



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Metodenotat til CO2Vision-fremskrivningsmodel

Holm, Jacob Rubæk; Bernth, Christian; Jensen, Karoline Gundersen B; Skipper, Emma

Publication date:
2023

Document Version
Også kaldet Forlagets PDF

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Holm, J. R., Bernth, C., Jensen, K. G. B., & Skipper, E. (2023). *Metodenotat til CO2Vision-fremskrivningsmodel*.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Metodenotat til CO2Vision-fremskrivningsmodel

Jacob Rubæk Holm
Christian Bernth
Karoline Gundersen Bjørnlund Jensen
Emma Skipper

Sidst kompileret: July 25, 2023

Indholdsfortegnelse

1	Indledning	2
2	Variable og data	2
2.1	Endogene variable: udbud, efterspørgsel og relativ løn	2
2.2	Eksogene variable: kvalifikationer i arbejdsstyrken og økonomisk aktivitet	3
2.3	Variabelkonstruktion fra data	3
2.3.1	Datatilgængelighed	5
3	Den formelle fremskrivningsmodel	5
3.1	Overblik over modellen	5
3.1.1	Laglänge	6
3.2	Modellens ligninger	8
3.2.1	Modellen på matriceform	9
3.3	Parameterestimer	10
4	Fremskrivninger	13
4.1	Præsentation af output	15
4.2	Scenarier	16
5	Afrunding	16

1 Indledning

CO2Vision-fremskrivningsmodellen har til formål at fremskrive arbejdsmarkedet i Danmark for de arbejdsfunktioner, der er centrale for CCUS-innovationsøkosystemet. Der er altså tale om fremskrivning af det samlede danske arbejdsmarked på trods af, at CO2Vision-projektet har fokus på CCUS i Nordjylland. Eftersom CCUS-aktiviteter er relativt begrænsede i omfang selv på national plan, og arbejdsfunktionerne indenfor CCUS er typiske funktioner i det samlede arbejdsmarked, må modelleringen nødvendigvis være på nationalt plan.

Modellen vil modellere udbud og efterspørgsel separat, og disse to variable er dermed modellens første to endogene variable. De endogene variable kan potentielt påvirke hinanden. Eksempelvis kan et lavere udbud af arbejdskraft for en given funktion få virksomheder til at undgå at efterspørge arbejdskraft af denne funktion gennem omorganisering af arbejdet eller gennem teknologisk automatisering. Og omvendt kan en lavere efterspørgsel få arbejdstagere til i stedet at udbyde deres arbejdskraft til andre funktioner, hvor de opfatter sandsynligheden for at finde arbejde som værende højere.

Modellen er en vektorautoregressiv- (VAR-) model. Det vil sige, at modellen indeholder flere endogene variable, som indbyrdes påvirker hinanden mod en ligevægt. Foruden udbud og efterspørgsel indeholder modellen relativ løn som endogen variabel. VAR modellen indeholder således modellering af forholdet mellem flere variable over tid. Den indeholder desuden eksogene variable, som ikke fremskrives internt i modellen, men hvor man derimod kan konstruere forskellige scenarier. I vores tilfælde er disse scenarier forskellige hypoteser om udrolning af CCUS i Danmark. De to eksogene variable i vores model er henholdsvis en variabel for relevante kvalifikationer i arbejdsstyrken samt en variabel for niveauet for økonomisk aktivitet i de relevante dele af den danske økonomi. Modellen vil strukturelt set være identisk for hver jobfunktion, og jobfunktionerne vil blive modelleret individuelt. Der er således tale om en partiel ligevægtsanalyse, som ignorerer afledte effekter på og fra den resterende danske økonomi.

2 Variable og data

2.1 Endogene variable: udbud, efterspørgsel og relativ løn

Som nævnt i afsnit 1 kan det forventes, at der er vekselvirkninger mellem udbud og efterspørgsel af arbejdskraft til en given funktion. En tredje endogen variabel vil være den relative løn til funktionen (Ibsen og Stamhus, 2016, kap. 2, 5 og 11). Løndannelse på det danske arbejdsmarked er mere end et spørgsmål om udbud og efterspørgsel, men det kan ikke benægtes, at udbud og efterspørgsel må forventes at have en vis effekt på lønnen. Lønnen påvirker udbud og efterspørgsel, da højere relativ løn for en given funktion alt andet lige vil medføre, at flere arbejdssøgende vil udbyde deres arbejdskraft indenfor netop denne funktion, mens arbejdsgivere vil søge billigere alternativer gennem omorganisering og au-

tomatiseringsteknologier såsom robotter og kunstig intelligens (Nielsen og Holm, 2020; Lorenz og Holm, 2021; Nielsen et al., 2021).

Forholdene mellem de tre endogene variable afspejler til dels trægheden i løndannelse. Hvis efterspørgsel efter en given type arbejdskraft er høj relativt til udbuddet, så bør lønnen i et frit marked stige, således at udbuddet stiger og efterspørgslen falder. Men på grund af træghed i løndannelsen vil denne effekt også observeres som en direkte effekt mellem udbud og efterspørgsel: hvis lønnen er relativt lav, og den er langsom til at tilpasse sig, vil arbejdsgivere opleve forgæves rekruttering. Fra arbejdstagernes synspunkt er dette ensbetydende med gode jobmuligheder, og de vil derfor i større grad udbyde deres arbejdskraft indenfor denne funktion. Hvis lønnen omvendt er relativt høj og ikke tilpasser sig, vil arbejdstagerne ikke kunne finde et arbejde. Fra arbejdsgivernes synspunkt er dette ensbetydende med, at der er lettilgængelig arbejdskraft for denne funktion, og de kan dermed med fordel omorganisere arbejdet, således at de kommer til at efterspørge netop denne funktion (Ibsen og Stamhus, 2016, kap. 6).

