

The Data Warehouse for Business Statistics

New and improved access to anonymised
microdata for research and analysis

Birgitte Anker
Statistics Denmark
30.11.2023

Access to microdata is a key priority



40 years' delivery on
drastically increased
demand for microdata for
research and analysis



The supply of microdata
is now extended
with access to
more business data and
better metadata

STRATEGIC OBJECTIVES

We provide new and improved statistics for the development of Danish society

DATA

We offer the most secure and easy access to data on Danish society

ORGANISATION

We are a competent and efficient organisation enabling others to succeed

Added relevance

We will match the present-day needs of users and thus cover new areas of relevance to society, develop early indicators and use new channels of communication.

One common entry point

We will create one common entry point to all data on society for researchers, analysts and authorities

Extended cooperation

We will initiate and enter into additional mutually beneficial partnerships and strengthen our relationship with stakeholders and users.

High quality

We will ensure that the quality of our statistics, communication and documentation is of the highest international standard.

New data sources

We will find new data sources, ensure their concordance with existing data, and turn them into new knowledge of relevance to society.

New competences

We will develop our talents and attract competent and dedicated employees and managers to make our vision come true.

Stronger development

We will pursue an ambitious and bold development agenda, experiment and test new opportunities whenever there is potential for improvement.

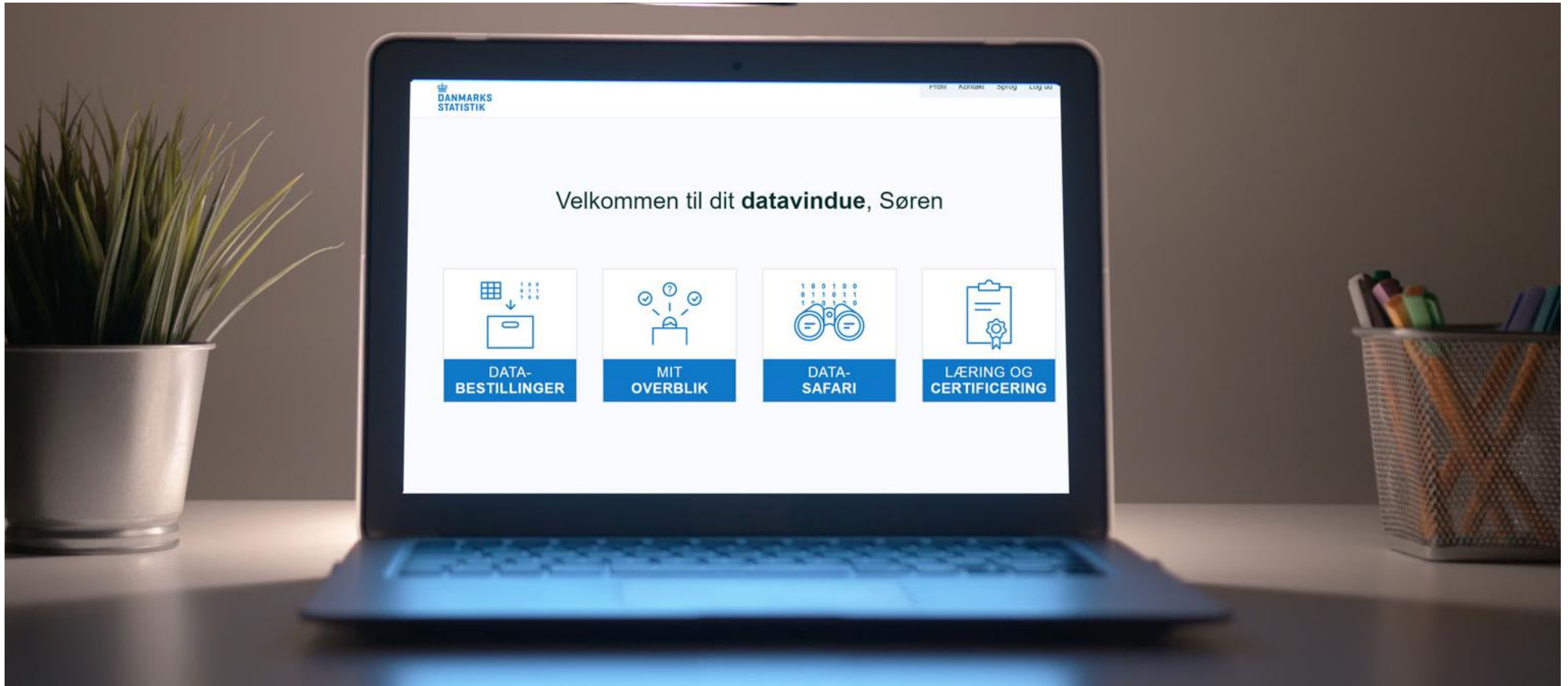
High data security

We will maintain a leading, well-established and certified data environment, where security is always a primary consideration

Improved efficiency

We will improve our efficiency using new technology and work towards efficient use of our own resources as well as those of others.

The new gateway: Danish Data Portal



What's inside the new Data Warehouse

- 19 different business statistics registers
- Of which 9 new
- Continuous updating
- Improved metadata and documentation
- We aim for adding
 - registers addressing the green economy
 - short term statistics



Connecting the dots – in new dimensions

Linking business statistics with other data sources

Privately held data:

- Electric meter data
- Scanner data
- Platform data

Social statistics registers:

- Education
- Employment
- Income
- Health



New / future admin data:

- CSR-D (ESG) data
- Traffic toll data
- Climate Atlas

Program

- 10:15 - 11:00 **Presentation of the Data Warehouse for Business Statistics**
Peter Bøegh Nielsen, consultant, and Kalle Emil Holst Hansen, senior advisor, Statistics Denmark
- 11:00 - 11:15 **Coffee Break**
- 11:15 - 11:45 **Testing the new documentation of source at variable level in Accounting Statistics**
Andreas Kuchler, principal economist, Danmarks Nationalbank
- 11:45 - 12:45 **Lunch break (sandwich offered)**
- 12:45 - 13:15 **Merchanting and processing firms**
Jakob Roland Munch, professor, University of Copenhagen
- 13:15 - 13:45 **Data on Enterprise Owners,**
Anders Sørensen, professor, Copenhagen Business School (CBS)
- 13:45 - 14:00 **Closing remarks** by Søren Schiønning Andersen, Director Business Statistics, Statistics Denmark

Data Warehouse for Business Statistics

Kalle Hansen khs@dst.dk
Peter Bøegh Nielsen pbn@dst.dk



What does the previous situation reflect?

- Only few administrative registers holding information on enterprises available
- A unique identifier for enterprises across administrative registers only implemented from 1999 onwards (CVR no.) – before different id no.s for VAT and Tax purposes
- Most business statistics are based on survey samples => only few observed units per sample and even fewer represented in several survey samples hampering the linking of micro data across statistical registers
- The majority of business statistics registers has a sectoral coverage only; not covering the total business economy as often is the case for social statistics

The main objectives of the project

- Improve the access to business statistics microdata by adding new registers:
 - In total 19 registers of which 9 new registers
- Establish a data set with longitudinal relations
- Include a private source KOB to establish a time series going back to 1994, incl. Match with old FIRM
- Improve the metadata and the documentation of each register, incl. especially a source documentation at variabel level of FIRE
- Investigate the pros and cons of increased co-ordination of the survey samples
- DST internal improvements:
 - Consistency at micro level by increased validation of unit and data consistency
 - Introduce a support IT system to easier extraction of the relevant data

Which registers are included?

ESR – Statistical
Business register

FIRE – Structural
business statistics +
treatment data

FIRM – General
enterprise statistics

DEMO – Business
demography

IFATS – Foreign-owned
enterprises in Denmark

OFATS – Danish
foreign affiliates

FUI – Research,
development and
innovation statistics

VITA – ICT usage in
enterprises

FIDA – employer-
employee link

UHV – Trade in goods

UHT – trade in services
including merchanting
and processing

KONC –
Group register

ORGOUT –
international
organisation and
outsourcing

BEST – Board
members and directors

Longitudinal registers

GL FIRMA

GL REGN

KOB

New and improved metadata

– Statistical documentation

- The existing documentation in STATBANK

– Register documentation

- New documentation area describing the data series for micro data purposes, e.g. changes in definitions or administrative changes, influencing comparability over time

– Variable dokumentation

- The existing Times documentation, but in a new and extended format

FIRE: From source documentation at unit level to documentation at variable level

- Previously source declaration at **unit level** (J-kod):
 - Observed values from data collection (E,X,V)
 - Data from Tax authorities (SKAT) (S)
 - Non-observed values (R)
- New source declaration at **variabel level**:
 - Reported values to DST or calculated on reported values
 - Reported values from other sources (Tax, Digital Annual accounts reported to the Danish Business Authorities or Statistical Business Register).
 - Variables calculated primarily based on data from Tax or Danish Business Authorities
 - Calculated variables based on Tax or Danish Business Authorities variables
 - **Imputed variables based on activity, legal form and size (employment or turnover)**

Unit consistency

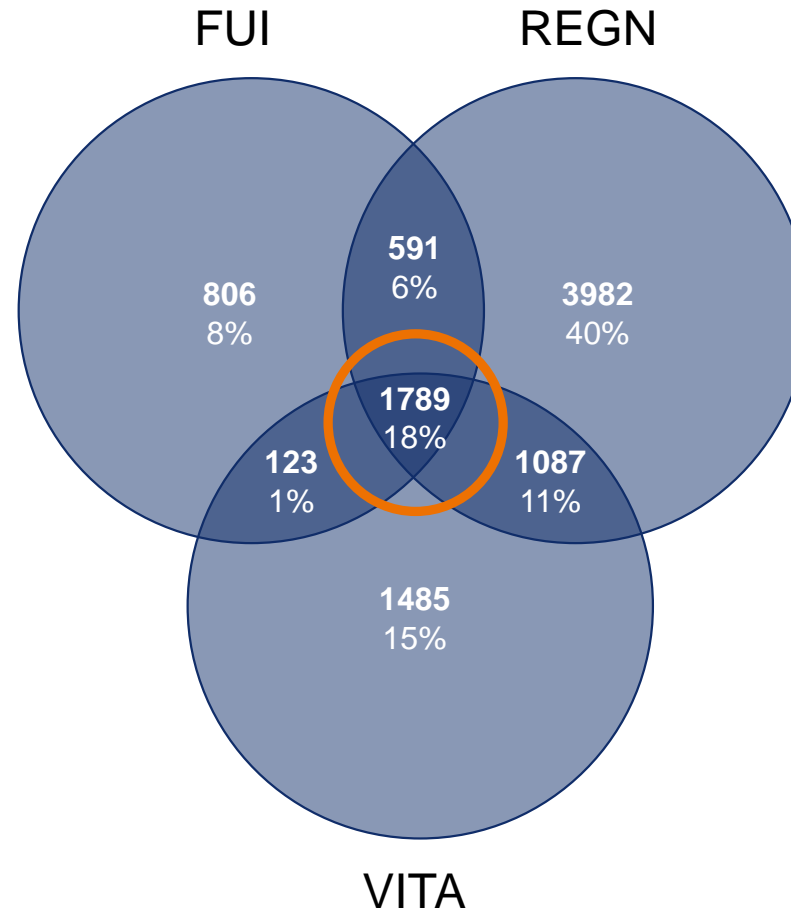
- Reasons for unit no match:
 - Different populations due to sampling criteria (activity or size coverage) or administrative rules (3-6-9 rule)
 - Demographic events between sampling and data collection
 - Different type of units (statistical and administrative units)
 - Complex units (enterprise groups) with several legal units in different activities
 - Use of different versions of the Statistical Business Register

Data consistency at micro level

- Possible reasons for inconsistencies
 - Different definitions (e.g. turnover)
 - Different periodicity
 - Inconsistencies in reporting to different statistics

Positive co-ordination of samples?

