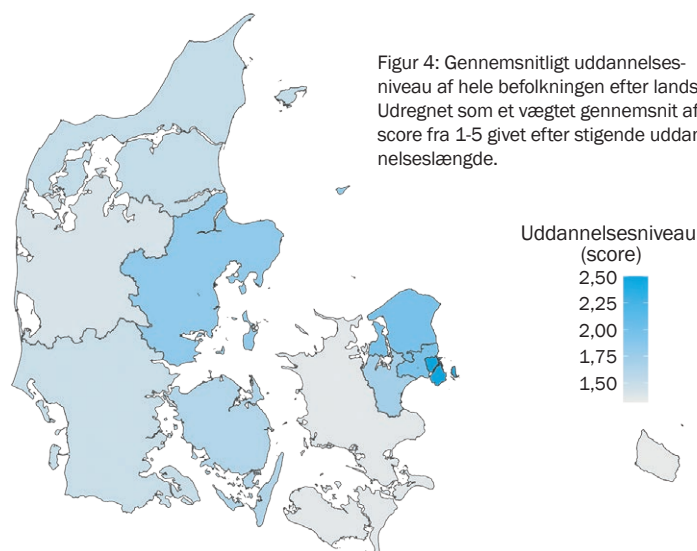
Figur 2: Gennemsnitligt antal barselsdage for fædre mod brancher efter årstal.



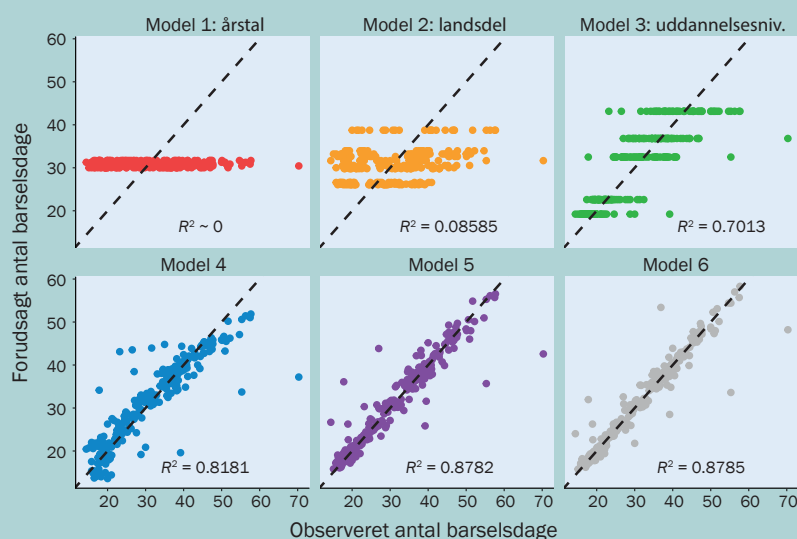
Figur 3: Gennemsnitligt antal barselsdage for fædre efter landsdel i år 2019.

at ansatte i den offentlige sektor har adgang til gode barselsforhold, mens der i landbruget ikke er samme muligheder, blandt andet fordi

andelen af selvstændige er højere i denne branche, og de derfor ikke har en overenskomst, som sikrer adgang til barsel.



Figur 4: Gennemsnitligt uddannelsesniveau af hele befolkningen efter landsdel. Udregnet som et vægtet gennemsnit af en score fra 1-5 givet efter stigende uddannelseslængde.



Figur 5: Figurene viser det forudsagte antal barselsdage mod observeret antal barselsdage vist med modeller af stigende kompleksitet.

Hvert punkt viser et gennemsnitligt antal barselsdage for et bestemt år i en bestemt landsdel, hvor faderen har et bestemt uddannelsesniveau. Den stiplede linje ved hver model viser en ret linje med hældning 1 og skæring i 0,0. Jo tættere punkterne ligger på denne linje, jo bedre forklarer modellen de observerede data.

Model 1, 2 og 3 viser gennemsnitligt antal barselsdage efter henholdsvis årstal, landsdel og uddannelsesniveau.