2.2 Eksogene variable: kvalifikationer i arbejdsstyrken og økonomisk aktivitet

Foruden de tre endogene variable udbud, efterspørgsel og relativ løn, indeholder modellen to eksogene variable: kvalifikationer i arbejdsstyrken samt økonomisk aktivitet. Kvalifikationer i arbejdsstyrken er en variabel, der påvirker udbuddet af arbejdskraft til en given funktion (Ibsen og Stamhus, 2016, kap. 4). Jo færre personer i den danske arbejdsstyrke, der har de typiske formelle kvalifikationer til en given funktion, jo lavere vil udbuddet af arbejdskraft til denne funktion være. Eksempelvis betyder et faldende antal personer med smedearbejde, at udbuddet af arbejdskraft til funktionen “smedearbejde” er faldende. Den økonomiske aktivitet i virksomhederne påvirker efterspørgslen. Jo højere økonomisk aktivitet i de virksomheder, der udfører smedearbejde, jo højere er efterspørgslen efter personer til at udføre smedearbejde.

2.3 Variabelkonstruktion fra data

Dette afsnit beskriver, hvordan data til modellens fem variable er indsamlet, og hvordan variablene er konstrueret fra data.

Hver funktion har en firecifret kode såsom “7221 Smedearbejde”. Man burde således skrive d_t^{7221} for at henvise til værdien af variabelen d for smedearbejde på tidspunkt t , men formålet her er at udvikle den generelle model til fremskrivning af alle 16 funktioner, og vi undlader derfor at henvise til den specifikke funktion.

Efterspørgsel, d_t . Efterspørgslen er antal nyopslåede stillinger i måned t indenfor den relevante funktion. Månedlige data på antal nyopslåede stillinger er leveret af Styrelsen for Arbejdsmarked og Rekruttering (STAR).

Udbud, s_t . Udbuddet er antal nye ledige i t , der ønsker job indenfor funktionen. Månedlige data på nye lediges jobønsker er leveret af STAR.

Relativ løn, w_t . Relativ løn er forskellen i gennemsnitlig månedsløn over personer med pågældende funktion i t fra det samlede nationale gennemsnit i t . Inden beregningen er lønnen korrigeret for alder, uddannelse, region og køn. I praksis er dette en regression efter mindste kvadraters metode udført for hver måned, hvor den afhængige variabel er månedslønnen og de forklarende variable er funktion, alder, uddannelse, region og køn. For funktion anvendes effektkodning således at effekten for hver enkelt funktion er afvigelsen fra det samlede gennemsnit. Data til denne beregning er registerdata hos Danmark Statistik (DST), og for hver måned anvendes alle fuldtidsansættelsesforhold i Danmark. Dataene er omregnet til 2005 kroner ved hjælp af forbrugerprisindekset. En positiv værdi af w_t betyder dermed, at personer med pågældende funktion har højere månedsløn end andre, og at dette ikke er forklaret ved alder, uddannelse, region eller køn.

Kvalifikationer i arbejdsstyrken, k_t . Kvalifikationer i arbejdsstyrken er den naturlige logaritme til antal personer på det danske arbejdsmarked i måned t med en relevant uddannelse til at varetage funktionen. De relevante uddannelser for en given funktion er alle de uddannelser, som observeres blandt mindst 5% af de personer, der bestred pågældende funktion i 2020. For nogle funktioner findes der med denne definition kun én relevant uddannelse, enten fordi størstedelen af personerne har samme uddannelse, eller fordi personerne er spredt over så mange uddannelser, at kun en enkelt uddannelse observeres hos minimum 5% af personerne. Disse to ekstremer er tilfælde, hvor der er hhv. en meget stærk og meget svag sammenhæng mellem formelle kvalifikationer og arbejdsfunktion. Registerdata hos DST på den samlede danske arbejdsstyrke er anvendt til at beregne k_t for hver måned.

Økonomisk aktivitet, y_t . Økonomisk aktivitet i måned t er gennemsnitsomsætning over de brancher, der har jobs indenfor funktionen. For at bestemme hvilken andel af den økonomiske aktivitet i en branche, der kan tilskrives hver funktion, har vi anvendt en simpel økonomisk model. Modellens udgangspunkt er ligning 1.

$$Y_i = a_i L_i. \quad (1)$$

I ligning 1 er Y_i omsætningen i branche i og L_i er antal beskæftigede i branche i . a_i er dermed gennemsnitlig omsætning per beskæftiget. Forskelle i a_i over branche reflekterer bl.a. forskelle i produktivitet og i kapitalintensitet på tværs af brancher. I hver branche er beskæftigelsen fordelt over et antal funktioner indekseret med j : $L_i = \sum_j L_{ij}$. Den samlede nationale omsætning produceret med funktion j er $Y_j = \sum_i a_i L_{ij}$, og eftersom $a_i = Y_i/L_i$ så er

$$Y_j = \sum_i Y_i \frac{L_{ij}}{L_i}. \quad (2)$$

Den naturlige logaritme til Y_j (ligning 2) fra denne simple model er vores y_t . Ligning 2 siger, at omsætningen tilskrevet funktion j er summen over omsætningen i alle de brancher, der anvender funktion j , vægtet med funktionens vægt i pågældende branches samlede beskæftigelse. Hvis eksempelvis CCUS

medfører en stigning i omsætningen i branche i på 1 million kroner, og 10% af arbejdskraften i branche i har funktion j , så stiger omsætningen for funktion j samlet set med 1 million kroner gange $0,1 = 100.000\text{kr}$.