Match and no-match between the different survey samples



Case 1: Measuring SMEs

- **Area of interest:**

- *Importance of SMEs for Danish exports of goods*
- *When is a SME an SME?*

- **Selection choices:**

- *Exporting enterprise*
- *Independent enterprise or part of a group*
- *Size: SME or large*

**International
Trade in goods**
- Exporter

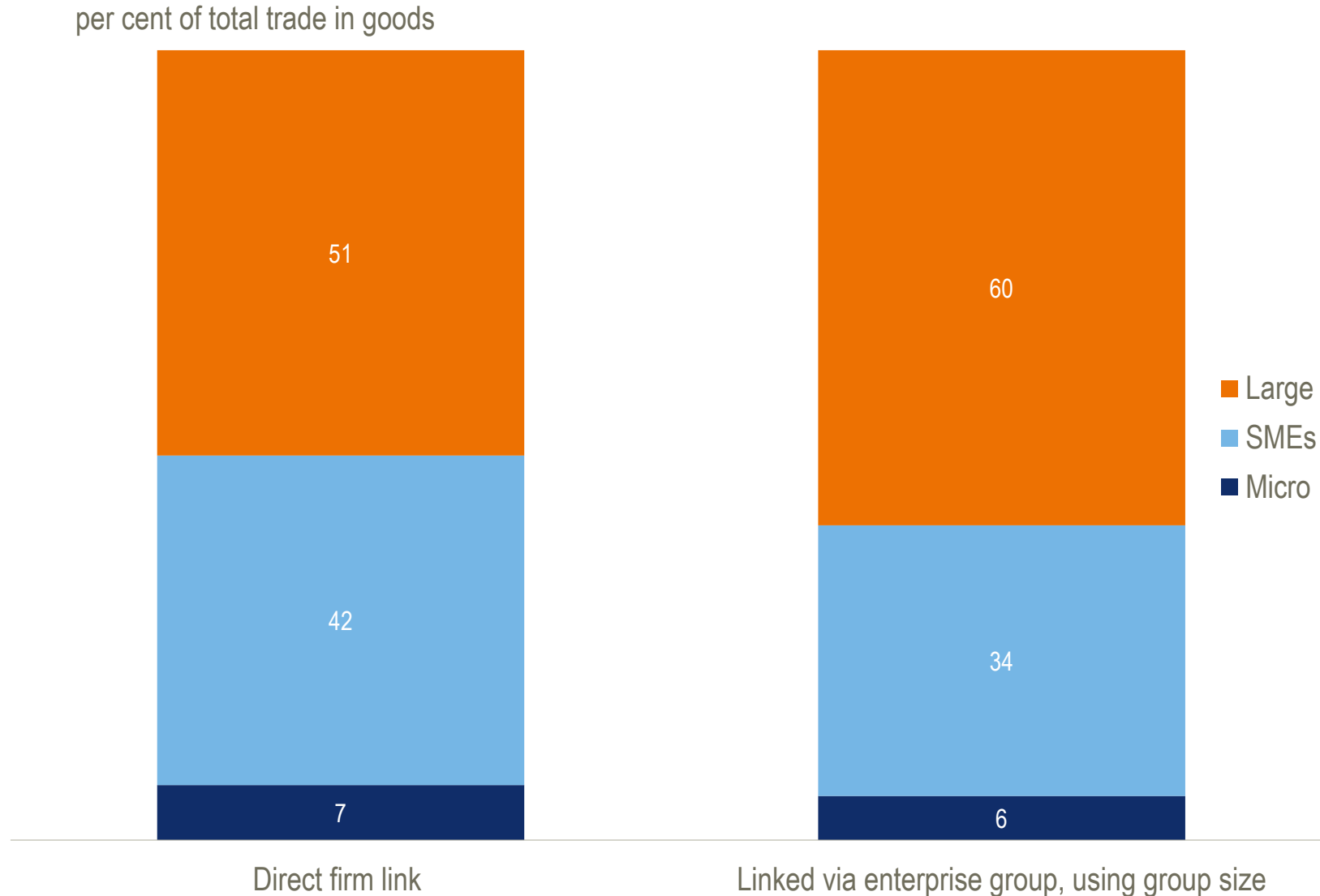


**Statistical
Business
Register**
- Size (employment)



**Enterprise
Group register**
- Group relation
- Size of group

How important are SMEs for Danish International trade in goods?



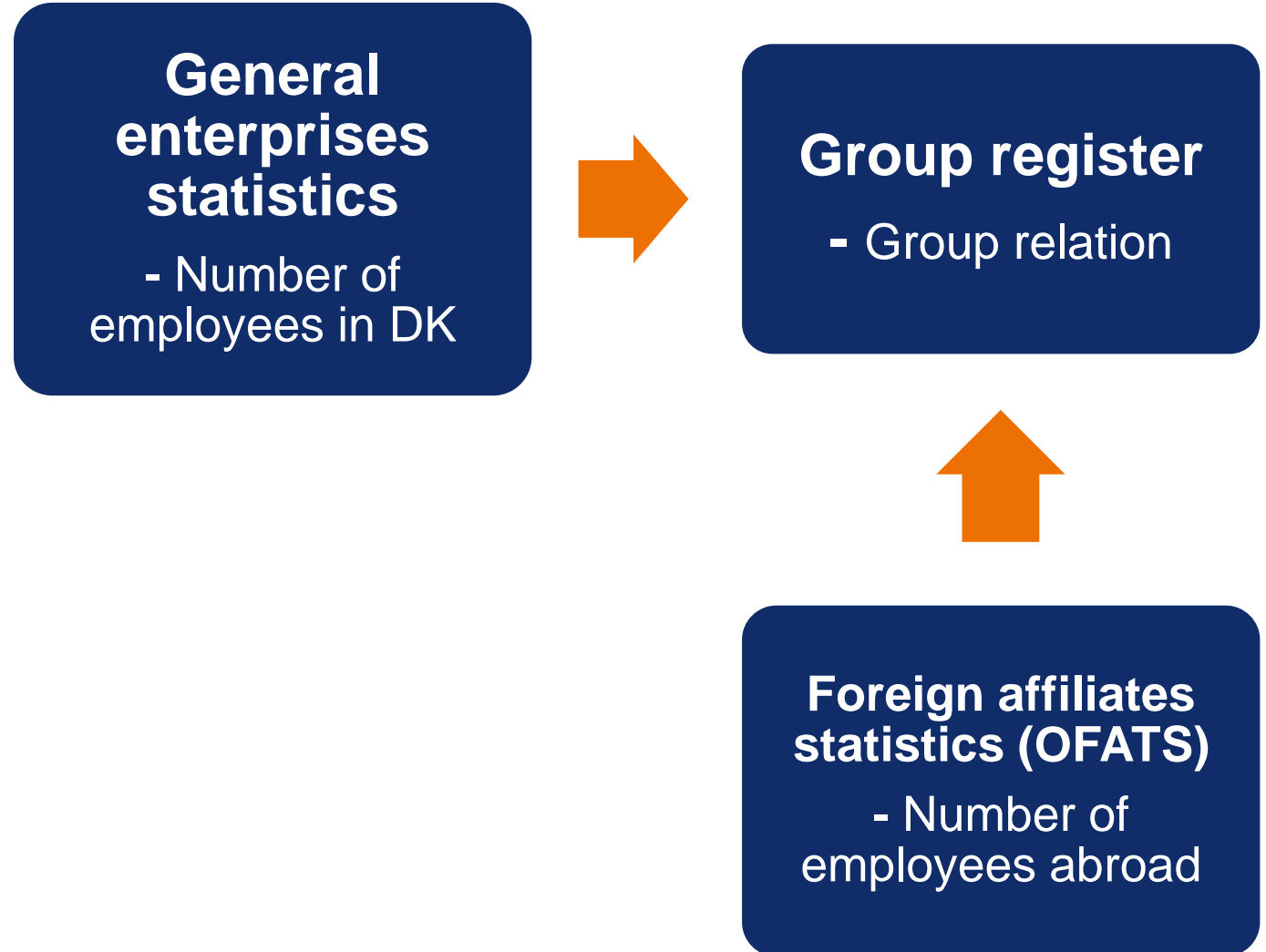
Case 2: Employment growth in manufacturing

- **Area of interest:**

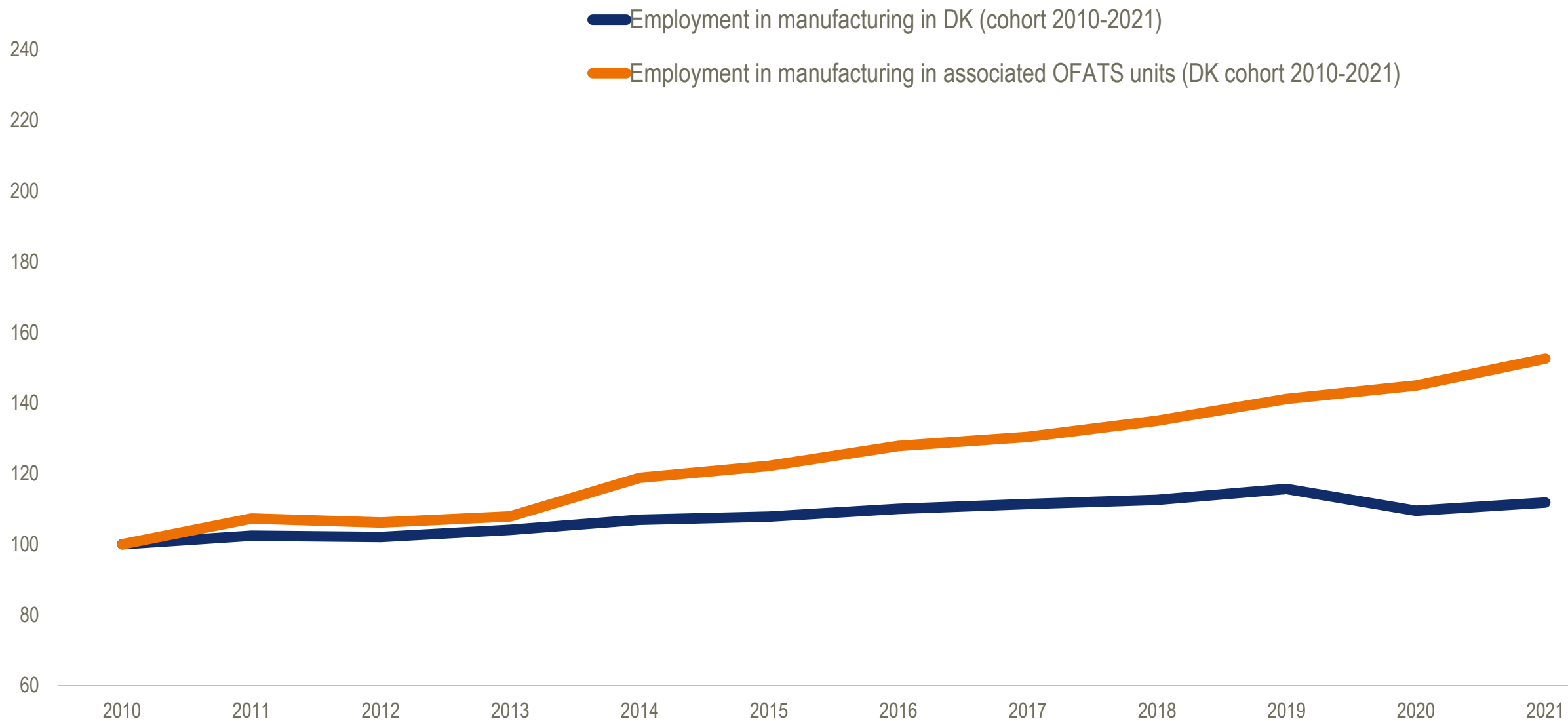
- *Firm organization*
- *Change in production*
- *Spatial division of labour*

- **Selection choices:**

- *Using register directly or using micro data linking?*
- *Using the group register as a go-between*
- *Which way - from DK to abroad or abroad to DK?*

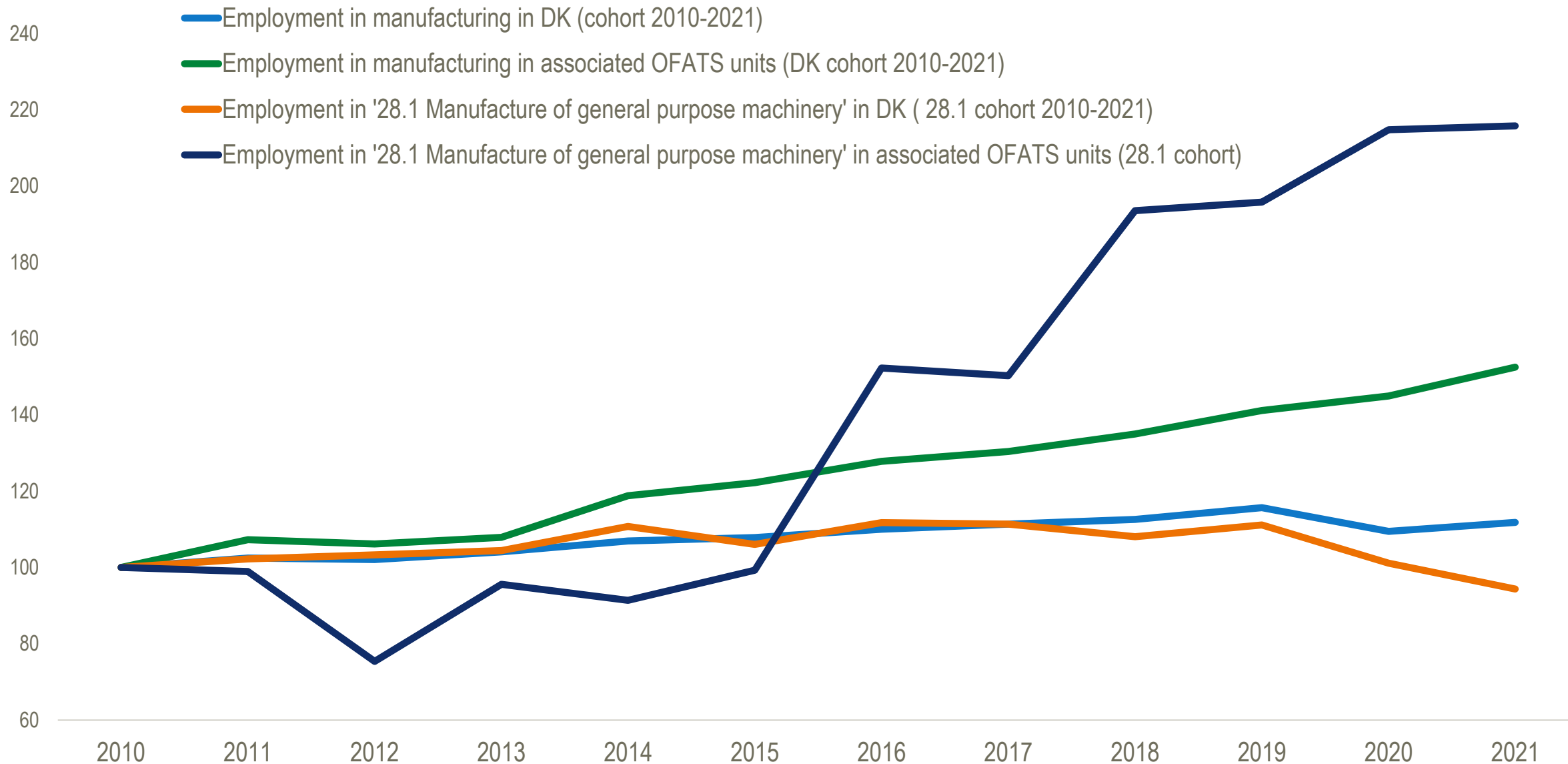


Employment growth in manufacturing (Cohort approach)



