The Data Warehouse for Business Statistics

Full content description

New and improved access possibilities to microdata for research and analysis

Introducing the new Data Warehouse for Business Statistics

In several fields of the social sciences, including political science, economics and sociology, as well as other scientific fields, there is a growing interest in micro-founded research. Micro-founded research is based on the characteristics and actions of individuals, households and enterprises for the study of, for example, educational choice, enterprise productivity, or - most central for a small open economy like the Danish economy - the economic globalization and associated enterprise choices. At the core of micro-founded research is the possibility of accessing and linking data across statistical registers by unique identifiers, as well as the option of linking them over time.

The access to anonymized high-quality microdata, facilitated by Statistics Denmark, has been fundamental for the Danish research community in recent decades. It has given Danish social science researchers a distinctive competitive edge, and research founded in this infrastructure is highly valued internationally. In an international comparison, the research infrastructure is unique and has been instrumental in attracting world-leading scholars to Danish universities.

Unique possibilities of linking data
So far, the available data at micro level has primarily been in the domains of social and health statistics, and to a lesser extent in the domain of business statistics. The new Data Warehouse for Business Statistics will widen and improve the access to anonymized enterprise microdata and give researchers access to unique data by new ways of linking data across a number of statistical registers. It will also facilitate the linking of business and social statistics at micro level, the so-called Linked Employer-Employee Data (LEED).

The Data Warehouse for Business Statistics

The new Data Warehouse for Business Statistics adds a variety of new statistical registers to the foundation of already existing high quality registers. This gives researchers more possibilities of linking data across registers and paves the way for new academic findings. To give a better overview of the data warehouse content, we have grouped it in four main categories of statistical registers that are described in the following pages:

Basic enterprise information

Enterprise structure, performance and dynamics

Globalisation

Innovation, R&D and ICT usage

Basic enterprise information

- **Statistical Business Register** (Det Erhvervsstatistiske register)

  Contains - for all legal units/enterprises in the private and public sector - information about the enterprise ID, the enterprise Group ID, administrative ID, start and end date for the enterprise ID and the enterprise Group ID, legal form, main activity of the enterprise and information on demographic relations (mergers etc.) and links to associated establishments. The register contains data on approx. 750,000 legal units.

- **Enterprise Group Register** (Koncernregister) **NEW!**
  Contains links between legal units and group heads and holds information on approx. 50,500 groups and nearly 330,000 legal units in the non-financial business sector with a Danish mother head.

- **Longitudinality Register** **NEW!**
  Contains information on demographic events and relations, such as birth and death of enterprises and cause of the event (merger, acquisitions or restructuring) or continuation between two years for all enterprises. The register contains possible links between approx. 500,000 enterprises.

DID YOU KNOW THAT...

+1,800 users

are accessing microdata at the Research Services, Statistics Denmark, each month?

Enterprise structure, performance and dynamics

- **Account Statistics for Non-Agricultural Private Sector** (Regnskabsstatistik for Byerhverv)

  Contains detailed accounting items such as turnover, value added at factor cost, total purchases of goods and services, wages and salaries and employment information. The statistics is partly survey based, partly based on administrative registers, including XBRL data (annual financial reports).

  Documentation of the source for each variable, e.g. if data is reported by survey, administrative data, calculated or imputed. It is possible to increase the number of observed data from approx. 8,000 to 150,000, depending on the variable. **NEW!**
  Read more in the documentation of statistics

- **General Enterprise Statistics** (Generel Firmastatistik)
  Contains economic and employment information concerning all sectors and industries, including agriculture and the public sector. The statistics is based on other statistical registers in Statistics Denmark and on administrative registers.
  Read more in the documentation of statistics

- **Business Demography** (Erhvervsdemografi)
  Contains enterprise births and deaths and provides information on turnover and employment. Data is based on information from the Statistical Business Register.
  Read more in the documentation of statistics

- **Annual Accounting and rating data from the private source KOB/Experian, covering 1992 -2018** **NEW!**
  These datasets are merged with data sets from the Statistical Business Register (Det Erhvervsstatistiske register) and Old Firm statistics data sets for the same period, allowing for longer time series of accounting data.