Til at konstruere variabelen y_t anvendes registerdata hos DST om alle virksomheders omsætning, hvilket vi har aggregeret til brancheniveau. Brancheniveauet beskrives af de første tre cifre i branchekoden.

Data for omsætning i registrene er årlig. Vi har fordelt omsætningen ud over årets 12 måneder ved at antage, at omsætning over året er proportional med beskæftigelse over året. Omsætning er omregnet til 2005-priser ved hjælp af forbrugerprisindekset.

Vi har konstrueret en variabel for faste bruttoinvesteringer på samme måde som variabelen for omsætning, og sammenholdt VAR-modeller med kun omsætning, kun investering eller begge. Økonometriske tests viste, at modellen med kun variabelen for omsætning er overlegen og variabelen for investeringer er derfor udeladt af fremskrivningsmodellen.

2.3.1 Datatilgængelighed

Projektet har haft adgang til registerdata fra DST frem til udgangen af 2020. Data fra STAR, dvs. data på udbud og efterspørgsel, løber derimod fra maj 2010 til januar-februar 2023.

Data for maj 2010 til december 2020 er derfor brugt til at estimere VAR modellen, og fremskrivningerne starter derfor med januar 2021. Vi vil kun præsentere fremskrivninger for udbud og efterspørgsel, og der vil altså være et overlap mellem data for fremskrivning i figurerne. Dette tjener delvist til at validere fremskrivningerne.

3 Den formelle fremskrivningsmodel

3.1 Overblik over modellen

I afsnit 2.1 og 2.2 beskrives modellens sammenhænge mens variablene i modellen er beskrevet i afsnit 2.3. I dette afsnit beskrives den formelle model.

Modellen har tre ligninger; en for hver endogen variabel. Hver endogen variabel er en funktion af fire grupper af effekter:

1. Tidligere værdier af sig selv.
2. Tidligere værdier af de andre endogene variable.
3. Samtidige og tidligere værdier af de eksogene variable.
4. En række deterministiske elementer forklaret senere i dette afsnit.

I afsnit 2.2 blev det argumenteret for, at mængden af kvalificeret arbejdskraft påvirker udbuddet mens den økonomiske aktivitet påvirker efterspørgslen.

Den matematiske model er dog komplet symmetrisk så begge eksogene variable påvirker alle tre endogene variable.

Indledningsvis er det et åbent spørgsmål hvor mange “tidligere værdier”, der skal inkluderes i modellen. Hvis kun én tidligere værdi inkluderes, så vil eksempelvis udbuddet i juni 2012, altså antal nye ledige der ønsker en bestemt jobfunktion i juni måned 2012, afhænge af: 1. udbuddet for samme funktion i maj måned, 2. af efterspørgslen i maj, 3. af den relative løn for funktionen i maj, 4. af den kvalificerede arbejdskraft i juni og maj, og 5. af omsætningen i relevante brancher i juni og maj.

3.1.1 Laglængde

Det er på ingen måde givet at kun én tidligere værdi skal anvendes i modellen. Antal tidligere værdier kaldets også laglængden. For hver af de 16 funktion har vi sammenholdt modeller med forskellig laglængde. I de fleste tilfælde, at der er begrænset forskel på at sætte laglængden lig 1 eller 2. Vi har derfor valgt at sætte laglængden lig 1. Som nævnt tidligere er modellen strukturelt set identisk for alle funktioner. Tabel 1 viser den optimale laglængde for hver arbejdsfunktion. Ved hver funktion er tre informationskriterier anvendt: Akaike, Bayes-Schwarz og Hannan-Quinn (jf. (Enders, 2014, kap. 5) og (Ghysels og Marcellino, 2018, kap. 6)). I de fleste tilfælde er alle tre kriterier enige om den optimale laglængde, men I alle tilfælde er mindst to af de tre enige.

	Arbejdsfunktion	Optimal laglængde
2141	Ingeniørarbejde i industri og produktion	1
2142	Ingeniørarbejde vedrørende bygninger og anlæg	1
2143	Ingeniørarbejde inden for miljø	1
2144	Ingeniørarbejde inden for mekaniske systemer	1
2149	Andet ingeniørarbejde (undtagen elektroteknologi)	1
2152	Ingeniørarbejde inden for elektronik	1
2161	Arbejde med bygningsarkitektur	1
3111	Teknikerarbejde inden for kemisk og fysisk arbejde	1
3112	Teknikerarbejde inden for byggeri- og anlæg	1
3115	Teknikerarbejde inden for det mekaniske område	1
3119	Andet teknikerarbejde inden for fysisk videnskab og ingeniørvidenskab	1
7213	Pladesmedarbejde	2
7221	Smedearbejde	2
7223	Operatør og opstillingsarbejde af metalforarbejdningsmaskiner	1
7233	Mekanikerarbejde, landbrugs- og industrimaskiner	1
8113	Boreanlægsarbejde	1

Tabel 1: Optimal laglængde for hver funktion

Laglængde er testet op til laglængde 4. Antal parametre og tidsseriens længde betyder, at højere laglængde ikke er muligt. Modellen indeholder $15p+48$ parametre at estimere, hvor p er laglængde, jf. opremsningen i punktform i afsnit 3.3. Tidsserien er fra maj 2010 til december 2020 begge inklusiv, altså 128 måneder mens $p = 4$ medfører $15 * 4 + 48 = 108$ parametre i modellen.