Employment growth in manufacturing (Cohort approach)

Index based on number of employees



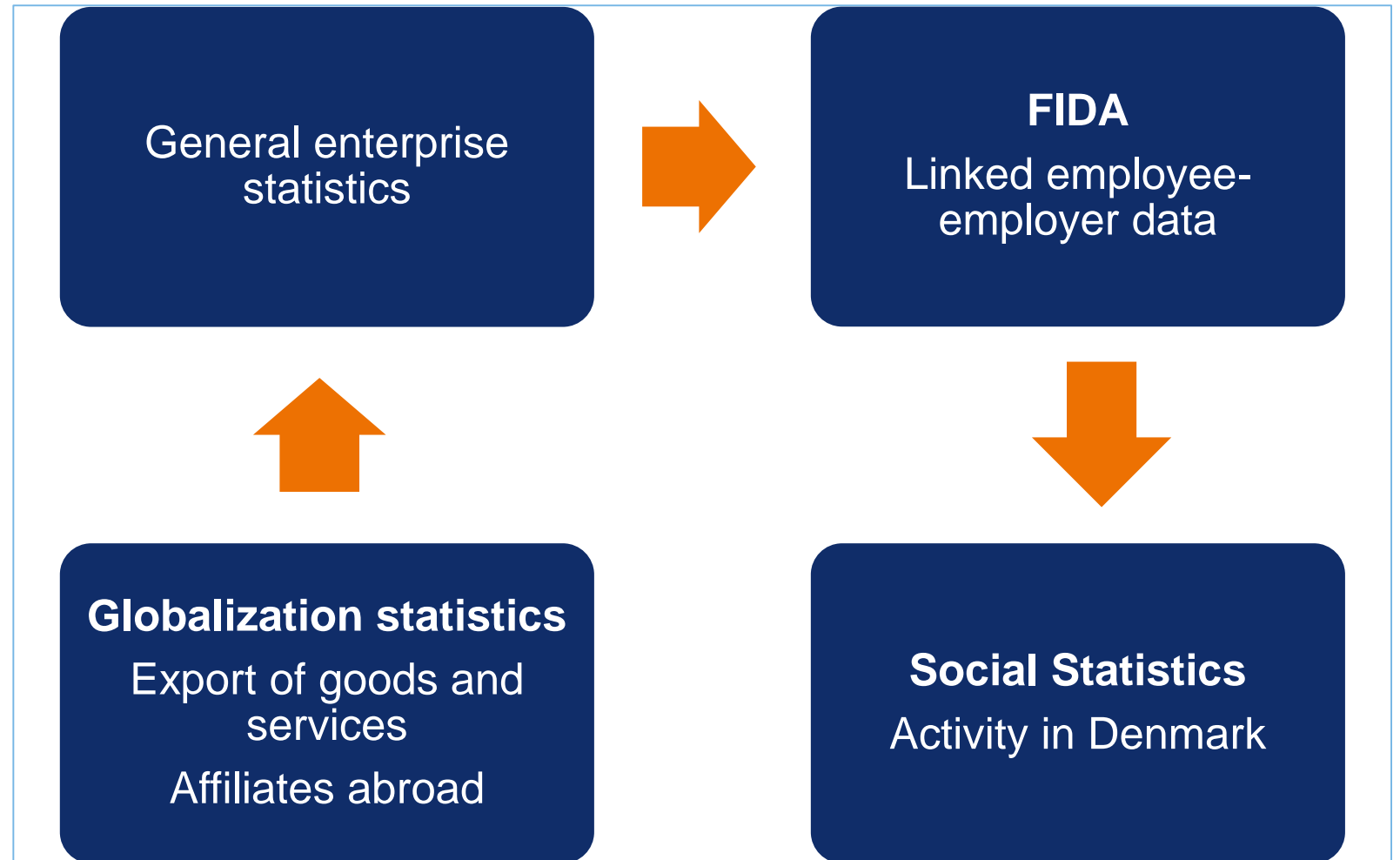
Case 3: Wage differences between globalized and non-globalised enterprises

- **Area of interest:**

- *Possible differences in wage structures in globalized versus non-globalized enterprises*
- *Matching of business and social statistics*

- **Selection choices:**

- *Exporters*
- *Danish companies with affiliates abroad*
- *Manufacturing*



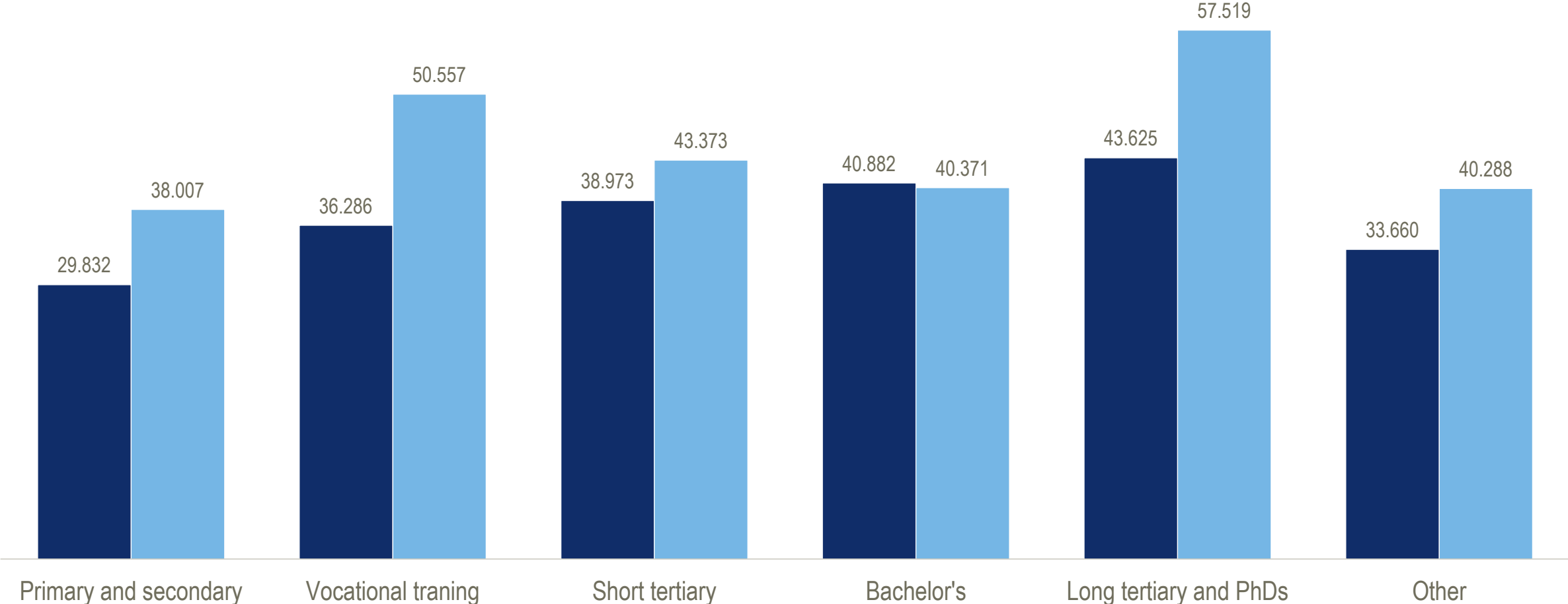
Wage difference by education in global and non-global enterprises

kr.

■ No global activities

■ Global sales (export) and or affiliates abroad

Enterprises within manufacturing activities



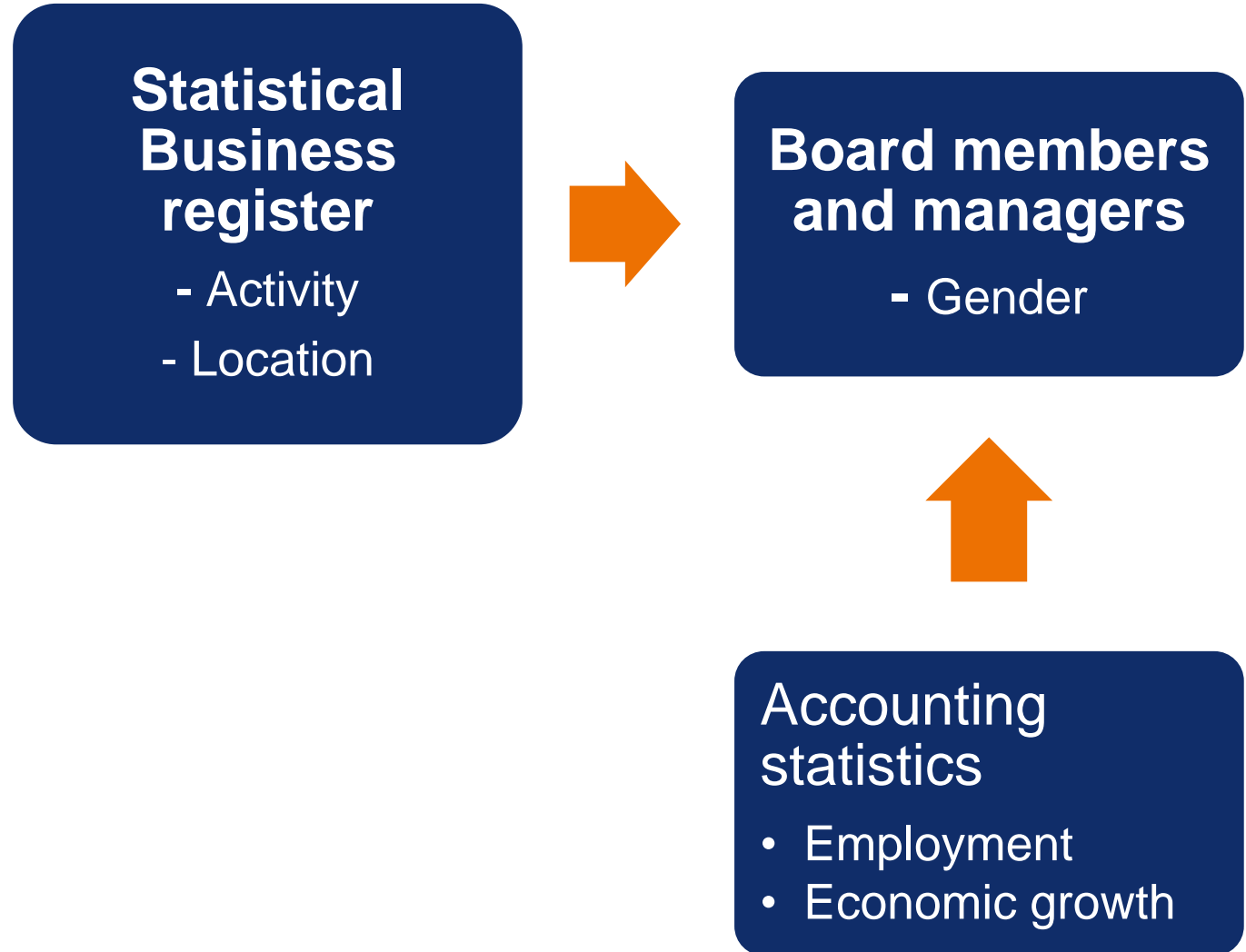
Case 4: Female directors

- **Area of interest:**

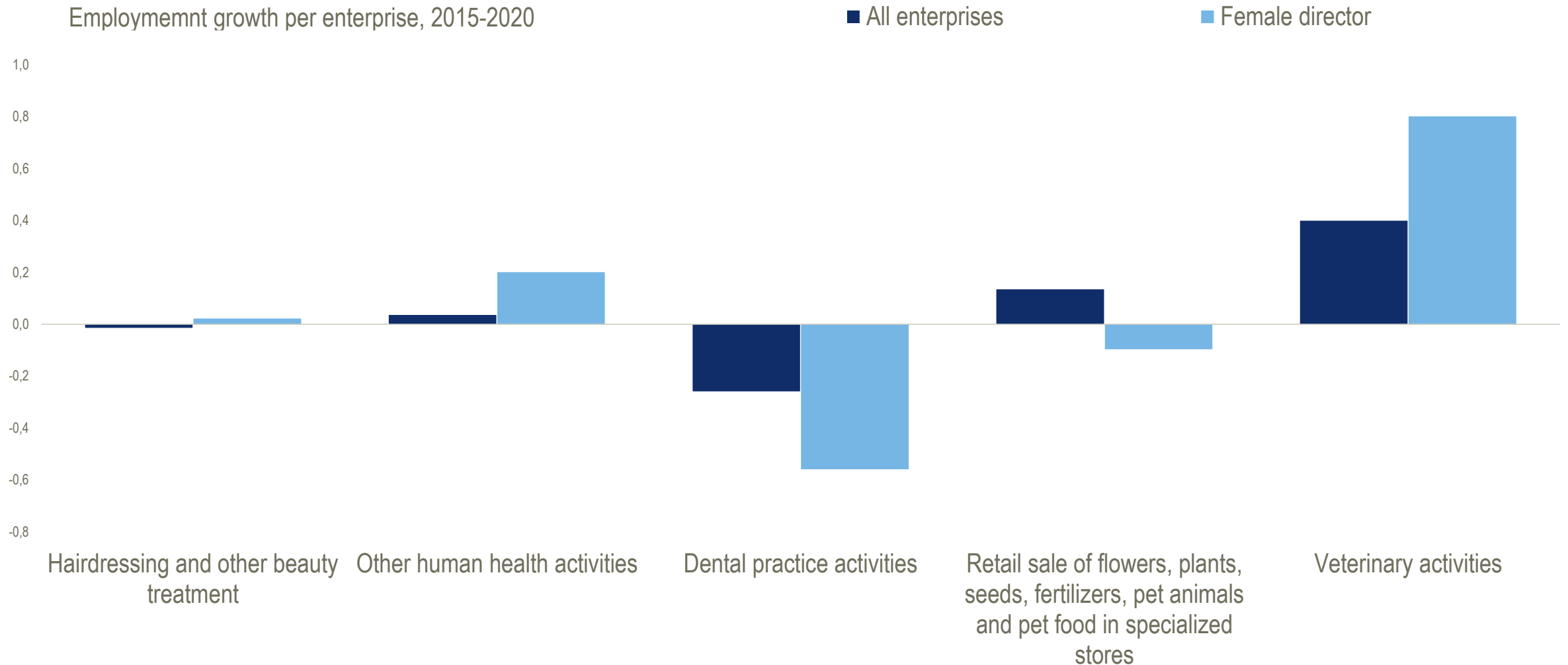
- *In which activities do we find the largest share of female directors?*
- *Are there differences in employment growth?*

