

## Baggrundsnotat til kapitel III i *Produktivitet, 2023*

**Dette baggrundsnotat giver yderligere detaljer bag analyserne om automatisering i fremstillingsvirksomhederne i afsnit 4-5 i kapitel III i *Produktivitet, 2023*.**

I afsnit III.4 i *Produktivitet, 2023* opgøres udviklingen i antallet af fremstillingsvirksomheder, som investerer i automatiserende teknologier. I analysen i afsnit III.5 estimeres effekten af at påbegynde investeringer i automatisering på forskellige variable i virksomheder.

De centrale elementer i metode, datagrundlaget og resultater er beskrevet i rapporten. Dette notat indeholder nogle uddybende detaljer og fordeler sig på fire afsnit:

1. Sampleudvælgelse
2. Beskrivelse af metoden
3. Definition af produktionsarbejdere
4. Yderligere resultater

### 1 Sampleudvælgelse

I afsnit III.4 og III.5 i *Produktivitet, 2023* anvendes virksomhedsdata fra Danmarks Statistik. I analyserne kombineres data fra regnskabsstatistikken (FIRE) med data om import fra udenrigshandelsregistret (UHDI). Analysen har fokus på fremstillingsvirksomheder, der ikke er i brancher med høj grad af videresalg af automatiserende maskiner. Desuden kræves det, at virksomhederne har indrapporteret skatteoplysninger eller årsregnskaber. Endeligt fokuseres der på virksomheder med mindst et årsværk. For analysen i afsnit 5 stilles en række yderligere krav til stikprøven, der beskrives nedenfor. Tabel 1 giver et overblik over stikprøvesammensætningen efter hvert udvælgelseskriterie.

<b>Tabel 1: sampleudvælgelse</b>					
	<b>Vsh. år</b>	<b>Unikke Vsh.</b>	<b>Automatiserende</b>	<b>BESK (2010)</b>	<b>OMS (2010)</b>
<b>Alle</b>	3712470	578372	16806	1.193	2.737.271
<b>Fremstillingssektoren</b>	298213	36738	4650	271	637.471
<b>Ikke i brancher med videresalg</b>	264393	32795	4101	245	594.989
<b>Skatteoplysninger</b>	166997	22636	3852	227	562.948
<b>BESK&gt;1</b>	131027	17331	3635	225	559.168
<b>Balanceret panel</b>	34570	1102	167		
Anm.: Vsh. år er antallet af (virksomhed-år) observationer. Unikke virksomheder er antallet af virksomheder, der observeres mindst en gang. Automatiserende er antallet af unikke automatiserende virksomheder. BESK er den samlede beskæftigelse fra regnskabsstatistikken i 2010 målt i 1000 årsværk. OMS er den samlede omsætning i regnskabsstatistikken i 2010 målt i mio. kr.					

I fremstillingssektoren var virksomhedernes samlede omsætning ifølge Statistikbanken på 627.508 mio. kr., mens den samlede beskæftigelse var på 270.621. Dermed dækker samplet, inden der fokuseres på et balanceret panel, 83 pct. af beskæftigelsen og 89 pct. af omsætningen i 2010 af omsætningen i danske fremstillingsvirksomheder i 2010. Dækningen er på samme niveau de øvrige år i perioden 2000-18.

### 1.1 Balanceret panel

I analysen i afsnit III.5 indgår et balanceret panel af virksomheder, der automatiserer på et tidspunkt i perioden 2005-18. Disse sammenlignes med virksomheder, der aldrig automatiserer. Udgangspunktet for analysen er virksomhederne med gyldige regnskabsoplysninger fra tabel 2. Men der stilles en række yderligere krav.