Modellen indeholder desuden en deterministisk del bestående af 11 dummier for måneder, som skal tage højde for sæsonudsvingninger, en lineær tidstrend, en konstant og en kontrol for COVID19-pandemien. Sidstnævnte er en dummy, der tager værdien 1 fra februar 2020 til september 2021 begge måneder inklusiv. Februar 2020 er måneden for første nedlukning i Danmark, mens september 2021 er måneden, hvor COVID19 for første gang blev erklæret “ikke længere samfundskritisk”.¹ Modellen er for hver af de 16 funktioner desuden underkastet den sædvanlige række af tests for overholdelse af VAR-metodens økonometriske antagelser. Fokus har været på at få middelrette estimater, mens forhold såsom autokorrelation, der påvirker præcisionen af estimaterne, har fået relativt mindre fokus. Det eneste kritiske problem er manglende stationaritet i den estimerede funktion for efterspørgslen efter den lidt vagt definerede funktion “2149 Andet ingeniørarbejde (undtagen elektroteknologi)”, men i og med at funktionen er så vagt defineret, vil der ikke blive lagt vægt på den i analysens konklusioner.

Middelrette estimater er prioriteret over præcise estimater da fokus i analysen er på fremskrivninger og ikke på de konkrete estimater og deres signifikans. I afsnit 3.3 præsenteres estimaterne for en af de 16 funktioner som et eksempel. I Bernth et al. (2023) undersøges fire af funktionerne i yderligere detaljer, de analyseres for Grangerkausalitet og Impulse-Response-funktioner (IRF) opstilles (Enders, 2014, kap. 5). Formålet er at vurdere effekter af choks til de endogene variable, men dette ligger ud over formålet med CO2Vision-analysen.

3.2 Modellens ligninger

Modellen har tre ligninger; en for hver endogen variabel. Højre side af ligningerne er ens. Dvs. at hver endogen variabel får lov til at afhænge af begge de eksogene variable k_t og y_t , og har ens deterministisk del. Ligning 3 er den første af de tre ligninger, ligningen for efterspørgslen. Efterspørgslen afhænger af 1 lags af sig selv, 1 lags af udbuddet, 1 lags af relativ løn, både samtidig og første lag af variablene for kvalifikationer i arbejdsstyrken og økonomisk aktivitet, samt den deterministiske del bestående af variabelen for COVID19, 11 dummier for måneder ($M_{m,t}$), en lineær tidstrend og en konstant.

$$\begin{aligned}
 d_t = & \beta_{d,d}d_{t-1} + \beta_{d,s}s_{t-1} + \beta_{d,w}w_{t-1} \\
 & + \theta_{d,k0}k_t + \theta_{d,k1}k_{t-1} + \theta_{d,y0}y_t + \theta_{d,y1}y_{t-1} \\
 & + \phi_d I_{COVID19}(t) + \sum_{m=1}^{11} \alpha_{d,m}M_{m,t} + \tau_d t + \alpha_{d,0} + \varepsilon_{d,t}
 \end{aligned} \tag{3}$$

¹Jf. tidslinje fra Statens Seruminstitut:

<https://www.ssi.dk/aktuelt/nyheder/2022/da-covid-19-ramte-verden-og-danmark-se-tidslinjen-her>
Sidst besøgt 19. juni 2023.

hvor $I_{COVID19}(t) = 1$ hvis $t \in [2020.2, 2021.9]$. Bemærk at $M_{m,t}$ er 11 dummier, hvor fx $M_{1,t} = 1$ hvis tidspunkt t henviser til januar et år, mens $m_{10,t} = 1$ hvis t henviser til oktober et år osv. Ligningerne for udbud og for relativ løn ligner meget ligningen for efterspørgslen, ligning 3. Den eneste forskel er parametrene og fejllede.

$$\begin{aligned} s_t &= \beta_{s,d}d_{t-1} + \beta_{s,s}s_{t-1} + \beta_{s,w}w_{t-1} \\ &+ \theta_{s,k0}k_t + \theta_{s,k1}k_{t-1} + \theta_{s,y0}y_t + \theta_{s,y1}y_{t-1} \\ &+ \phi_s I_{COVID19}(t) + \sum_{m=1}^{11} \alpha_{s,m} M_{m,t} + \tau_s t + \alpha_{s,0} + \varepsilon_{s,t} \end{aligned} \quad (4)$$

$$\begin{aligned} w_t &= \beta_{w,d}d_{t-1} + \beta_{w,s}s_{t-1} + \beta_{w,w}w_{t-1} \\ &+ \theta_{w,k0}k_t + \theta_{w,k1}k_{t-1} + \theta_{w,y0}y_t + \theta_{w,y1}y_{t-1} \\ &+ \phi_w I_{COVID19}(t) + \sum_{m=1}^{11} \alpha_{w,m} M_{m,t} + \tau_w t + \alpha_{w,0} + \varepsilon_{w,t} \end{aligned} \quad (5)$$

3.2.1 Modellen på matriceform

Skriver man i stedet alle tre ligninger sammen på matriceform skal man først definere matricerne. Her er vektorer og matricer med fed skrift. Data skrives endvidere med stort. \mathbf{X}_t er de endogene variable og β er parametrene for de endogene variable. Tilsvarende er \mathbf{z}_t de eksogene variable med parametre θ_0 og θ_1 .

$$\begin{aligned} \mathbf{X}_t &= \begin{bmatrix} d_t \\ s_t \\ w_t \end{bmatrix}, \beta = \begin{bmatrix} \beta_{d,d} & \beta_{d,s} & \beta_{d,w} \\ \beta_{s,d} & \beta_{s,s} & \beta_{s,w} \\ \beta_{w,d} & \beta_{w,s} & \beta_{w,w} \end{bmatrix}, \\ \mathbf{z}_t &= \begin{bmatrix} k_t \\ y_t \end{bmatrix}, \theta_0 = \begin{bmatrix} \theta_{d,k0} & \theta_{d,y0} \\ \theta_{s,k0} & \theta_{s,y0} \\ \theta_{w,k0} & \theta_{w,y0} \end{bmatrix}. \end{aligned}$$