- **Selection choices:**

- *Gender*
- *Activity*
- *Employment*



Employment growth in activities with a high share of female directors



Case 5: Old FIRM and KOB

- **Area of interest:**

- *Establish a longer time series holding accounting information than possible with DST data only*

- **Selection choices:**

- *Only limited liability and private limited companies*
- *Only activities within the non financial business economy subject to VAT*

Old FIRM

- Legal form:
 - Activity
- Selected accounting variables
- Employment



KOB

- Accounting variables

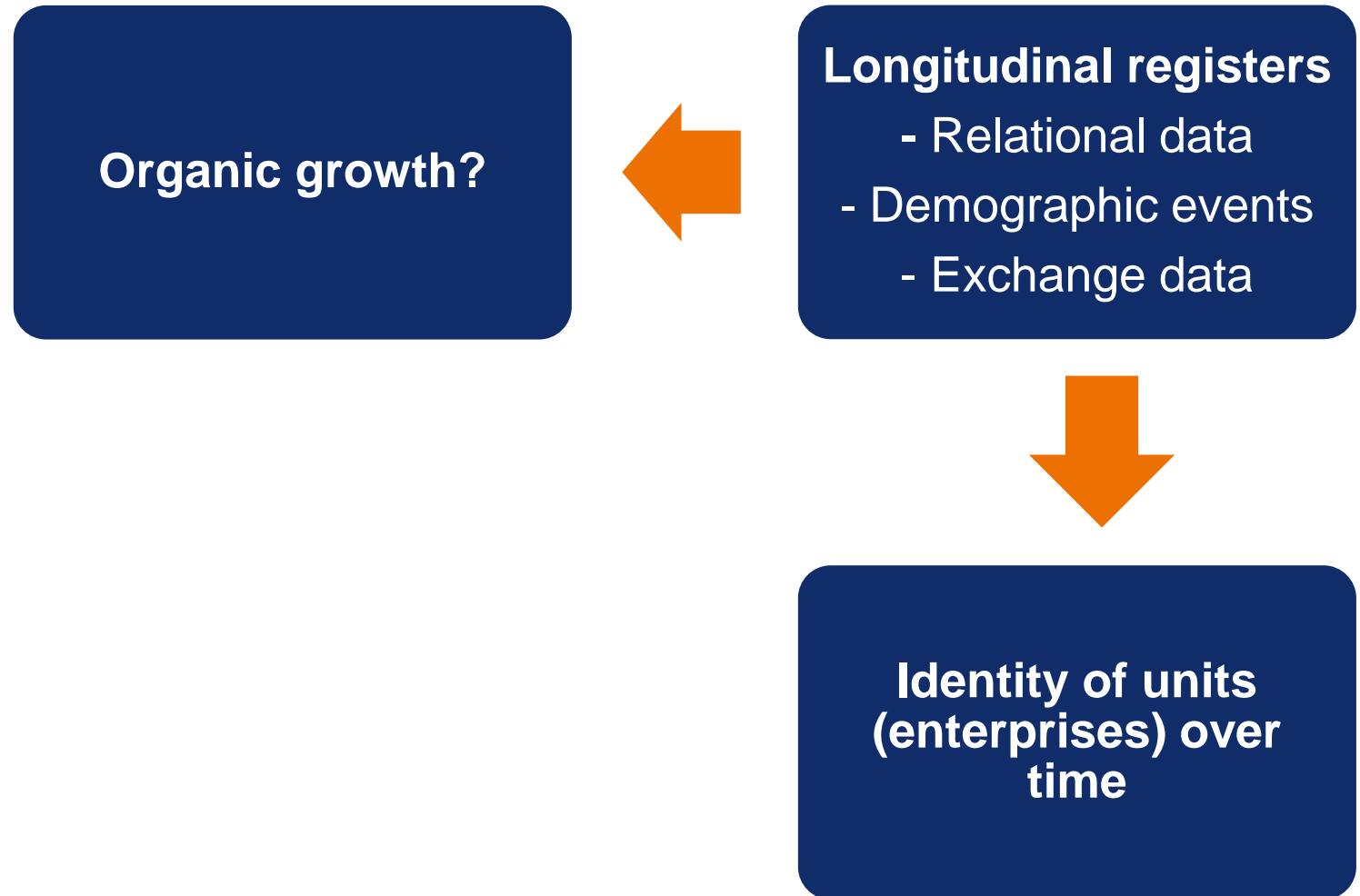
Case 6: Year-to-Year statistics or a longitudinal approach?

- **Area of interest:**

- *Activity specific growth analysis*
- *Correcting data for demographic noise*

- **Selection choices:**

- *Using legal units (old)*
- *Using enterprise IDs (new)*



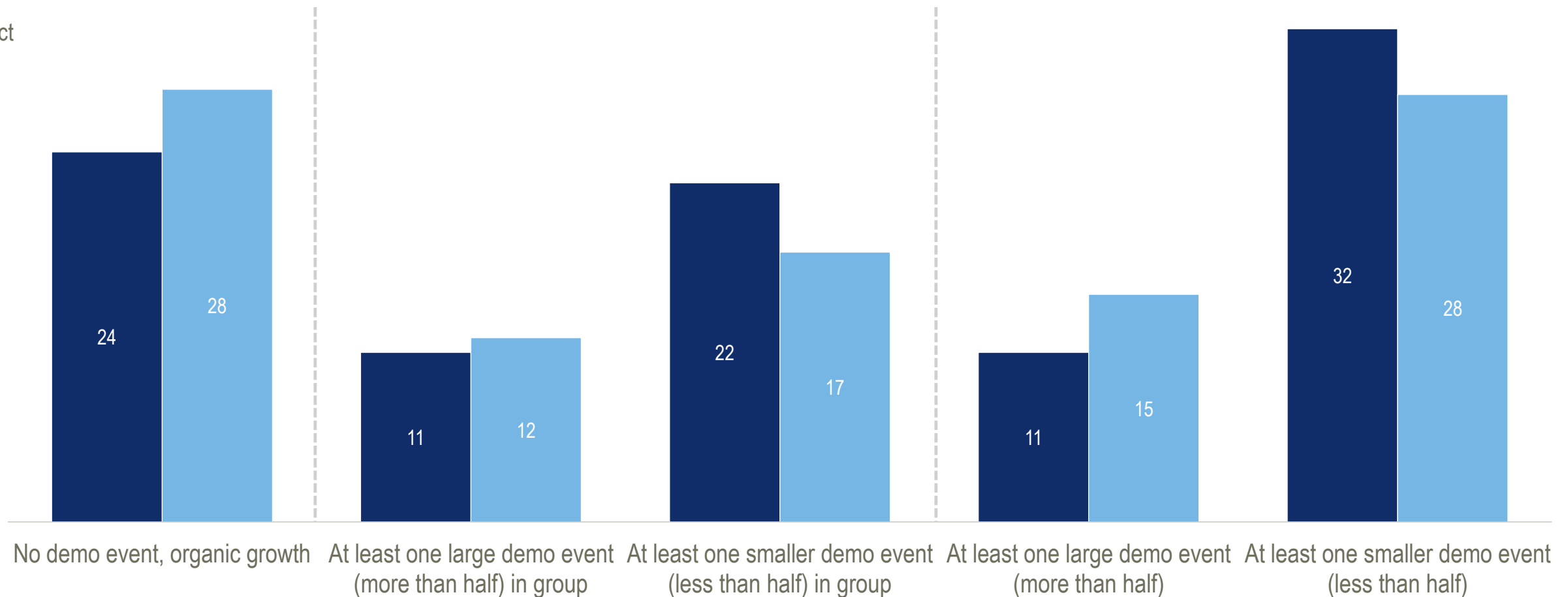
The 100 enterprises with the highest employment growth 2015-2020

The 100 enterprises with the largest employment growth 2015-2020

■ Share of enterprises

■ Share of population employment growth

pct



Quality challenges to be continuously addressed

- Complex enterprise structures
 - Monitor the largest groups
 - Introduce aGroup ID no.?
 - Eliminate intragroup transactions
 - How to determine activity of groups?
- Improve capturing of demographic events
- Increased data confrontation before final editing of data
- Addressing the conflicting goals of improving micro data research and minimise respondent burden

The (possible) road ahead

- **Work in progress:** International Trade in Services 2009 -2015, incl. merchandising and processing
- **Work in progress:** New register of owners
- New registers holding variables on "green economy"
- Include registers on short term statistics
- Inclusion of FDI register Danmarks Nationalbank?
- Linking to international SBRs such as EGR?
- Investigating the possibility of remote access to selected registers in the other Nordic Statistical Institutes
- Continuation of close co-operation and exchange of experiences and needs with the main users

Thank you

– Any questions?

DANMARKS
NATIONALBANK

New data on source of individual variables in the Accounting Statistics – tested by Danmarks Nationalbank

Andreas Kuchler
DRDS Conference, 30 November 2023

Background

The accounting statistics dataset (FIRE) provides balance sheet and income statement data for all active firms in Denmark in the non-primary, non-financial industries (“Private byerhverv”)

The data is collected through different sources, and some of it is imputed:

- DST-survey of firms, large firms overrepresented
- The business registry (Statistics Denmark)
- Data from Skat (the Danish tax agency)
- XBRL-data from Erhvervsstyrelsen (The Danish business authority)
- Data from Lægemiddelstyrelsen
- “Calculated” data based on the above sources
- Imputed data

New treatment data

DST has produced a new dataset, that for each data point in FIRE flags the source of data (e.g. reported, data from other sources, calculated, imputed)

This is a huge step forward compared to the existing JKOD (- 2016) and KODE (2017 -) variables, which did not specify exactly which variables were treated in what way

Nationalbanken has tested the new data. It is very simple to use and a great addition to the FIRE datasets. Though still in a testing phase, it has already allowed us to use a larger number of observations (with greater confidence) in our analyses

Previously...

GENEREL ERHVERVSSTATISTIK

2010:1 • 23. juni 2010

Regnskabsstatistik for firmaer 2008

- baseret på regnskaber afsluttet i perioden 1. maj 2008 - 30. april 2009

Resumé: Statistikken viser bl.a., at egenkapitalens forrentning i danske virksomheders faldt fra 20 pct. i 2007 til 10 pct. i 2008. Tallene er eksklusive landbrug og finansieringsvirksomhed.

Kort om statistikken: Statistikken har til formål at vise struktur og økonomisk udvikling i de private byerhverv.

Fordsætninger

5.6 Kvalitetsvurdering

Statistikken giver et pålideligt billede af det danske erhvervslivs økonomiske situation, da det bygger på en bred stikprøve, administrativ data og er baseret på detaljerede regnskaber.

Nogle af oplysningerne i regnskabsstatistikken er mere detaljerede end de poster, der kræves ifølge Årsregnskabsloven. Det gælder fx oplysninger om omkostninger til leje af arbejdskraft. Det kan betyde, at firmaerne kan have svært ved give disse oplysninger, og det er derfor sandsynligt, at data vedrørende disse oplysninger er undervurderet, udgifterne vil i stedet blive placeret på andre poster fx eksterne omkostninger eller vareforbrug, der dermed vil blive overvurderet.

Investeringsoplysninger fremgår heller ikke direkte af poster på regnskaber, der følger Årsregnskabsloven, men de kan som regel udledes af en tabel i noterne til regnskabet. Man må derfor også påregne en vis sandsynlighed for, at investeringerne er undervurderet blandt de firmaer, som udfylder og indsender spørgeskemaer til Danmarks Statistik. Først fra og med 2005 fås der fra SKAT oplysninger om investerings til- og afgang, så pålideligheden af investeringerne må antages at blive bedre fra dette år.

Regnskabsstatistikken er mere usikker på arbejdsstedsniveau end på firmaniveau, idet fordelingsmetoden hviler på antagelser. På trods af usikkerheden vurderes det dog, at oplysningerne er pålidelige på hovedbranche- og landsdelsniveau.