Globalisation

- **International Trade in Goods Statistics** (Udenrigshandel med varer)
  Contains data on imports and exports of goods by places of origin or destination, respectively, as well as by goods categories (Combined Nomenclature, CN), and the industry of the enterprise. The statistics is partly survey based with a sample of approx. 6,000 enterprises combined with data from administrative registers.
  Read more in the documentation of statistics

- **International Trade in Services Statistics** (Udenrigshandel med tjenester) **NEW!**
  Contains data on imports and exports of services by places of origin or destination, respectively, as well as by service categories (EBOPS), and the industry of the enterprise. The data sets also include data on so-called processing and merchanting. The statistics is survey based with a sample of approx. 3,000 enterprises.
  Read more in the documentation of statistics

- **Danish Foreign Affiliates** (Danske datterselskaber i udlandet) **NEW!**
  Contains data about the enterprise's foreign affiliates by industry and host country, the number of persons employed and turnover in foreign affiliates and the industry of the mother enterprise. The statistics is based on a census and the population (approx.5,000 mother enterprises) identified by use of administrative registers and privately held data.
  Read more in the documentation of statistics

- **Foreign Owned Enterprises** (Udenlandsk ejede firmaer)
  Contains data on the enterprise's country of ownership, account and employment information and the industry of the enterprise in Denmark. The statistics is based on administrative registers delineating the population combined with statistical registers concerning variables.
  Read more in the documentation of statistics

- **International Organization and Sourcing Survey** (International organisering og outsourcing)
  Contains data on international sourcing by business functions, incl. no. of job sourced, and geographical destination, international backsourcing and relocation, and the industry of the enterprise. The statistics is survey based with a sample of approx. 3,600 enterprises with 50+ employees.
  Read more in the documentation of statistics

Innovation, R&D and ICT usage

- **Innovation in the Private Sector** (Innovation i erhvervslivet)
  Contains data on product, innovation by technology level, process innovation by technology level, organizational innovation by technology level, marketing innovation by technology level. The statistics is survey based with a sample of approx. 3,500 enterprises with at least 10 full-time employees.
  Read more in the documentation of statistics

- **Business Enterprise Research and Development** (Forskning og udvikling i erhvervslivet)
  Contains data on engagement in R&D, R&D costs, R&D employment of the enterprises. The statistics is survey based with a sample of approx. 3,000 enterprises.
  Read more in the documentation of statistics

- **ICT Usage in Enterprises** (IT-anvendelse i virksomheder)
  Contains information about different types of ICT usage by the enterprise and amount of e-commerce. The information collected is dynamic and different topics are addressed annually. The statistics is survey based with a sample of approx. 4,560 enterprises with at least 10 full-time employees.
  Read more in the documentation of statistics

DID YOU KNOW THAT...

+17,000 files

containing research results are transferred from Research Services to users in one month alone?
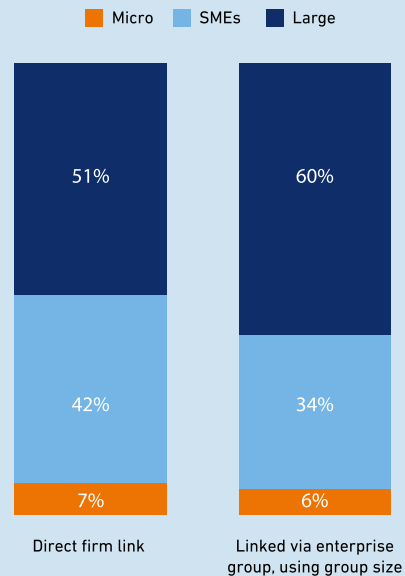
The importance of SMEs to the Danish trade in goods

As international trade is carried out by enterprises – not countries – it is relevant to analyse which firms are actually exporting. The traditional method has been to link information about the size of the exporting enterprises with the International Trade in Goods Register. This shows that close to half of total exports are carried out by SMEs (Small and Medium-sized Enterprises = less than 250 employees). But are SMEs always SMEs? Many small exporting enterprises are actually special purpose entities within a larger group. These enterprises are "small" only due to the organizational structure of the group.