Dermed kan modellens tre ligninger (ligning 3, 4 og 5) skrives kort som ligning 6, der er den traditionelle økonometriske opstilling af modellen:

$$\begin{aligned} \mathbf{X}_t &= \beta \mathbf{X}_{t-1} \\ &+ \theta_0 \mathbf{z}_t + \theta_1 \mathbf{z}_{t-1} \\ &+ \phi I_{COVID19}(t) + \mathbf{A} \mathbf{M}_t + \mathbf{T} t + \mathbf{A}_0 + \boldsymbol{\varepsilon}_t \end{aligned} \quad (6)$$

Hvor den deterministisk del er: en vektor af tre konstanter \mathbf{A}_0 , en vektor af 11 dummier for måneder \mathbf{M}_t , en 3x11 matrice \mathbf{A} af tilhørende månedseffekter og en vektor af tre tidseffekter \mathbf{T} . $\boldsymbol{\varepsilon}_t$ er en vektor af tre fejled, som hver især

har middelværdi nul, er homoskedastiske og ikke udviser autokorrelation. De tre fejlede kan dog være indbyrdes korrelerede.²

3.3 Parameterestimer

VAR-modellen i ligning 6 indeholder 63 parametre, som skal estimeres:

- 9 estimater i β
- 6x2 estimater i alt i θ_0 og θ_1
- 3 estimater i ϕ
- 33 estimater i A
- 3 estimater i T
- 3 estimater i A_0

Tabel 2 viser de 63 parameterestimer og tilhørende standardfejl, når modellen anvendes for funktionen “smedearbejde”. Af pladshensyn er modelinformation placeret i tabel 3.

²Brud på disse antagelser om fejlede vil typisk være forbundet med mindre præcise estimater (Ghysels og Marcellino, 2018, kap. 6), men som argumenteret for i afsnit 3.1.1 er dette sekundært i den forhåndenværende analyse.

	<i>Ligning (endogen variabel):</i>		
	1 (d_t)	2 (s_t)	3 (w_t)
Efterspørgsel, d ved $t - 1$	0,287*** (0,094)	-0,112 (0,280)	2,027 (4,141)
Udbud, s ved $t - 1$	0,011 (0,021)	0,713*** (0,062)	0,493 (0,915)
Relativ løn, w ved $t - 1$	0,001 (0,002)	-0,005 (0,007)	-0,016 (0,101)
Kvalifikationer, k ved t	81,232 (432,277)	806,522 (1.288,215)	16.805,470 (19.027,660)
Aktivitet, y ved t	40,440 (34,409)	-118,329 (102,540)	-556,907 (1.514,568)
Kvalifikationer, k ved $t - 1$	-242,355 (436,639)	-303,251 (1.301,214)	-31.361,480 (19.219,660)
Aktivitet, y ved $t - 1$	-46,087 (35,609)	104,967 (106,116)	-633,862 (1,567,392)
COVID19	-17,300 (10,805)	119,839*** (32,198)	-882,027* (475,587)
Måned 1	-9,890 (6,182)	-29,688 (18,422)	1,515,232*** (272,103)
Måned 2	-6,782 (8,515)	-59,136** (25,376)	-502,894 (374,818)
Måned 3	-0,243 (6,898)	1,751 (20,557)	-564,171* (303,645)
Måned 4	-7,879 (6,693)	1,118 (19,945)	829,143*** (294,594)
Måned 5	-7,903 (6,852)	19,354 (20,419)	-12,870 (301,593)
Måned 6	-12,328* (6,326)	31,047 (18,852)	1,486,178*** (278,448)
Måned 7	-16,892** (7,619)	50,863** (22,705)	1,468,739*** (335,368)
Måned 8	-11,508 (8,765)	66,445** (26,121)	-82,763 (385,817)
Måned 9	-10,339 (6,276)	23,806 (18,702)	203,514 (276,242)
Måned 10	-16,435*** (6,126)	22,786 (18,255)	292,905 (269,643)
Måned 11	-2,862 (6,684)	6,686 (19,920)	194,864 (294,226)
Konstant	1.957,901 (2.652,202)	-5.398,837 (7.903,740)	192.984,500 (116.742,700)
Tidstrend	0,205 (0,229)	1,045 (0,682)	-22,073** (10,077)

Note:

*p<0,1; **p<0,05; ***p<0,01

Tabel 2: Parameterestimer for funktion 7221 Smedearbejde

	<i>Ligning (endogen variabel):</i>		
	1 (d_t)	2 (s_t)	3 (w_t)
Observationer	126	126	126
R ²	0,629	0,956	0,733
Adj. R ²	0,558	0,948	0,682
F-test (df = 20; 105)	8,889***	114,440***	14,424***

Tabel 3: Modelinformation for funktion 7221 Smedearbejde

Som man kan se afhænger efterspørgsel og udbud begge positivt af deres eget lag. Det vil sige, at antallet af nye stillinger og nye ledige for smedearbejde er højere, hvis det også var relativt højt i forrige måned. Derudover er der ingen statistisk signifikante effekter blandt de endogene og eksogene variable. Som diskuteret i afsnit 3.1.1 er fokus her på modellens fremskrivninger og ikke de enkelte parameterestimater, og man skal være forsigtig med at fortolke for meget på det enkelte estimat, da kryds effekter mellem ligninger og modellens dynamiske struktur betyder, at der er mange veje fra en given variabel til den anden.³