Fra og med 2005 fås flg. oplysninger fra SKAT: omsætning, vareforbrug, afskrivninger, ordinært resultat for finansiering og ekstraordinære poster, årets resultat før skat, selskabsskat, anlægsaktiver i alt, varebeholdning i alt, egenkapital, aktiver/passiver i alt, investeringstilgang og investeringsafgang. Disse regnskabsoplysninger skal indberettes af selskaber med en årsomsætning mellem ½ og 100 mio. kr. og af personligt ejede firmaer med en årsomsætning mellem 0,3 og 25 mio. kr. som en del af deres selvangivelse. Derudover fås fra SKAT arbejdsgivernes indberetning af lønudbetalingen til deres ansatte.

3.4 Datavalidering

Omfanget af og metoderne til datavalidering afhænger af kilden:

A) Der indhentes oplysninger af Danmarks Statistik direkte fra firmaerne vha. online indberetning fra de firmaer, der indgår i stikproven. Spørgeskemaet følger i så høj grad som muligt Årsregnskabsloven, og de indhentede oplysninger gennemgår en systematisk fejlsøgning i Danmarks Statistik. I forbindelse med fejlsøgningen af oplysningerne er der opbygget et automatisk fejlsøgningssystem, der giver meddelelse om fejl, hvis den indre konsistens, der skal være i et regnskab, ikke er opfyldt. Derudover giver fejlsøgningssystemet meddelelse, hvis der er oplysninger, der forekommer usandsynlige, når de sammenholdes med evt. tidligere afgivne oplysninger eller oplysninger fra lignende firmaer. Disse fejl eller usandsynlige oplysninger kontrolleres og rettes, evt. ved at firmaet kontaktes. For disse firmaer må regnskabsoplysningerne betragtes som meget pålidelige, og denne gruppe tegnede sig i 2021 for 78 pct. af den samlede omsætning og 70 pct. af de samlede passiver (inkl. firmaerne i B nedenfor).

B) Alle danske apoteker indsender regnskabsoplysninger til Lægemiddelstyrelsen, som sender Danmarks Statistik kopier heraf. Disse data fejlsøges meget lidt, da de betragtes som valide.

C) Regnskabsoplysningerne fra SKAT er ikke så detaljerede som på spørgeskemaerne. Regnskabsposterne fra SKAT betragtes i udgangspunktet som pålidelige, da de benyttes i forbindelse med skatteligningen. Dataene bliver fejlsøgt overordnet fx i forhold til andre kilder og tidligere år, regnskaber, der vurderes som usandsynlige, udledes fra statistikken. De firmaer, som i statistikken dækkes af oplysningerne fra SKAT, tegnede sig i 2021 for 11 pct. af den samlede omsætning og 2 pct. af de samlede passiver.

D) Regnskabsoplysningerne fra Erhvervsstyrelsen i XBRL-format er heller ikke så detaljerede som på spørgeskemaerne. Regnskabsposterne fra Erhvervsstyrelsen betragtes i udgangspunktet som pålidelige, da de benyttes i forbindelse med indberetning af årsregnskab til Erhvervsstyrelsen, dog kan der være lavet fejl i udarbejdelsen af regnskabet i elektronisk format, så det ikke stemmer med det trykte regnskab. Dataene bliver fejlsøgt i forhold til andre kilder og tidligere år, regnskaber der vurderes som usandsynlige udledes enten fra statistikken eller oprettes fx ved en manuel sammenligning med det trykte regnskab. De firmaer, som i statistikken dækkes af oplysningerne fra Erhvervsstyrelsen, tegnede sig i 2021 for 1 pct. af den samlede omsætning og 24 pct. af de samlede passiver.

I forbindelse med offentliggørelse af de foreløbige tal, foretages en begrænset datavalidering, der særligt er fokuseret på de største og mest betydende enheder. Denne offentliggørelse er derfor baseret på det indkomne materiale for dem i stikproven, som vi har nået at fejlsøge ca. 14 dage før offentliggørelsen af den foreløbige regnskabsstatistik. Samt det regnskabsmateriale som er modtaget fra SKAT og de elektroniske regnskaber som er indleveret af firmaer til Erhvervsstyrelsen i XBRL-format.

3.5 Databehandling

Målet for bearbejdningen er at danne regnskabstal svarende til spørgeskemaets detaljeringsgrad for hvert enkelt firma i populationen. Bearbejdningen sker trinvis:

På grundlag af de indberettede oplysninger på spørgeskema-niveau fra de stikprøvedtrukne firmaer dannes for hvert stratum forskellige fordelings- og forholdstal, der skal bruges til at beregne regnskabsposter for de firmaer, hvor der anvendes regnskabsoplysninger fra Erhvervsstyrelsen. For de firmaer, hvor der anvendes regnskabsoplysninger fra Erhvervsstyrelsen, fastfrysers for det enkelte firma de driftsøkonomiske poster herfra, og de regnskabsposter, som herudover er med i spørgeskemaet, beregnes ud fra stratificerede fordelingsstal baseret på de firmaer, som har indberettet på spørgeskema. På grundlag heraf beregnes herefter for hvert stratum regnskabstal pr.

← Gå til oversigt

JKOD

Navn

JKOD

Beskrivende navn

Journaliseringskode

Gyldighed

Gyldig fra: 01-01-1994

Kode tekst

- E Firmaet har udfyldt regnskabsskema elektronisk gennem Erhvervs- og Selskabsstyrelsens digitale indberetningssystem
- R Firmaet er opregnet på baggrund af tal fra stikprøvedtrukne firmaer og firmaer fra SKAT
- S Firmaet har nogle oplysninger fra SKAT
- V Firmaet har udfyldt regnskabsskema elektronisk gennem virk.dk
- X Firmaet har udfyldt regnskabsskema elektronisk i XBRL format
- 0 Startværdi
- 1 Udsættelse
- 2 Udsættelse overskredet
- 3 Nægter
- 4 Burde indsende, men er fritaget af forskellige årsager (fx konkurs, likvidation, sygdom mv.)
- 5 Firmaet tilhører ikke populationen, fx fordi det har vist sig, at firmaets korrekte branche ikke ligger inden for regnskabsstatistikens brancheområde, eller at der er tale om en ren momsafregningsenhed uden egentlig erhvervsaktivitet
- 6 Skemaet indsendt i udfyldt stand (+ evt. regnskab medsendt)
- 7 Regnskab indsendt
- 8 Udfyldt af Dst alene på grundlag af alternative kilder (fx SKAT, E og S, Lægemiddelstyrelsen)
- 9 Ventehyde

New codes available for each data point in FIRE

1. Indberettet til Regnskabsstatistikken eller beregnede på indberettede variable. Kan være rettet i forbindelse med fejlsøgning eller justeret
2. Data fra andre kilder herunder ESR, SKAT eller Erhvervsstyrelsen (XBRL). Kan være rettet i forbindelse med fejlsøgning eller justeret.
3. Beregnede variable, dog meget støttet fra variable fra SKAT eller Erhvervsstyrelsen (XBRL)
4. Beregnede variable, dog støttet fra variable fra SKAT eller Erhvervsstyrelsen (XBRL)
5. Imputerede variable på baggrund af branche, selskabsform, størrelse (årsværk eller omsætning)

Example – value added

Value added is calculated as

Omsætning (OMS)

+ Arbejde udført for egen regning og opført under aktiver (AUER)

+ Andre driftsindtægter (ADR)

+ Forøgelse/formindskelse af lagervareforbrug

- Køb af varer til videresalg (vareforbrug, handelsvarer) (KVV)

- Køb af råvarer, hjælpematerialer, færdigvarer og emballage (KRHE)

- Køb af energi (KENE)

- Køb af lønarbejde og underentrepriser (KLOE)

- Udgifter til husleje (UDHL)

- Udgifter til anskaffelse af småinventar/driftsmidler med kort levetid (UASI)

- Udgifter til vikarbureauer (UDVB)

- Udgifter til langtidsleje og operationel leasing (ULOL)

+ Ordinære tab på debitorer (konstaterede tab samt ændring i hensættelse til imødegåelse af tab (Post 11)) (OTDE)

- Eksterne omkostninger i øvrigt (bortset fra poster af sekundær karakter) (Post 12) (EKUD)

- Sekundære udgifter (SEUD)

Source: Documentation of GF_VTV (generel firmastatistik)

In principle no new information in data – but MUCH more user friendly presentation

. tab kode xvt_bhl

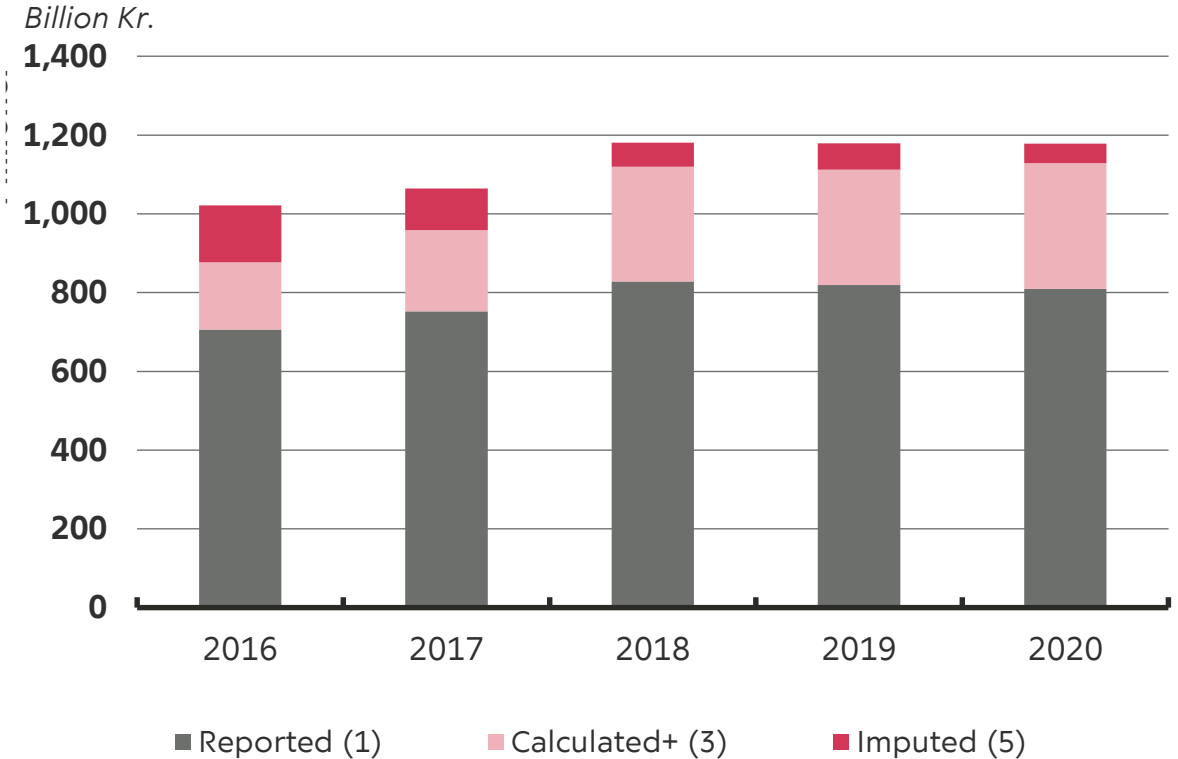
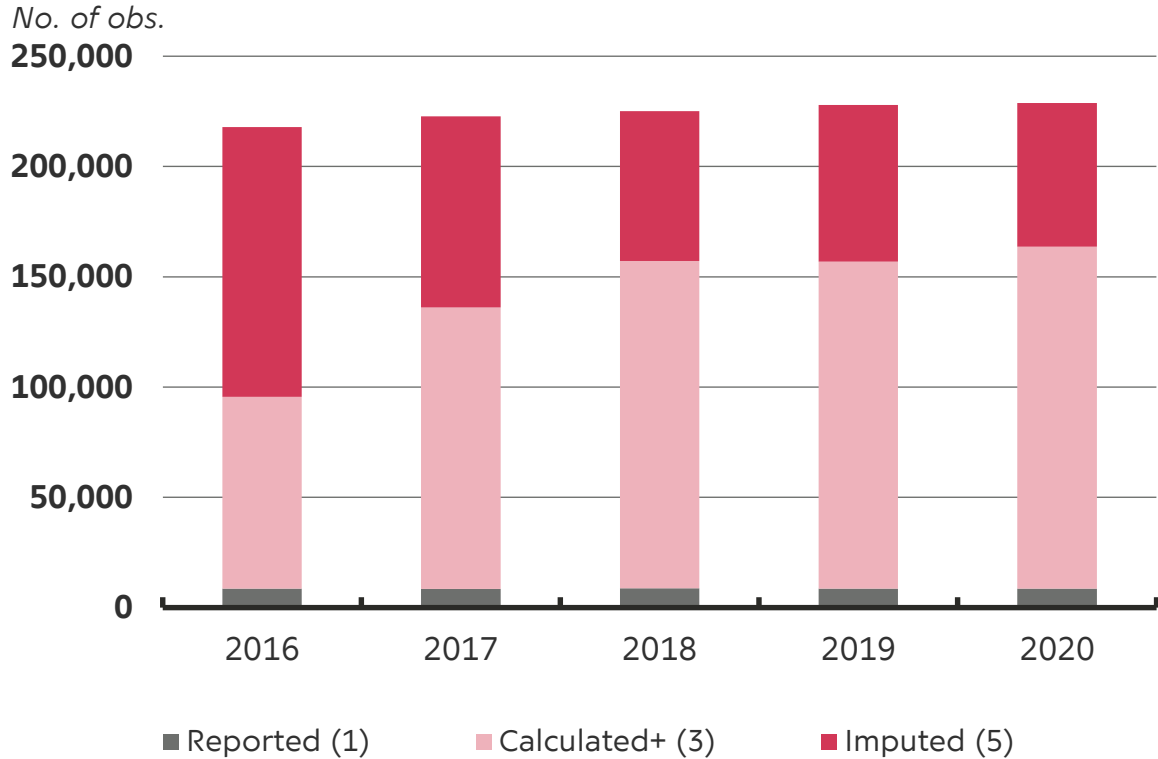
Datakilde	xvt_bhl			Total
	1	3	5	
3	0	251,901	0	251,901
4	0	174,431	0	174,431
5	0	29,727	0	29,727
14	0	106,993	0	106,993
15	0	16,185	0	16,185
16	0	0	263,894	263,894
17	0	0	26,981	26,981
67	28,652	0	0	28,652
991	874	0	0	874
995	851	0	0	851
999	4,024	0	0	4,024
Total	34,401	579,237	290,875	904,513