Such division of SMEs can be identified by linking the International Trade in Goods Register with both the Statistical Business Register and the Group register. This is possible via the Data Warehouse for Business Statistics, and as a result, the export share of the SMEs – either as independent enterprises or as part of a group with less than 250 employees – decreases with 9 percentage points, to 40 per cent.

Percentage of total trade in goods, 2020

Micro   SMEs   Large

	Direct firm link	Linked via enterprise group, using group size
Large	51%	60%
SMEs	42%	34%
Micro	7%	6%

Sources: International Trade in Goods Register, Statistical Business Register and Enterprise Group Register
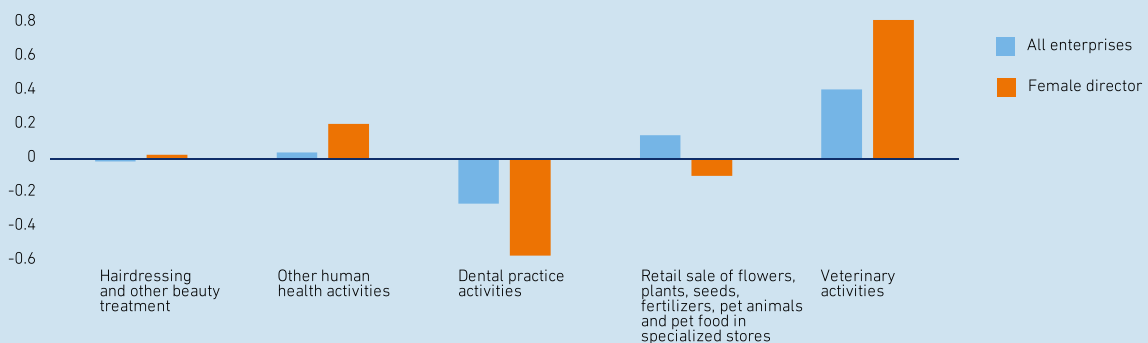
Enterprises with female directors

One of the Sustainable Development Goals of the UN is to "ensure women's full participation in leadership at all levels of decision-making". One dimension of this is to ensure female participation in the daily running of enterprises. This area can be studied via the statistics on Board Members and Directors. Simultaneously, we can assess the growth of enterprises by director gender via microdata, linking this register with e.g. the General Enterprise Statistics.