I modellens deterministiske del kan man se, at der i måneder påvirket af Covid19 var højere udbud (flere søgte smedearbejde) og lavere relativ løn. Der ses desuden nogle sæsonudsvingning i alle tre endogene variable. Nederst ses det, at der er en statistisk signifikant negativ tidstrend for relativ løn. Den negative trend for relativ løn kan potentielt afspejle, at en voksende andel af de personer, der udfører smedearbejde, ikke har de formelle kvalifikationer, der udløser højere løn. Det kan med andre ord i sig selv være et tegn på et stramt arbejdsmarked, hvor smedearbejde udføres af personer uden de rette formelle kompetencer.⁴

4 Fremskrivninger

Modellen fremskriver de endogene variable ved hjælp af parameterestimaterne - eksempelvis anvendes estimaterne i tabel 2 til at fremskrive efterspørgsel, udbud og relativ løn for smedearbejde. For at kunne fremskrive kræver modellen dog også værdier for de eksogene variable. Det vil sige antagelser om hvordan k_t , kvalifikationer i arbejdsstyrken, og y_t , økonomisk aktivitet, udvikler sig i fremtiden. Dette er vores mulighed for at antage forskellige scenarier, jf. afsnit 4.2.

Til den indledende fremskrivning antager vi, at udviklingen i de to variable fortsætter som hidtil. Denne fremskrivning er reference for de senere scenarier. k_t og y_t fremskrives begge som en lineær tidstrend med sæsonudsvingninger og korrektion for Covid19. Til brug for denne fremskrivning er de tilgængelige data for kvalifikationer i arbejdsstyrken og for økonomisk aktivitet regresseret på variable for Covid19, måneder og tid. Resultatet ses i tabel 4.

³Eksempelvis vil en stigning i antal personer i arbejdsstyrken med de rette kvalifikationer på 1% på tidspunkt t øge udbuddet med 806,5 ledige, men det vil sænke udbuddet igen perioden efter med $-303,3$ ledige. Bemærk at en stigning på 1% oftest er en relativt stor stigning på tusindevis af personer.

På samme vis vil stigningen i kvalifikationer først øge og sidenhen sænke den relative løn, hvilket også påvirker udbuddet gennem den laggede effekt af den relative løn på udbud. Netop denne komplicerede struktur er et godt argument for eksempelvis at analysere Grangerkausalitet og IRF, som i Bernth et al. (2023), hvis fokus er på variabelnes effekter på hinanden.

⁴Den relative løn er korrigeret for uddannelse, jf. afsnit 2.3, men dette er uddannelse på et relativt grovkornet niveau med fokus alene på uddannelsens længde.

<i>Afhængige variabel:</i>		
	Kvalifikationer, k_t	Aktivitet, y_t
COVID19	0,001 (0,002)	-0,431*** (0,042)
Måned 2	0,001 (0,002)	0,035 (0,049)
Måned 3	0,002 (0,002)	0,069 (0,049)
Måned 4	0,003 (0,002)	0,016 (0,049)
Måned 5	0,004* (0,002)	0,042 (0,049)
Måned 6	0,006** (0,002)	0,054 (0,049)
Måned 7	0,008*** (0,002)	-0,049 (0,049)
Måned 8	0,009*** (0,002)	0,011 (0,049)
Måned 9	0,010*** (0,002)	0,057 (0,049)
Måned 10	0,008*** (0,002)	0,058 (0,049)
Måned 11	0,009*** (0,002)	0,052 (0,049)
Måned 12	0,010*** (0,002)	-0,011 (0,049)
Konstant	11,461*** (0,002)	20,755*** (0,039)
Tidstrend	-0,001*** (0,00001)	0,005*** (0,0003)
Observationer	141	132
R ²	0,974	0,713
Adj. R ²	0,972	0,682
F-test	368,987*** (df = 13; 127)	22,578*** (df = 13; 118)
<i>Note:</i>	*p<0,1; **p<0,05; ***p<0,01	

Tabel 4: Parameterestimer for eksogene variable for smedearbejde

For kvalifikationer kan man se i tabel 4, at der ofte er flere kvalifikationer tilgængelige sidst på året. Dette kan skyldes en kombination af sæsonarbejdsløshed, hvor personer forlader arbejdsstyrken om vinteren, og tilgang af nyuddannede i efteråret. Man kan desuden se en statistisk signifikant tidstrend på -0.001 . Dette betyder, at antallet af personer med kvalifikationer til at tage smedearbejde faldt med i gennemsnit $0,1\%$ om måneden i årene 2010-2020,⁵ og denne trend vil derfor blive en del af reference-fremskrivningen. For økonomisk aktivitet kan man se, at måneder påvirket af Covid19 havde $43,1\%$ mindre økonomisk aktivitet⁶ Der er ingen sæsonudsvingninger for økonomisk aktivitet relateret til smedearbejde, men der er en positiv tidstrend som viser, at den økonomisk aktivitet stiger med $0,5\%$ om måneden.

4.1 Præsentation af output

De rå data for efterspørgsel og udbud er månedlige. Eftersom der kan være en del tilfældigheder involveret i hvilken måned et specifikt job slås op (efterspørgsel), og hvilken måned en person registrer sig som arbejdssøgende, så er der store udsvingninger i disse to tidsserier. Dertil skal selvfølgelig lægges sæsonudsvingninger. Disse rå data vil derfor ikke blive præsenteret som en del af analysen. I stedet præsenteres antal nye ledige per nye stilling over de forudgående 12 måneder. Denne brøk er relativt nem at fortolke, når fokus er på risikoen for mangel på arbejdskraft, og ved at summere over 12 måneder udjævnes de mange tilfældige udsving.

Som eksempel bliver de rå tal dog præsenteret her for funktionen "smedearbejde" som et kig ned i motorrummet på modellen. Figur 1 side 18 viser efterspørgslen, dvs. nye stillinger per måned, og udbuddet, dvs. nye ledige per måned. Den vandrette akse i figurerne er en abstrakt tidsakse for 1 er maj 2010 mens 121 er 10 år senere, maj 2020 osv.