. tab jkod xvt_bhl

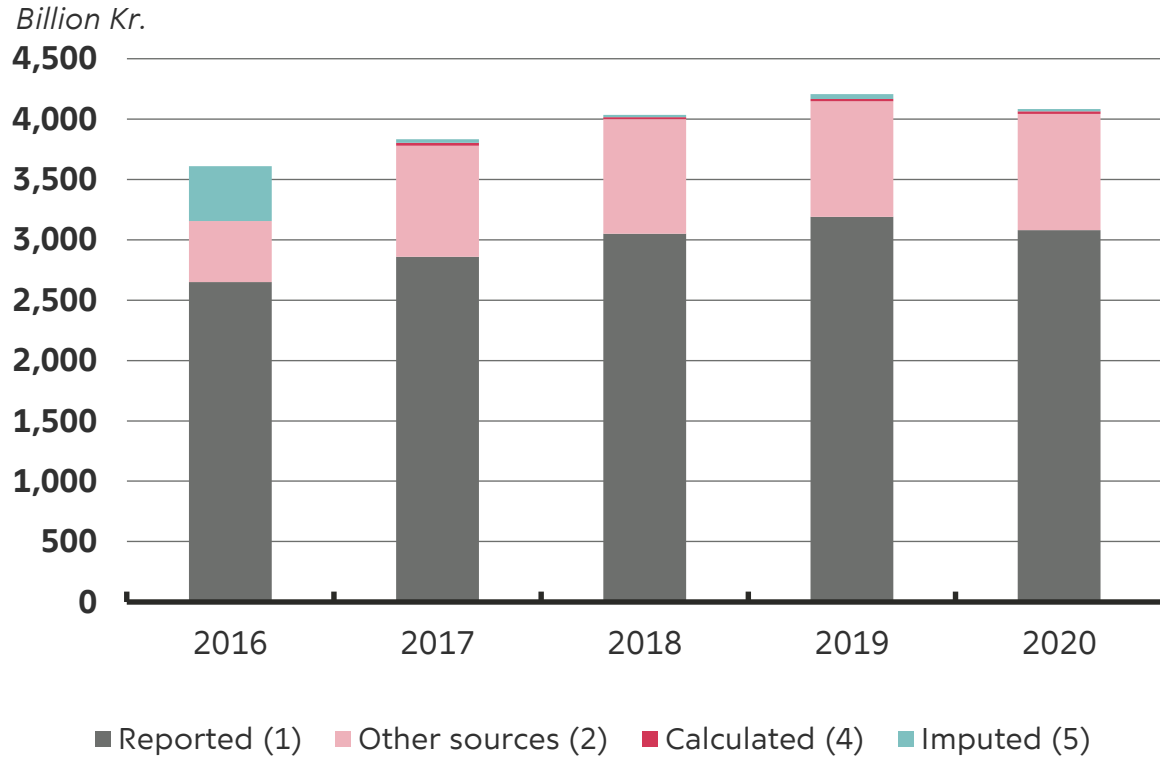
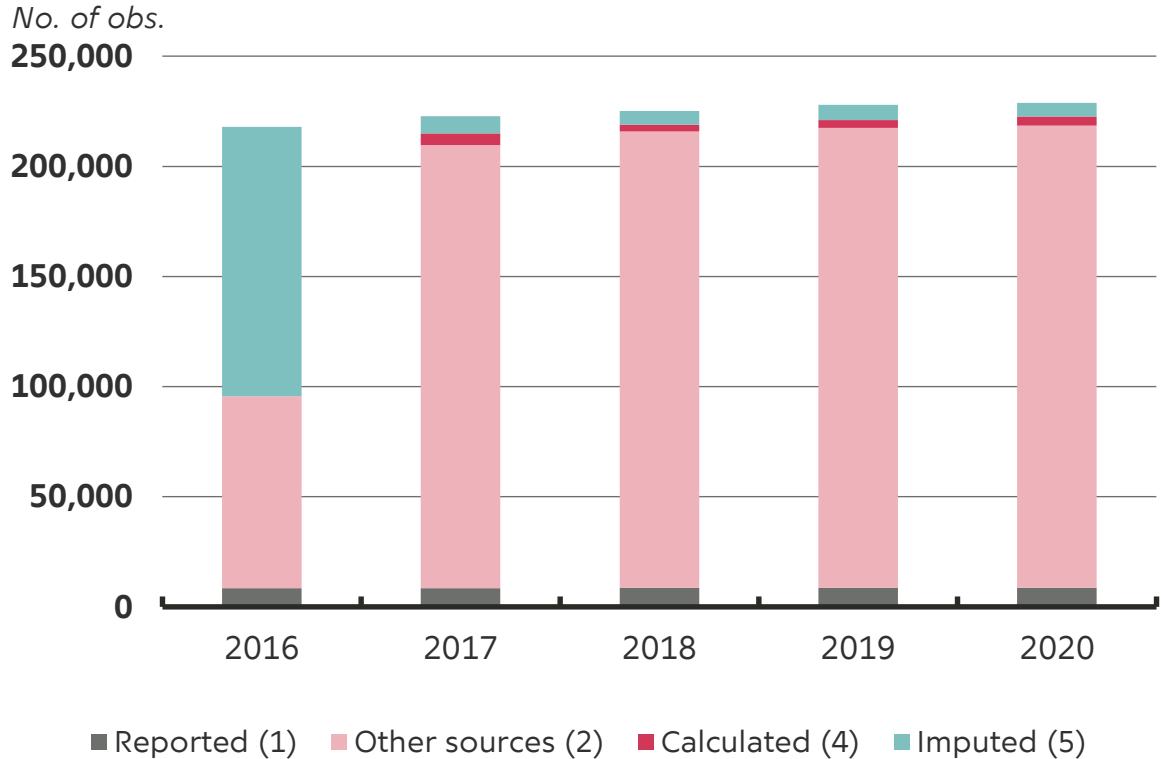
Journalise ringskode	xvt_bhl			Total
	1	3	5	
0	4	0	0	4
6	1	0	0	1
7	2,221	0	0	2,221
8	671	0	0	671
R	0	0	122,357	122,357
S	0	87,000	0	87,000
V	5,214	0	0	5,214
X	424	0	0	424
Total	8,535	87,000	122,357	217,892



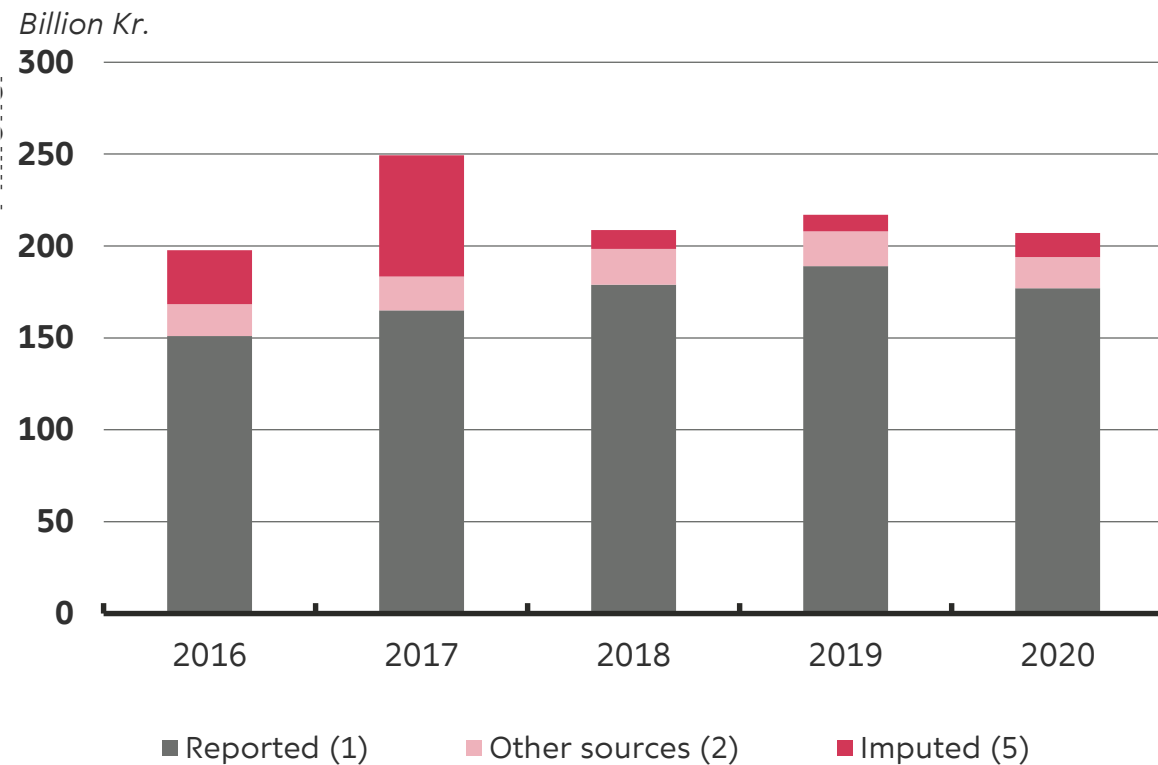
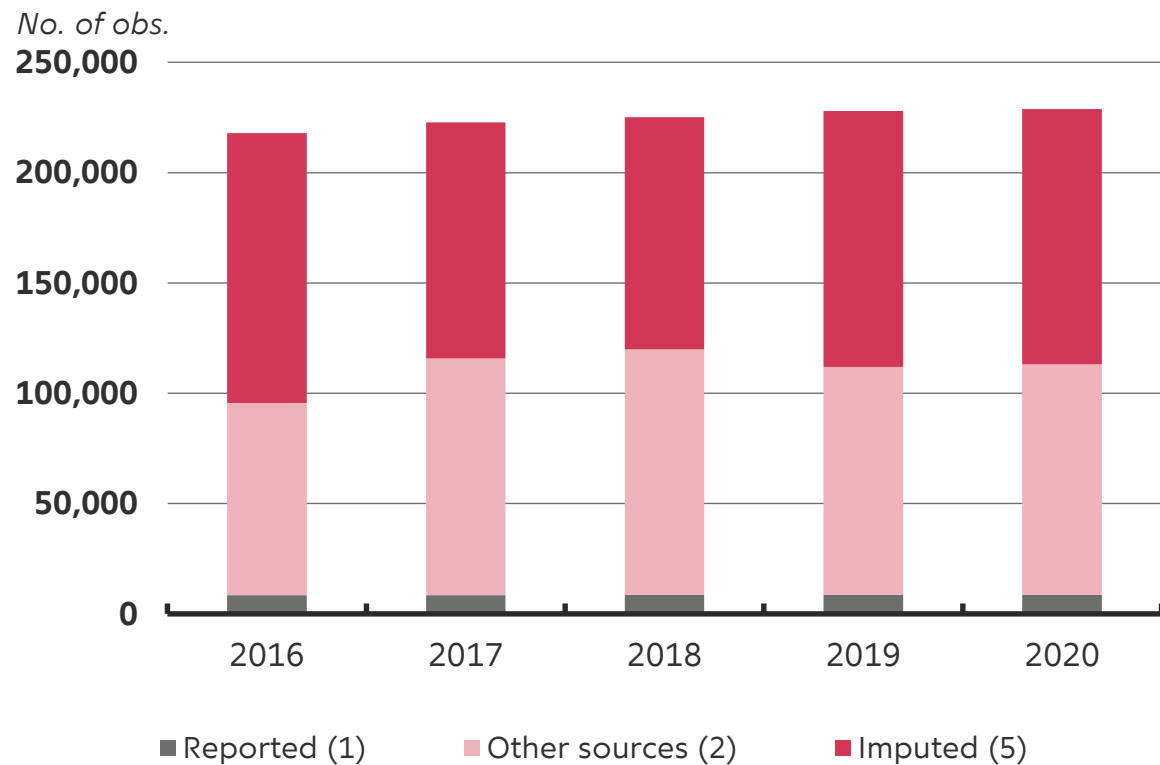
Value added (XVT)



Turnover (OMS)



Investment (ATIT)



Example: Calculation of investment ratio

. tab atit_bhl kode												
atit_bhl	3	4	5	14	15	Datakilde 16	17	67	991	995	999	Total
1	0	0	0	0	0	0	0	28,652	874	851	4,024	34,401
2	251,901	174,431	0	0	0	0	0	0	0	0	0	426,332
5	0	0	29,727	106,993	16,185	263,894	26,981	0	0	0	0	443,780
Total	251,901	174,431	29,727	106,993	16,185	263,894	26,981	28,652	874	851	4,024	904,513