1. For virksomheder, der faktisk automatiserer, kræves det, at de observeres i mindst tre år inden og tre år efter året med den første investering, altså i alt syv sammenhængende år. For virksomheder, der aldrig automatiserer, kræves det, at disse observeres i minimum syv sammenhængende år, da de dermed kan parres med en virksomhed, som automatiserer i midten af syvårsperioden.
  - a. Bemærk at kravet om tilgængeligt data indebærer, at det er muligt at matche virksomheden med IDAN-registret i alle årene, for at få oplysninger om stillingskategorier for de ansatte i virksomheden. Dette er nødvendigt for at beregne lønandelen for produktionsarbejdere.
2. For virksomheder, der automatiserer, kræves det, at de har mindst 10 årsværk i året inden den første investering. For virksomheder, der aldrig automatiserer, kræves det, at de har mindst en observation med 10 årsværk, der har tre års tilgængeligt data både før og efter. Hvis en virksomhed har flere observationer, der lever op til dette, kan den indgå flere gange i analysen.

Disse kriterier giver et balanceret panel med 167 automatiserende virksomheder og 935 potentielle kontrolvirksomheder, der i alt har 3.338 sammenhængende forløb af syv års gyldige observationer.

## 2 Metode

I analysen antages det, at det er muligt at konstruere en relevant kontrolgruppe for de automatiserende virksomheder ved at kombinere alle de virksomheder, der aldrig automatiserer. Antagelsen er, at når der er kontrolleret tilstrækkeligt for observerbare forskelle mellem de automatiserende og ikke-automatiserende virksomheder, så ville de automatiserende virksomheder have udviklet sig som kontrolgruppen, hvis de aldrig havde automatiseret.

I analysen parres hver automatiserende virksomhed med kontrolvirksomheder, som aldrig automatiserer, men som observeres i årene før og efter året, hvor den pågældende virksomhed automatiserer. Tag to virksomheder, der observeres i perioden 2010 til 2018. Den ene af dem automatiserer i 2015. Det giver et naturligt referencetidspunkt. Den anden automatiserer aldrig og har dermed som udgangspunkt intet naturligt referencetidspunkt. Løsningen er at sammenligne virksomhederne, der automatiserer i 2015, med alle virksomheder, der observeres i både 2014 og 2015, men som aldrig automatiserer. Det samme gøres i de øvrige år i perioden 2008-15. Dermed indgår de automatiserende virksomheder kun en gang i samplet, men de ikke-automatiserende virksomheder kan indgå flere gange. I praksis kræves det, at virksomhederne observeres i tre år før og tre år efter automatiseringsbegivenheden, så det er muligt at beregne effekter på længere sigt i en balanceret stikprøve. For en virksomhed, der aldrig automatiserer, betyder det, at den for at være med i kontrolgruppen skal observeres i minimum syv sammenhængende år. Hvis den derimod observeres i otte sammenhængende år, kan den anvendes som kontrol to gange, osv. Tag eksempelvis en virksomhed, der aldrig automatiserer, men som observeres i perioden 2010-18. Den vil indgå i kontrolgruppen tre gange med tre referencetidspunkter: 2013, 2014 og 2015. Derfor laves tre kopier af virksomheden med forskellige referencetidspunkter. I analysen antages det, at effekten af at automatisere i gennemsnit er den samme på tværs af brancher og over tid. Derfor samles alle virksomhederne i en beregning, hvor effekterne beregnes relativt til referenceåret, f.eks. 2015.

Analysen anvender en type matching kaldet *entropy balancing* til at kontrollere for observerbare forskelle, jf. Hainmueller (2012). Metoden går ud på at genvægte alle de virksomheder, der aldrig automatiserer, så gruppen ligner virksomhederne, der automatiser, på en række momenter, f.eks. gennemsnittet af logaritmen til omsætning.

Metoden forsøger at udvælge vægtene, så alle de valgte momenterne balancerer i de to grupper. Metoden forsøger dog samtidigt at sprede vægtene ud på så mange kontrolobservationer som muligt, da det betyder, at mest muligt information anvendes, hvilket giver mindre standardfejl i analysen.