Den første del af graferne i figur 1 er datapunkter. Man kan se efterspørgslen vokse fra ganske få nye jobs om måneden over tid. Udbuddet vokser også og har tydelige sæsonudsvingninger. Den store pukkel midt i grafen for udbud er sammenfaldende med Covid19. Den sidste del af grafen er skraveret - heraf navnet fanchart. Dette er fremskrivningen, og skraveringen viser den statistiske usikkerhed. Både udbud og efterspørgsel stiger, men det ser ud til, at efterspørgslen stiger kraftigere end udbuddet, dog på et meget lavere niveau.

Figur 2 side 19 er det format, i hvilket fremskrivningen vil blive præsenteret i analysen. Da der er tale om sum over de forudgående 12 måneder, er de første datapunkter i 2011, hvor der er næsten 200 ledige, som søger smedearbejde, for hver ny stilling indeholdende smedearbejde. Kigger man "ned i motorrummet" i figur 1 kan man se, at der ikke var synderligt mange ledige, som søgte smedearbejde på dette tidspunkt, men derimod meget få nye stillinger med smedearbejde. Dermed blev antal ledige per stilling meget højt.

Covid19-puklen kan også tydeligt ses i figur 2. Efter Covid19 viser fremskrivningen et lavt og svagt faldende antal ledige per ny stilling frem mod 2030,

⁵ k_t er den naturlige logaritme til antallet af personer med kvalifikationer for smedearbejde.

⁶ y_t er den naturlige logaritme til den gennemsnitlige omsætning i relevante brancher.

hvilket er konsistent med den kraftigere stigning for efterspørgsel end for udbud i figur 1.

4.2 Scenarier

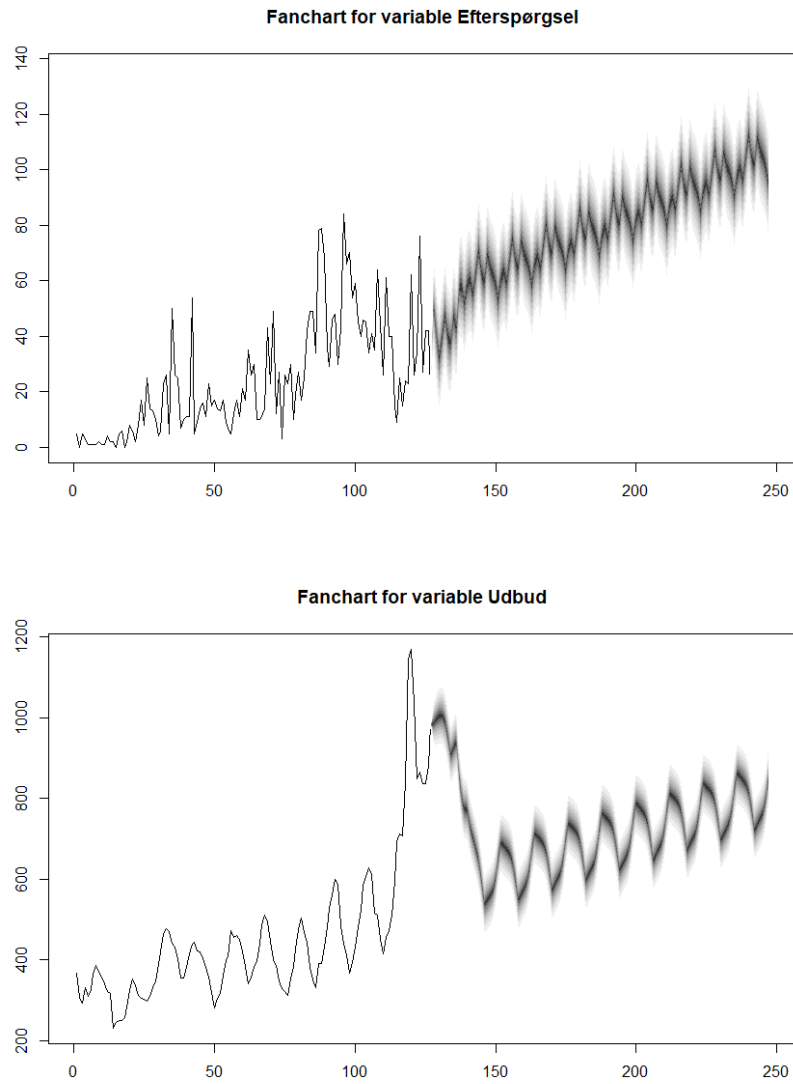
I analysen vil tre forskellige scenarier blive sammenholdt med referencefremskrivningen. Der vil være to scenarier for stigning i den økonomiske aktivitet forbundet med fangst, anvendelse og lagring af kuldioxid. Gennem CO2Vision-projektet er de brancher identificeret, der bidrager til CCUS, og det kan derfor kvantificeres, hvorledes variabelen y_t vil udvikle sig for de forskellige arbejdsfunktioner i fremtiden, når CCUS rulles ud i Danmark. Forskellen på de to scenarier vil være omfanget af udrulningen. I de tredje scenario antages i stedet en negativ demografisk udvikling, hvorfor væksten i k_t begrænses eller faldet i k_t intensiveres, alt efter hvad tilfældet er for den konkrete arbejdsfunktion.