. tab xvt_bhl kode												
xvt_bhl	3	4	5	14	15	Datakilde 16	17	67	991	995	999	Total
1	0	0	0	0	0	0	0	28,652	874	851	4,024	34,401
3	251,901	174,431	0	0	0	0	0	0	0	0	0	579,237
5	0	0	29,727	106,993	16,185	263,894	26,981	0	0	0	0	290,875
Total	251,901	174,431	29,727	106,993	16,185	263,894	26,981	28,652	874	851	4,024	904,513

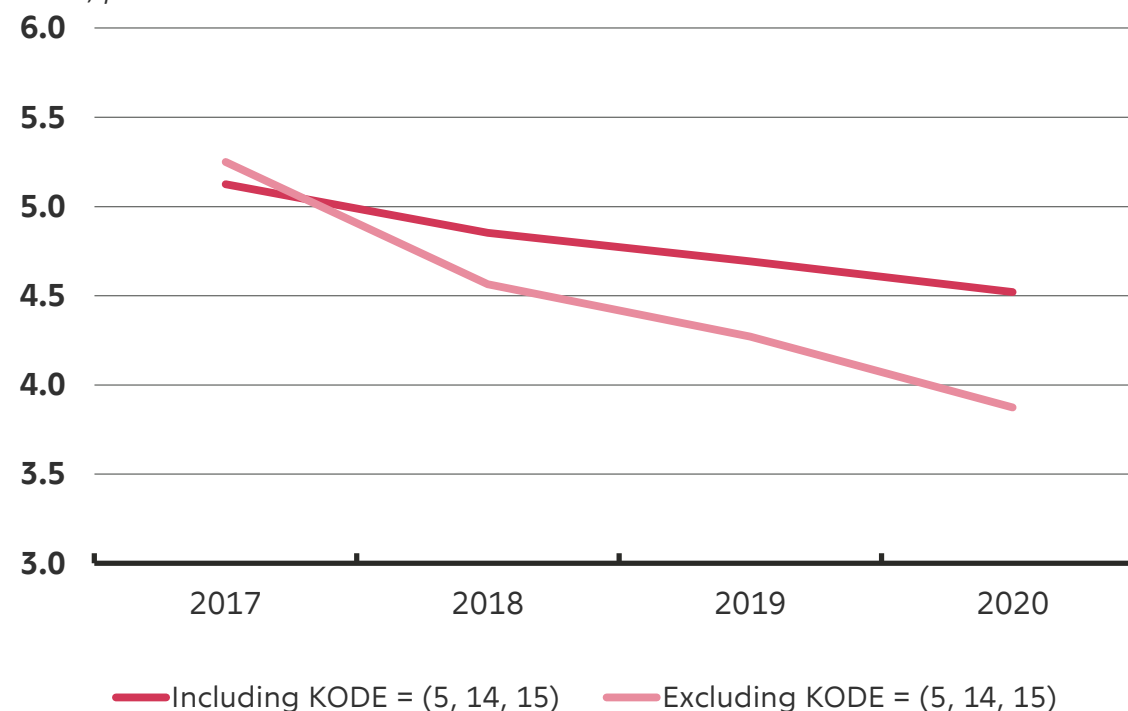
For kode 5, 14 and 15, ATIT (investment) is imputed but XVT (value added) is calculated.

Suppose we trust observations from surveys, SKAT etc. and derived variables strongly supported by these sources.

With KODE (and JKOD), one had to be very careful and could have mistakenly included those with KODE 5, 14 and 15.

With the new data, it is more clear in every case what is safe to include and what should be excluded -> probability of errors reduced.

Investment ratio, per cent



Conclusions

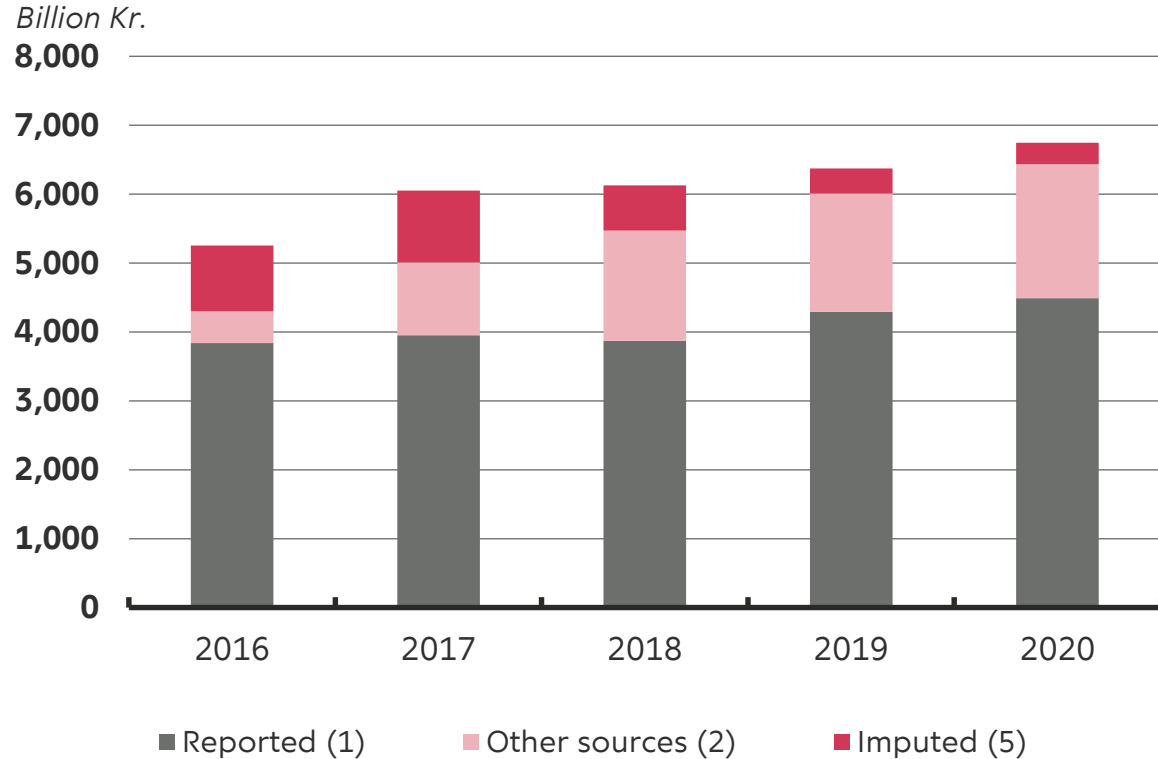
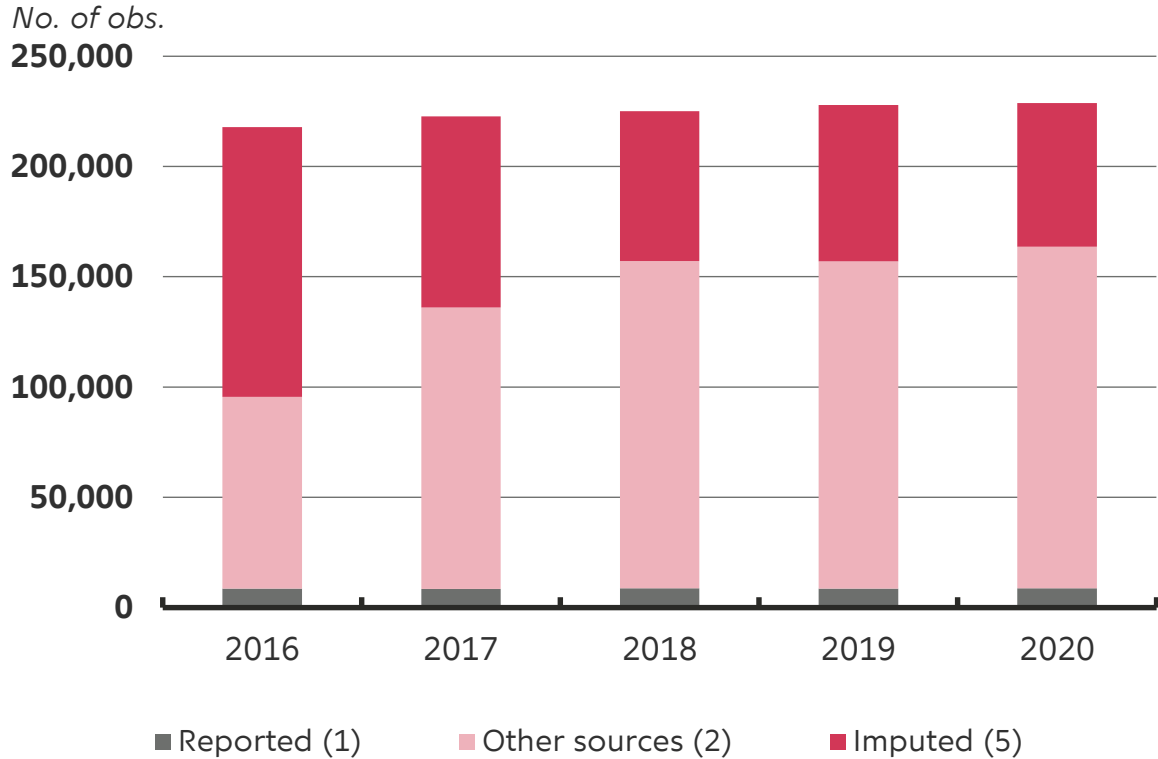
New source data is a very welcomed increase in user experience and would likely lead to

- A larger part of data being used when data quality can be assessed variable by variable
- Fewer errors
- Greater confidence in results
- Improval of weighting routines etc.

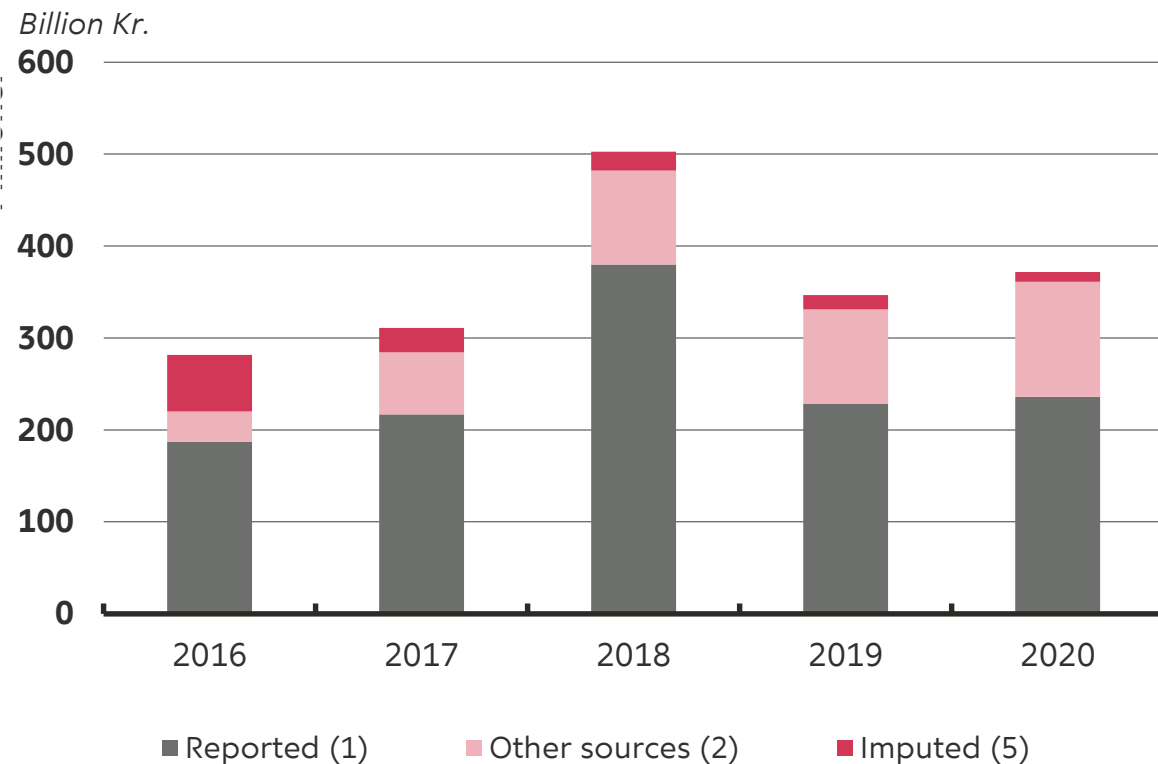
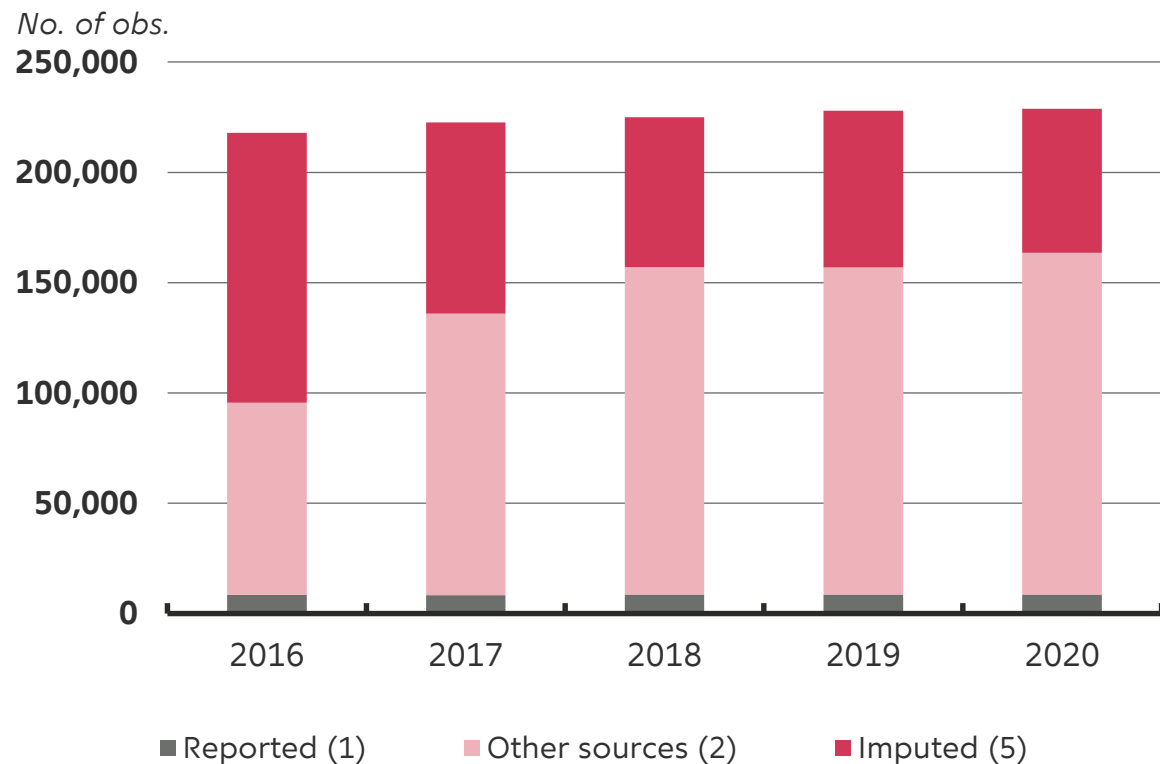
Thank you!