Metoden består i at minimere en funktion, der straffer for at samle vægtene ved en enkelt observation under betingelse af, at hver de vægtede momenter i kontrolgruppen svarer nogenlunde til de tilsvarende momenter i gruppen af automatiserende virksomheder.

Hvis  $m_j(X)$  er et givet moment i gruppen af automatiserende virksomheder, der er beregnet som  $\sum_{i \in \text{auto}} a_{ij}(X) / N$ , hvor  $N$  er antallet af automatiserende virksomheder, så er  $m_j^k(X)$  det tilsvarende vægtede moment i kontrolgruppen beregnet som  $\sum_{i \in \text{kontrol}} w_i a_{ij}(X) / N$ , hvor  $w_i$  er vægten.

Hvis momentet eksempelvis er gennemsnittet af omsætning, bliver  $a_{ij}(X) = oms_i$ . Hvis der derimod er tale om variansen af omsætning, bliver  $a_{ij}(X) = (oms_i - \overline{oms})^2$ , hvor  $\overline{oms}$  er (det vægtede) gennemsnit af omsætning. Det giver anledning til minimeringsproblemet:

$$\min_{\{w_i\}} w_i \log\left(\frac{w_i}{q_i}\right) \text{ s. t. } m_j^k(X) = m_j(X) \text{ for alle momenter } j$$

Der stilles desuden krav til, at alle vægte er positive, og at de summer til antallet automatiserende virksomheder ( $N$ ).  $q_i$  er en på forhånd valgt referencevægt, der sættes til 1. Kontrolvirksomheder, der indgår som kopier med flere referencetidspunkter, vægtes dog ned, så kopierne samlet set indgår med vægt 1. Dermed straffer minimeringsproblemet løsninger, der lægger meget vægt på forskellige kopier af den samme kontrolvirksomhed, hvilket minimerer standardfejl, da der i analysen clustres på virksomhedsniveau.

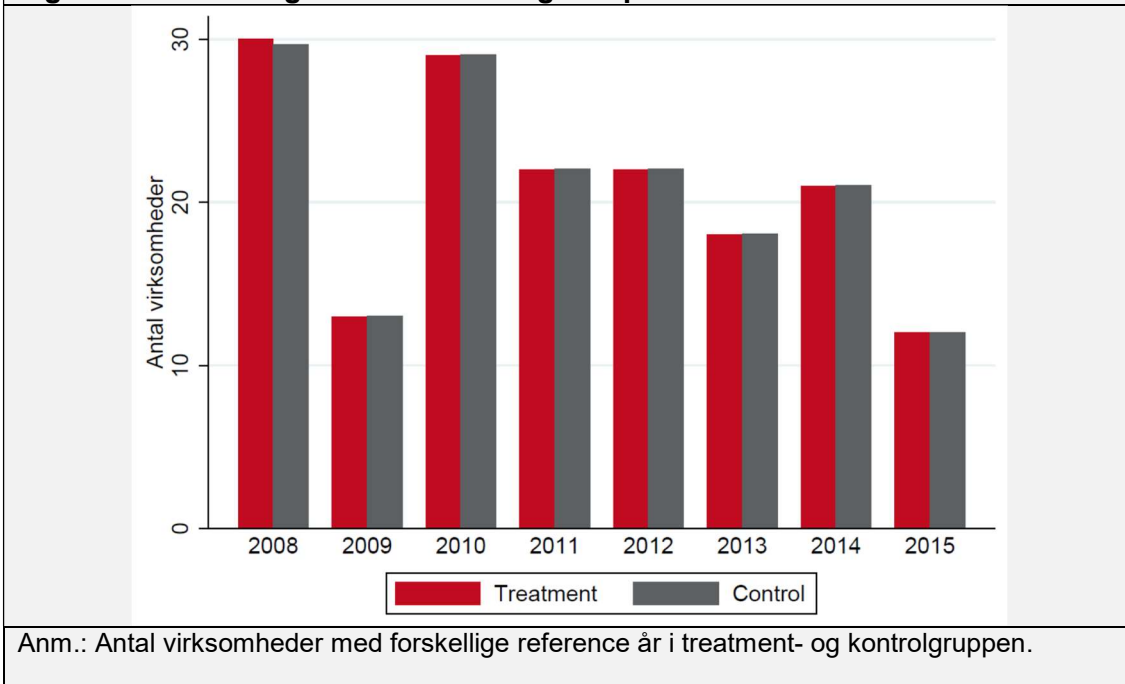
I praksis kræves kun, at bibetingelserne opfyldes approksimativt. I estimationen angives således et konvergenzkriterium, som bestemmer, hvor store afvigelserne mellem momenterne i kontrol- og treatmentgruppe kan være. Estimationen udføres ved brug af *ebalance* i Stata og det dertilhørende standardniveau for konvergenzkriteriet.