Model 4 viser barselsdage efter den lineære kombination af både årstal, landsdel og uddannelsesniveau.

Model 5 viser barselsdage efter den lineære kombination af både årstal, landsdel og uddannelsesniveau, samt interaktionen af landsdel og uddannelsesniveau, hvilket dermed tillader, at effekten af uddannelsesniveau på antal barselsdage afhænger af landsdelen.

Model 6 viser gennemsnitligt antal barselsdage efter interaktionen mellem både årstal, landsdel og uddannelsesniveau. Dette tillader dermed, at effekten af uddannelse på antal barselsdage afhænger af både året og landsdelen. Denne model viser sig dog ikke at være signifikant bedre end model 5 til at beskrive dataene, hvorfor den er farvet grå.

Ved hver model er desuden angivet forklaringsgraden, R^2 , justeret efter kompleksiteten af modellen.

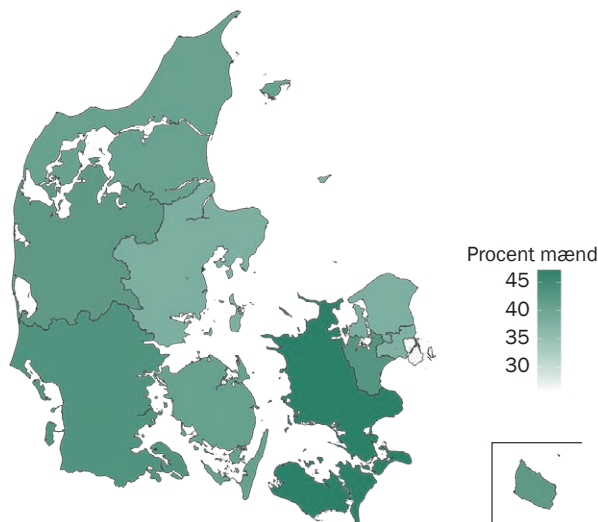
Uddannelse har ikke samme effekt i alle landsdele

Ser vi mere geografisk på emnet, er der også meget store forskelle i længden af fædrebarsel på tværs af de forskellige landsdele i Danmark (figur 3). De største kontraster ses mellem landsdelene med uddannelsesbyer og resten af landet. I landsdelen "Byen København" tager fædre i gennemsnit cirka 47 dages barsel, i Nordsjælland 36 dage, mens Vestjylland og Sydjylland begge ligger på cirka 25 dage. Der er mange mulige forklaringer, da hver landsdel jo har forskellige sammensætninger i forhold til uddannelsesniveau, branchefordeling og alder ligesom der måske kan være kulturelle forskelle landsdelene imellem.

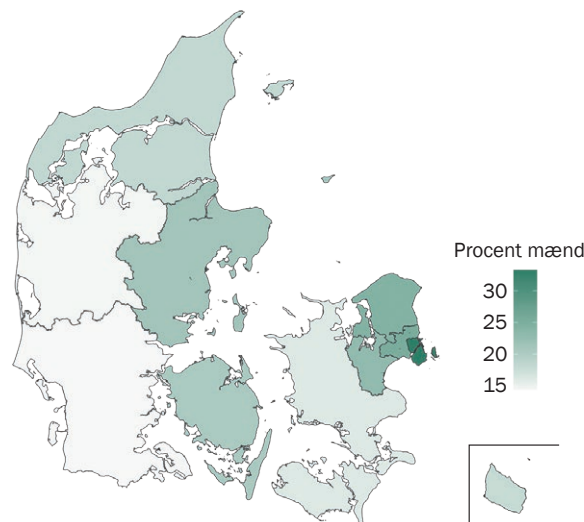
For at undersøge, om uddannelsesniveauet er forskelligt i landsdelene, har vi udregnet et vægtet gennemsnitligt uddannelsesniveau for hver landsdel (figur 4). Geografiske forskelle i uddannelsesniveau viser samme tendens som gennemsnitligt antal barselsdage. Dette er ingen overraskelse, da uddannelsesniveau jo hænger tæt sammen med antallet af barselsdage. Så kan det generelle uddannelsesniveau i landsdelene sikkert forklare en stor del af forskellen i antal barselsdage på tværs af landsdelene.