Alle tre scenarier ventes at føre til strammere arbejdsmarkeder, hvor antallet af ledige per nye job er lavere end i referencetilfældet. Dog med den begrænsning i baghovedet, at CCUS har et relativt begrænset omfang sammenholdt med det samlede danske arbejdsmarked. Scenarierne vil derfor blive præsenteret relativt til referencefremskrivningen. Det vil sige, vi ikke indsætter en alternativ fremskrivning i figur 2 for hvert scenario, men i stedet præsentere den procentvise afvigelse fra referencefremskrivningen.

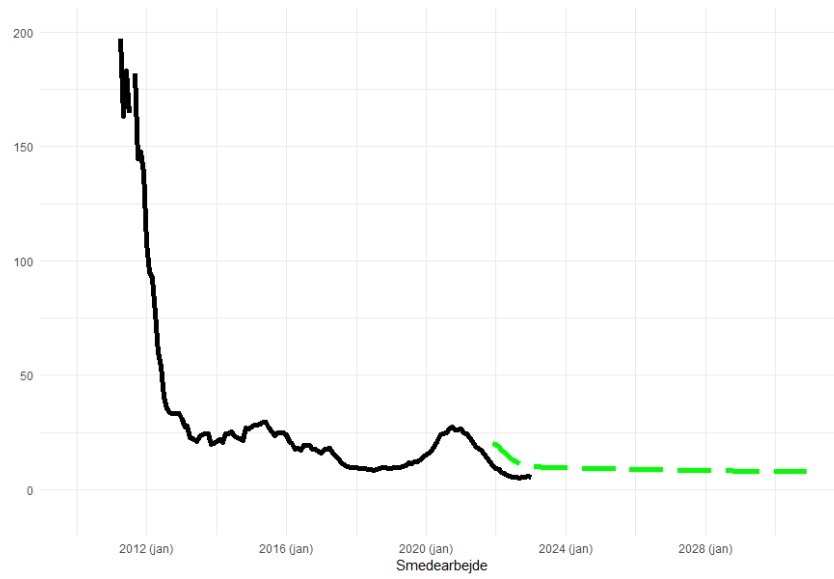
5 Afrunding

CO2Vision-fremskrivningsmodellen udviklet i dette metodenotat har fordele og ulemper. Den faste struktur på tværs af de 16 arbejdsfunktioner er en stor fordel, når modellen skal anvendes, omend det havde været optimalt at lave unikke modeller for hver funktion fra et fagligt synspunkt. Som så mange andre VAR-modeller lider denne model under få frihedsgrader, fordi mange parametre skal estimeres. Endvidere er især de eksogene variable ofte statistisk insignifikante. I et dedikeret projekt udelukkende fokuseret på at udvikle modeller til fremskrivning af arbejdsmarkedet kunne modellens eksogene del være udviklet yderligere. Desuden kunne modellens fremskrivningsevne være analyseret mere systematisk. Til det forhåndenværende formål vurderes modellen dog at være tilstrækkelig, ikke mindst sammenlignet med andre modeller, der er anvendt til at fremskrive arbejdsmarkedet i forbindelse med den grønne omstilling, jf. diskussionerne i Holm (2022). Konkret kan man pege på, at CO2Vision-modellen er en statistisk model, der er baseret på historiske lovmæssigheder fremfor enkelte eksperter overslag for jobskabelsen ved konkrete investeringsprojekter. Modellen her skiller sig også ud ved at modellere udbud, efterspørgsel og relativ løn samtidigt, således at der tages højde for krydseffekter mellem de tre variable. Og slutteligt, hvor andre fremskrivninger fokuserer på antal personer i arbejdsstyrken med en specifik uddannelse, så er fokus i CO2Vision-modellen på arbejdsfunktion eller stilling. Det betyder, at denne model er et skridt tættere på det egentlige udbud af, og efterspørgsel efter,

kompetencer, da den tager højde for at arbejdsgivere kan substituere mellem personer med forskellige uddannelser, når en given funktion skal udføres.



Figur 1: Rå fremskrivning for smedearbejde



Figur 2: Endelig afbildning af fremskrivning

Referencer

- Bernth, C., Skipper, E. og Karoline, G. B. J. (2023). Chok til arbejdsmarkedsbalancen for smede. Bachelorprojekt, Aalborg University Business School.
- Enders, W. (2014). *Applied econometric time series*. John Wiley & Sons, 4th udgave.
- Ghysels, E. og Marcellino, M. (2018). *Applied economic forecasting using time series methods*. Oxford University Press.
- Holm, J. R. (2022). Udbud af og efterspørgsel efter CCUS-relateret arbejdskraft og kompetencer: Relevante eksisterende analyser. Rapport, CO2Vision. <https://www.co2vision.dk/analyserapport-fra-aau-vedr-kompetencebehov/>.
- Ibsen, F. og Stamhus, J. (2016). *Arbejdsmarkedsøkonomi*. Djøf Forlag.
- Lorenz, E. og Holm, J. R. (2021). Work organisation, innovation and the quality of working life in Denmark. I Christensen, J. L., Gregersen, B., Holm, J. R. og Lorenz, E., redaktører, *Globalisation, New and Emerging Technologies, and Sustainable Development: The Danish Innovation System in Transition*, side 169-188. Routledge.
- Nielsen, P. og Holm, J. R. (2020). Arbejdsmarkedspolitik i automatiseringens tidsalder. I Klindt, M. P., Rasmussen, S. og Jørgensen, H., redaktører, *Aktiv arbejdsmarkedspolitik: Etablering, udvikling og fremtid*, side 325-356. Djøf Forlag.
- Nielsen, P., Holm, J. R. og Lorenz, E. (2021). Work policy and automation in the fourth industrial revolution. I Christensen, J. L., Gregersen, B., Holm, J. R. og Lorenz, E., redaktører, *Globalisation, new and emerging technologies, and sustainable development: The Danish Innovation System in Transition*, side 189-207. Routledge.