Der balanceres på gennemsnit og standardafvigelse af logaritmen til omsætning, logaritmen til værditilvækst, logaritmen til beskæftigelse, logaritmen til værditilvækst pr. beskæftiget, logaritmen til lønkvoten og produktionsarbejdernes lønandel (der kan være 0). Desuden balanceres der, så andelen af virksomheder i treatment- og kontrolgrupperne, der er i en given branche et givent år, er ens. Tabel 2, samt figur 1 og 2 viser, at momenterne i kontrolgruppen og gruppen af automatiserende virksomheder er ens i referenceåret, når de beregnede kontrolvægte anvendes. I praksis balanceres der på kombinationen af branche og år, men resultaterne vises i figur 1 og 2 på et mere aggregeret niveau af diskretionshensyn.

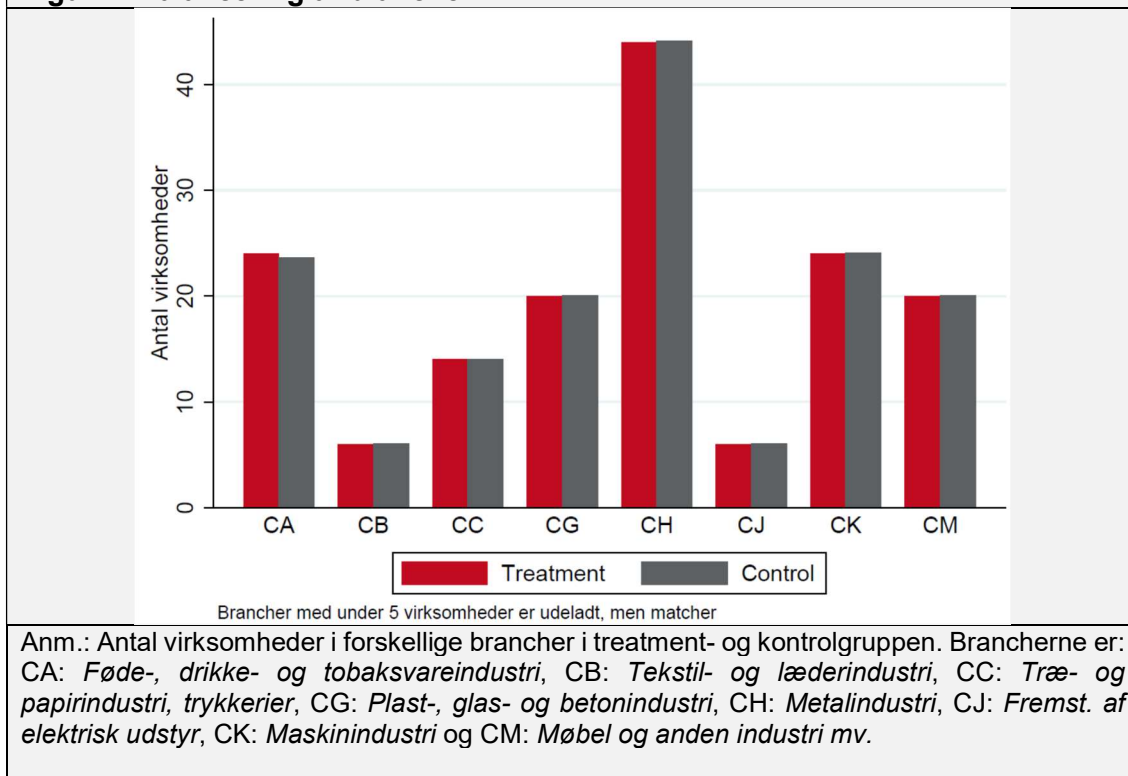
**Tabel 2: Balancering af samples**

	Gennemsnit		Standardafvigelse	
	Auto	Kontrol	Auto	Kontrol
Log. Omsætning	10,95	10,95	1,09	1,09
Log. Værditilvækst	9,96	9,96	0,9	0,9
Log. Beskæftigelse	3,62	3,62	0,83	0,83
Log. Værditilvækst pr. beskæftiget	6,34	6,33	0,36	0,36
Log. Lønkvote	-0,37	-0,37	0,3	0,3
Produktionsarbejdernes lønandel	0,52	0,52	0,22	0,22

Anm.: Gennemsnit og standardafvigelse i treatment- og kontrolgruppen efter balancering.