For at undersøge dette nærmere kan man bruge matematisk modellering, hvor der tillades forskellig effekt for hver af kombinationerne af landsdel og uddannelsesniveau. Til det formål har vi opstillet 6 modeller (såkaldte multiple regressionsmodeller) med stigende kompleksitet (figur 5). Disse viser effekten af de forskellige variable "uddannelse", "landsdel" og "tid (årstal)" – alene eller i kombination og med uden interaktion mellem variable – på det gennemsnitlige antal barselsdage.

Det viser sig, at en model, der indeholder alle variable på én gang og som også tager højde for interaktionen mellem uddannelse og landsdel (og dermed tillader, at effekten



Figur 6: Procent af mænd i hver landsdel, som arbejder inden for én af brancherne, hvor der bliver holdt færrest barselsdage.



Figur 7: Procent af mænd i hver landsdel som arbejder inden for én af brancherne, hvor der bliver holdt flest barselsdage.

af uddannelse er forskellige i for eksempel Vestjylland og København) forklarer hele 88% af variationen i vores data. Denne model viser eksempelvis, at effekten af en lang videregående uddannelse på længden af fædrebarsel kun er halvt så stor i Vestjylland som i Nordsjælland.

Branchen har større betydning end uddannelsesniveauet

Vi forventer som tidligere nævnt en stærk korrelation mellem, hvilken branche fædrene arbejder i og deres uddannelsesniveau. Derfor har vi opstillet en anden matematisk model, hvor vi undersøger effekten af de brancher, som indgår i vores data, social status, tid, samt interaktionen af social status og branche på det gennemsnitlige antal barselsdage for fædrene. Social status er angivet som enten "Lønmodtager på grundniveau", "Lønmodtager på mellemniveau", eller "topledere/lønmodtager på højeste niveau". Denne model forklarer hele 95% af variationen i dataene. Modellen fortæller blandt andet, at inden for "Industri, råstofindvinding og forsyningsvirksomhed", har man meget stor gevinst af at være topledere eller lønmodtager på højeste niveau, mens der inden for "offentlig administration, undervisning og sundhed" ikke ses en statistisk signifikant effekt af samme sociale status. Igen kan dette formodentlig forklares af forskellige overenskomster og forhold brancherne imellem.

Ydermere kan vi se, at der er stor forskel på fordelingen af ansatte i de forskellige brancher på tværs af landsdelene (se figur 6 og 7). I landsdele med få barselsdage er procentdelen af mænd i brancher med kort barsel stor. Modsat er procentdelen af mænd i brancher med et højt antal barselsdage stor i de landsdele med et højt antal barselsdage. Dette underbygger den store effekt af branche på fædrebarsel.

Sammenligning af landsdele gør forskellene tydelige

Det kan være svært at forholde sig til matematiske modeller, så

for at gøre modellernes resultater mere tydelige, har vi valgt at sammenligne landsdelene Vestjylland og Nordsjælland, da det er to landsdele med stor kontrast indenfor de undersøgte parametre. Med udgangspunkt i disse to landsdele, undersøger vi, hvor mange der arbejder i de forskellige brancher. Den største forskel finder vi i branchen med færrest barselsdage "Landbrug, skovbrug og fiskeri" samt branchen med flest barselsdage "Finansiering og forsikring". Således er procentdelen af personer i branchen "Landbrug, skovbrug og fiskeri" markant større