Looking solely at activities with more than 100 female directors in Denmark, it is only in enterprises within 'Hairdressing and other beauty treatment' (63.3 pct.) that the female share of directors is more than 50 pct. Subsequently the activities with the highest share of female directors are 'Other human health activities' (44.0 pct.), 'Dental practice activities' (43.9 pct.), 'Retail sale of flowers, pet animals, etc. in specialized stores' (43.6 pct.), and 'Veterinary activities' (42.3 pct.).
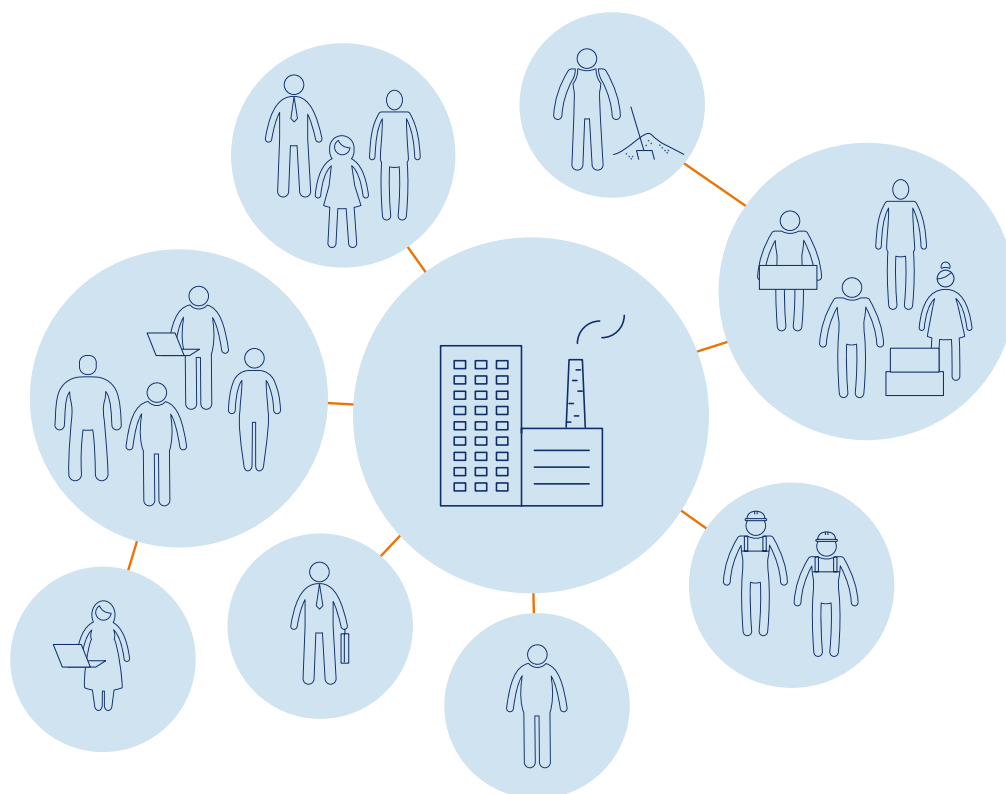
Looking at the employment growth of these activities, it is apparent that the female led enterprises are generally following and amplifying the growth trend of the specific activity as a whole.

Employment growth per enterprise, 2015-2020

All enterprises
Female director

Sources: Register of Board members and managers and General Enterprise statistics

Linking enterprise statistics and social statistics

The new Data Warehouse for Business Statistics also facilitates the linking of business and social statistics at micro level. This goes for the following registers/statistics:

FIDA NEW!
Contains the key between General Enterprise Statistics (Generel Firmastatistik) and Employment in Businesses (EiB) (Erhvervsbeskæftigelsen) with the purpose of linking data about the enterprise (via Enterprise ID) with data about the employees (Person ID) at enterprise level. The Person ID can be used for further linking with the social statistics, so-called Linked Employer-Employee Data (LEED). The register contains the link between approx. 350,000 enterprises and 3.4 million jobs.

Board members and managers
(Bestyrelsesmedlemmer og direktører)
NEW!
Contains information about industry and size of the enterprise and the gender, age, education and place of residence of each board member or manager. The statistics is based on statistical and administrative registers.
Read more in the documentation of statistics

Register of Enterprise owners
(Ejerregisteret) NEW!
Contains information about the industry and size of the enterprise and the gender, age, education and place of residence of the ultimate owner and the founder of the enterprise. The register combines business and social statistical and administrative registers.

How to gain access

The Data Warehouse for Business Statistics is part of the research infrastructure project Danish Research Data for Social Sciences (DRDS) co-funded under the Danish Research Infrastructure program with the purpose of establishing a data infrastructure utilizing the high quality statistical registers at Statistics Denmark (DST) for new and groundbreaking research. The Data Warehouse mainly holds registers covering the period from 2009.

Research Services at Statistics Denmark (Forskningsservice) provides access to researchers, ministries and companies wanting to analyse microdata based on projects that are deemed to be societally relevant.

Project requests are sent to Research Services via Danish Data Portal (Danmarks Datavindue, DDV). DDV offers a full overview of data available, including data outside DST. It should be noted that access to business data differs between public and private users. For more information, please contact Research Services.

Users of DSTs microdata services can choose to work on DSTs server, get a hosted server or work on a DST-dedicated part of a high performance computing center.

Go to Research Services