**Figur 1: Balancering af automatiseringstidspunkt**

**Figur 2: Balancering af brancher**



Med baggrund i de beregnede kontrolvægte ( $w_i$ ), beregnes effekten af automatisering på logaritmen til en given variabel ( $y$ )  $p$  år efter automatisering som

$$\text{Effekt i år } p = \sum_{i \in \text{auto}} \frac{y_{i,p} - y_{i,-1}}{N} - \sum_{i \in \text{kontrol}} w_i \frac{y_{i,p} - y_{i,-1}}{N}$$

Den umiddelbare effekt i året for automatisering findes ved at sætte  $p = 0$ , men det er også muligt at beregne effekten i perioder, der ligger før kontroltidspunktet ( $p = -1$ ), hvilket gør det muligt vurdere, om der er forskelle i udviklingen mellem grupperne i årene før automatisering.

### 3 Definition af produktionsarbejdere

I kapitlet defineres produktionsarbejdere i fremstillingssektoren ud fra DISCO-koder fra 88-nomenklaturen. DISCO-koderne blev opdateret i 2009, hvilket er undervejs i sampleperioden og der er ingen naturlig overgang fra DISCO-08 til DISCO-88. Derfor anvendes en statistisk overgangstabel, som er blevet brugt på danske data før i Humlum (2022).<sup>1</sup> I praksis defineres produktionsarbejdere fra koderne:

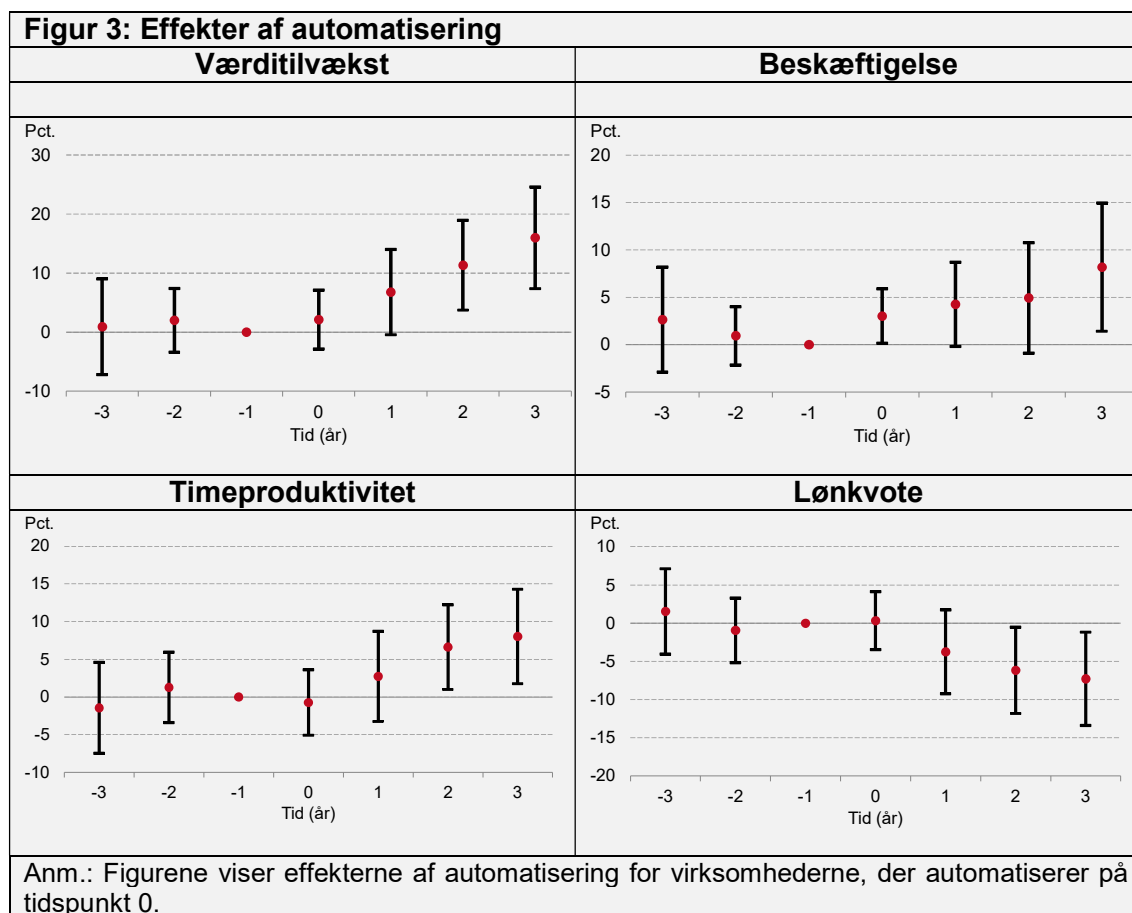
- 7XXX (Håndværkspræget arbejde)
- 932X (Manuelt arbejde inden for fremstillingsvirksomhed)

<sup>1</sup> Overgangstabellen kan hentes på hjemmesiden <https://www.andershumlum.com/codes>

- 81XX (Arbejde med stationære procesanlæg)
- 82XX (Betjening af industrimaskiner)

## 4 Yderligere resultater

I dette afsnit vises resultaterne fra analysen, herunder effekten på værditilvæksten, som er udeladt fra kapitlet. Effekten på værditilvæksten er positiv og signifikant. Tre år efter automatisering er værditilvæksten steget med ca. 15 pct.



Effekterne i procent kan omregnes til kronebeløb for den gennemsnitlige virksomhed:

**Tabel 3: Effekter omregnet til niveau**

	Gennemsnit	Efter 0 år	Efter 1 år	Efter 2 år	Efter 3 år
Værditilvækst pr. beskæftiget	600,1	-4,3	16,4	39,8	48,2
Beskæftigelse (årsværk)	56,2	1,7	2,4	2,8	4,6
Lønkvote (pct.point)	72,7	0,2	-2,7	-4,5	-5,3
Værditilvækst	32.475,7	695,7	2.206,7	3.684,1	5.188,9

Anm.: Tabellen viser effekterne af automatisering, når de omregnes til niveaueffekter for den gennemsnitlige virksomhed.

## 5 Litteraturliste

Acemoglu, D. og P. Restrepo (2022): Demographics and Automation, *The Review of Economic Studies*, 89(1), January 2022, s. 1–44

De Økonomiske Råds Formandskab (2023): Produktivitet, 2023

Hainmueller, J. (2012): Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies, *Political Analysis*, 20(1), s. 25-46.

Humlum, A. (2022). Robot Adoption and Labor Market Dynamics. The ROCKWOOL Foundation Research Unit, Study Paper No. 175.