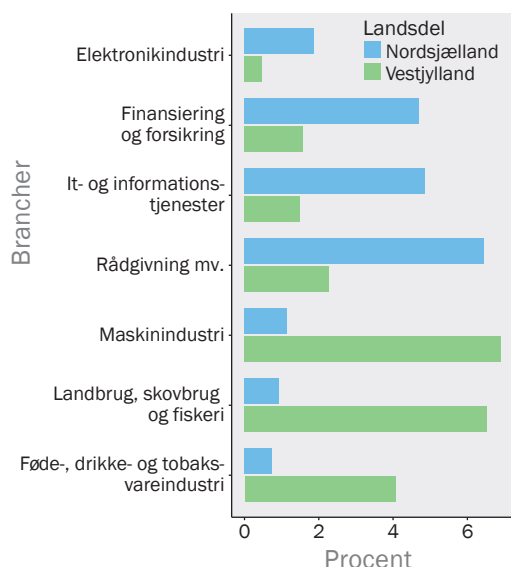
Multipel regression

Lineær regression er en matematisk metode til at undersøge sammenhængen mellem to variable – den uafhængige (også kaldet forklarende) variabel og den afhængige variabel. Da der i simpel lineær regression kun er en enkelt uafhængig variabel, kan man plote den afhængige variabel (y) mod den uafhængige variabel (x) i et almindeligt 2-dimensionalt koordinatsystem. Multipel regression er en naturlig udvidelse af lineær regression, men her er der to eller flere uafhængige variable til at beskrive den afhængige variabel. For mere end to uafhængige variable er det derfor ikke muligt at visualisere punkterne i hverken et 2-dimensionalt eller 3-dimensionalt koordinatsystem. Nedenfor er vist forskellen på en simpel lineær regressionsmodel og en multipel regressionsmodel (med k afhængige variable):

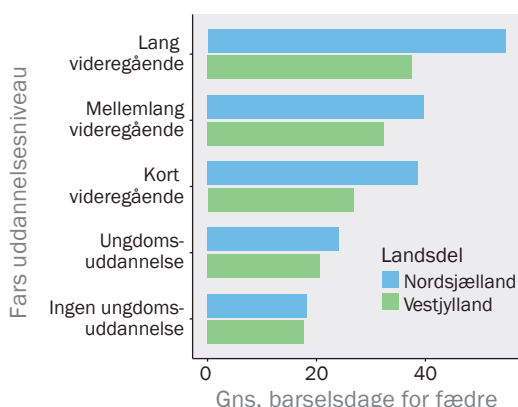
$$y = \beta_0 + \beta_1 x$$

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

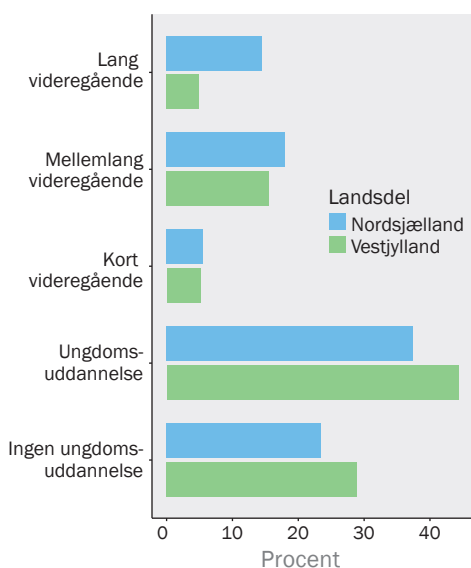
Når man undersøger, hvor godt en model passer til data, kigger man på R^2 -værdien, som er et mål for den samlede afvigelse fra model til data – altså hvor godt et fit, modellen laver. En justeret R^2 -værdi tager højde for modellens kompleksitet og vil derfor være lavere end den ikke-justerede R^2 .