Extra figures

Total assets (AT)



Result (AARE)



Data on Enterprise Owners

Anders Sørensen, CBS

Many registers – missing ownership data

- DST has many registers
 - **Important dataset:** matched employer-employee dataset (FIDA) provides link between workers (PNR) and firms (CVR).
 - **Missing dataset:** matched enterprise-owner dataset with provides link owners (PNR) and firms (CVR).
 - Needed to study research questions related to ownership.
- In the past, researchers have used external data from, e.g., Experian/KOB to study ownership.
- This shortcoming will soon be removed as DST plans to provide data on “enterprise ownership”.

Examples of analyses using ownership data

- Entrepreneurial firms
 - Who opens new corporations?
 - Which new firms become successful/large/profitable?
- Ownership structure and dynamics
 - Characteristics of owners and their firms
 - How does ownership change over time?
- Teams of ownership
 - Teams vs. single owner firm.

Studies of entrepreneurs/owners

- Requires a key between CVR and PNR.
 - Existing options from DST:
 - Entrepreneurship data: IVPE and IVPS including new firms that are truly active.
 - Entrepreneurs are identified from various sources: founders/board members/directors/owners/employees...
 - An alternative approach is to identify the entrepreneur as the director in LLC firms, or the founder of sole proprietorships and partnerships in IDAN.
 - External data, e.g., Experian/KOB.

Outline - Description of planned ownership dataset

- Definitions of owner
 - Owners of corporations – include A/S, ApS, IVS, K/S, I/S.
 - Legal and real owner:
 - Real owner is a new definition was introduced to increase transparency in the ownership structures of Danish firms and prevent money laundering and tax havens.
 - Mandatory to disclose the real owners in CVR.dk from May 23, 2017.
- Data
 - Source is CVR.dk - the Central Business Register.
 - Variables in the dataset
 - Spell data for ownership.

Definitions of owner

- **Legal owners:**
 - Must disclose legal owners if >5% of the capital/voting rights.
 - Legal owners can be persons or firms.
- **Real owners:**
 - Real owners are *persons* who ultimately directly or indirectly own or control a sufficient part of a business.
 - Own a sufficient part of the shares/voting rights. Or exercise control by other means.
 - Must disclose its real owners if they have more than 25% of the ownership/voting rights, or if they exercise control by other means.

The data source: CVR.dk - Central Business Register

- CVR.dk encompasses all legal entities and businesses in Denmark. It contains current and historical information.
- Ownership data will be available from the “participation” (“deltager”) register from CVR.dk.
- Plan to provide ownership data for firms that exists in 2018 and forward for which data on real owners exist.

Variables in the Ownership Data

- CVR – firm-id
- Deltager-PNR – person-id's for legal and real owners
- Deltager-CVR – firm-id's for legal owners
- Part of shares
- Part of voting rights
- From and to dates of ownership
- Information on foreign ownership

Variable: From and to dates of ownership

- A few examples of ownership data
 - One person or firm owes the firm
 - More persons and/or firms can owe the firm.
 - Ownership can change over time.
 - Self-reported information by firms – misinformation.

One owner

CVR - Det Centrale Virksomhedsregister

Søg i CVR

Hjælp til CVR

 Kurv

Vild med maling ApS

CVR-nummer	42369764
Adresse	Vejlevej 106
Postnummer og by	6000 Kolding
Startdato	07.05.2021
Virksomhedsform	Anpartsselskab
Reklamebeskyttelse	Ja - Vilkår for brug
Status	Normal

One Owner

Legale ejere

[Laust Anders](#)

Egernvej 19

Strandhuse

6000 Kolding

Ejerandel: 100% (07.05.2021 -)

Stemmerettigheder: 100% (07.05.2021 -)

Ændringsdato: 07.05.2021

One legal owner

Reelle ejere

[Laust Anders](#)

Egernvej 19

Strandhuse

6000 Kolding

Ejerandel: 100,00% (07.05.2021 -)

Stemmerettigheder: 100,00% (07.05.2021 -)

Ændringsdato: 07.05.2021

One real owner

Two Owners and Changing Ownership

CVR - Det Centrale Virksomhedsregister

Søg i CVR

Hjælp til CVR

 Kurv

KREATIVE MALERE APS

CVR-nummer	39991799
Adresse	C/O FlexBiz Group Danshøjvej 52B
Postnummer og by	2500 Valby
Startdato	01.11.2018
Virksomhedsform	Anpartsselskab
Reklamebeskyttelse	Ja - Vilkår for brug
Status	Normal

Two Owners and Changing Ownership

Legale ejere

[Michella Smith](#)
Danshøjvej 52B, st. th
2500 Valby
Ejerandel: 50-66,66% (30.12.2019 -)
Stemmerettigheder: 50-66,66% (30.12.2019 -)
Ændringsdato: 30.12.2019

[M. Holding APS](#)
C/O FlexBiz Group
Danshøjvej 52B
2500 Valby
Ejerandel: 50-66,66% (01.11.2018 -)
Stemmerettigheder: 50-66,66% (01.11.2018 -)
Ændringsdato: 01.11.2018

[K. Holding APS](#)
C/O FlexBiz Group
Danshøjvej 52B
2500 Valby
Ejerandel: 50-66,66% (01.11.2018 -)
Stemmerettigheder: 50-66,66% (01.11.2018 -)
Ændringsdato: 01.11.2018

Reelle ejere

[Michella Smith](#)
Danshøjvej 52B, st. th
2500 Valby
Ejerandel: 50,00% (01.11.2018 - 02.01.2019)
Ejerandel: 50,00% (30.12.2019 -)
Stemmerettigheder: 50,00% (01.11.2018 - 02.01.2019)
Stemmerettigheder: 50,00% (30.12.2019 -)
Særlige ejerforhold: Har indirekte besiddelser (01.11.2018 - 02.01.2019)
Særlige ejerforhold: Har indirekte besiddelser (30.12.2019 -)
Ændringsdato: 30.12.2019

[Kaya](#)
Kongevejen 201, 1. tv
2830 Virum
Ejerandel: 50,00% (01.11.2018 -)
Stemmerettigheder: 50,00% (01.11.2018 -)
Særlige ejerforhold: Har indirekte besiddelser (01.11.2018 -)
Ændringsdato: 01.11.2018

▼ Ophørt ejerskab

Ophørte reelle ejere

[Michael](#)
Ejerandel: 50,00% (02.01.2019 - 30.12.2019)
Stemmerettigheder: 50,00% (02.01.2019 - 30.12.2019)
Særlige ejerforhold: Har indirekte besiddelser (02.01.2019 - 30.12.2019)
Ændringsdato: 30.12.2019
Afhændelsesdato: 30.12.2019

Two legal owners

Misinformation
Requires data cleaning

Two real owners

Many Owners and Changing Ownership

CVR - Det Centrale Virksomhedsregister

[Søg i CVR](#)

[Hjælp til CVR](#)

[Kurv](#)

DecoFarver ApS

CVR-nummer	35142592
Adresse	M.P. Allerups Vej 45H Fragdekærby
Postnummer og by	5220 Odense SØ
Startdato	18.03.2013
Virksomhedsform	Anpartsselskab
Reklamebeskyttelse	Ja - Vilkår for brug
Status	Normal

[Se alle enheder på adressen](#)

Many Owners and Changing Ownership

Legale ejere

LoveMondays ApS

C/O Repeatly Leasing A/S
Dampfærgevej 8, 2. tv
2100 København Ø
Ejerandel: 10-14,99% (30.12.2020 -)
Stemmerettigheder: 10-14,99% (30.12.2020 -)
Ændringsdato: 30.12.2020

T3 Invest ApS

C/O Jesper Stensgaard Christensen
Vanløsevej 9
8000 Aarhus C
Ejerandel: 20-24,99% (02.01.2018 - 29.12.2020)
Ejerandel: 10-14,99% (30.12.2020 - 05.10.2021)
Ejerandel: 15-19,99% (06.10.2021 -)
Stemmerettigheder: 20-24,99% (02.01.2018 - 29.12.2020)
Stemmerettigheder: 10-14,99% (30.12.2020 - 05.10.2021)
Stemmerettigheder: 15-19,99% (06.10.2021 -)
Ændringsdato: 06.10.2021

BDSJ Holding ApS

Svendborgvej 85
5750 Ringe
Ejerandel: 33,33-49,99% (21.12.2017 - 29.12.2020)
Ejerandel: 25-33,32% (30.12.2020 -)
Stemmerettigheder: 33,33-49,99% (21.12.2017 - 29.12.2020)
Stemmerettigheder: 25-33,32% (30.12.2020 -)
Ændringsdato: 30.12.2020

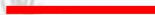




J.Laursen Holding ApS

Linde Allé 10
5750 Ringe
Ejerandel: 33,33-49,99% (21.12.2017 - 29.12.2020)
Ejerandel: 25-33,32% (30.12.2020 -)
Stemmerettigheder: 33,33-49,99% (21.12.2017 - 29.12.2020)
Stemmerettigheder: 25-33,32% (30.12.2020 -)
Ændringsdato: 30.12.2020

CoolGroup A/S

Loftbrovej 28
9400 Nørresundby
Ejerandel: 10-14,99% (15.06.2023 -)
Stemmerettigheder: 10-14,99% (15.06.2023 -)
Ændringsdato: 15.06.2023

From two to
five legal owners

December 21, 2017 
February 1, 2018 
December 30, 2020 
October 6, 2021 
June 15, 2023 

Many Owners and Changing Ownership

Two real owners

Reelle ejere

Brian Dennis Jensen

Svendborgvej 85

5750 Ringe

Ejerandel: 28,00% (18.03.2013 - 01.01.2018)

Ejerandel: 27,99% (02.01.2018 -)

Stemmerettigheder: 28,00% (18.03.2013 - 01.01.2018)

Stemmerettigheder: 27,99% (02.01.2018 -)

Særlige ejerforhold: Har indirekte besiddelser (18.03.2013 -)

Særlige ejerforhold: Ret til at udpege ledelsesmedlemmer
(18.03.2013 - 02.01.2018)

Ændringsdato: 02.01.2018

Jesper Laursen

Linde Allé 10

5750 Ringe

Ejerandel: 28,00% (18.03.2013 - 01.01.2018)

Ejerandel: 27,99% (02.01.2018 -)

Stemmerettigheder: 28,00% (18.03.2013 - 01.01.2018)

Stemmerettigheder: 27,99% (02.01.2018 -)

Særlige ejerforhold: Har indirekte besiddelser (18.03.2013 -)

Særlige ejerforhold: Ret til at udpege ledelsesmedlemmer
(18.03.2013 - 02.01.2018)

Ændringsdato: 02.01.2018

Legal and real owners in 2018

- Legal owners
 - for around 315,000 CVRs
 - around 80% of CVRs have one legal owner
- Real owners
 - for around 305,000 CVRs
 - around 80% of CVRs have one real owner
- FIRM2018 has are around 366,000 corporate CVRs

New firms founded in 2018

- Which firms become active firms? Need both active and inactive new corporations and their owners.
- Data excluding Holding companies.

Number of owners	Active	Non-Active	Total #cvr
1	4715	4457	9172
2	1270	1273	2543
3	223	283	503
4	46	55	101
5 or more	5	15	20
Total #cvr	6259	6083	12342

Summary and conclusions

- A highly potential new dataset on Enterprise Owners
- Spell dataset on owners of corporations
- Data will become available from 2018
- Data on sole proprietorships are not included

Closing remarks and topics to follow-up

Søren Schiønning Andersen
Director, Business Statistics



- Thank you!
- Today's presentations and an 'extended leaflet' will be uploaded to the website behind the QR code
- Doors will open 2nd January to most of the registers (data and metadata)
- Our dialogue continues:
 - At strategic level In the Advisory Board for the Danish Data Portal (DDV)
 - At functional level in the user forum
 - We will gladly come to you for more in-depth discussions
- We aim to expand the Data Warehouse in the coming years as mentioned
 - Let's maintain the active involvement of researchers in the reference group