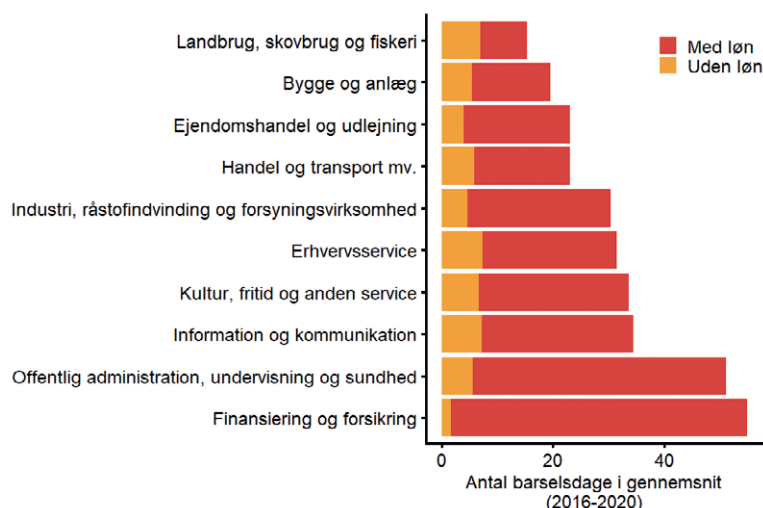
Figur 8: Brancher med den største procentvise forskel i antal ansatte i Nordsjælland og Vestjylland.



Figur 9: Gennemsnitlige antal barselsdage for fædre med forskellige uddannelser i Nordsjælland og Vestjylland.



Figur 10: Andel af personer med forskellige uddannelsesniveauer i Nordsjælland og Vestjylland.



Figur 11: Fordeling af barselsdage på brancher.

i Vestjylland med over 6 % i forhold til Nordsjælland, hvor det er under 1 %. Omvendt er der næsten 5 % indenfor "Finansiering og forsikring" i Nordsjælland, og kun knap 2 % i Vestjylland (figur 8). Det er også en interessant observation, at brancherne, som flest arbejder indenfor i Vestjylland, er mere fysisk krævende jobs, mens det i Nordsjælland i højere grad består af stillesiddende jobs, som kræver en højere uddannelse. Disse tal understøtter branchernes indflydelse på antal barselsdage, da fædre i Vestjylland gennemsnitligt tager markant kortere barsel end fædre i Nordsjælland.

Endvidere er der stor forskel på længden af barsel i de to landsdele for fædre med samme uddannelsesniveau (figur 9). I Nordsjælland tager fædre med en lang videregående uddannelse over 50 dages barsel, hvor de i Vestjylland tager under 40 dage. Brancherne, og jobtyperne inden for disse, kan give en mulig forklaring, da vi må forvente, at der er vidt forskellige barselsvilkår.

Det er svært at konkludere på, om dette udgør hele forklaringen på de forskelle, vi ser, eller om der også er en kulturel forskel på landsdelene, som også har noget at sige, når det kommer til fædres barselstendenser.

Adgangen til barsel er udslagsgivende

Vores undersøgelse af data viser alt i alt, at der er mange nuancer og store forskelle indenfor fædrebarsel på tværs af uddannelse, brancher og landsdele. Ikke overraskende tegner der sig et billede af, at det er adgangen til barsel, som er udslagsgivende. Denne adgang er i høj grad bestemt af branche, social status og uddannelsesniveau, og den er ikke den samme, når vi sammenligner landsdelene i Danmark. Vi ser derfor et meget mere nuanceret billede af fædrebarsel end der præsenteres i medierne, når vi dykker ned i data og kigger på årsager og sammenhænge, og det er tydeligt, at alle fædre ikke kan skæres over én kam, når vi diskuterer barsel. I data viser det sig også, at antal barselsdage uden løn kun er ca. 2 uanset branche, mens det er antallet af dage med løn der varierer (figur 11). Dette understøtter blot hvad vi allerede har set om branchernes effekt samt at adgangen til barsel er det udslagsgivende.

Kilder
Danmarks Statistik:
BARSEL04
BARSEL11
LIGEAB4
HFUDD11

Om aftalen om
øremærket barsel:
kortlink.dk/2fyhr

Politisk aftale indgået:
kortlink.dk/2htfg

Barsel gennem tiden
i Danmark: kortlink.
dk/2htfk

Næsten hver anden
er imod 11 ugers øre-
mærket barsel til far:
Ritzau, 22/9, 2021.