Cancels & replaces the same document of 6 June 2019

Use of business statistics – business statistics indicators for the 21st century

The 16th Meeting of the Committee on Statistics and Statistical Policy
24-25 June 2019
OECD Conference Centre, Paris

There is a growing demand for more granular data on businesses. Whilst current data on structural business statistics provide a relatively high degree of granularity that can respond to many needs, these cannot respond to all, as witnessed by the increasing demands for access to official micro-data, whether remotely (for example the co-ordinated access mechanisms used by OECD's Dynemp and MultiProd) or physically (through access at micro-data labs), and use of private data sources (such as BvD's ORBIS database).

Many of these additional demands, which create burdens for both compilers and users, and also challenges for comparability and quality (especially when non-official sources are used), could be met through smarter aggregations of existing official business statistics based on internationally agreed definitions that could be disseminated as a new suite of official statistics.

This paper describes (and proposes) a set of new internationally comparable indicators based on existing structural business statistics that could form the basis of a new set of official statistics and a new data collection by SDD.

CSSP delegates are asked to:
- Comment on the range of proposed indicators;
- Consider the possibility of a work-shop to identify a list of core indicators and to develop best practice in compilation.

This paper is for discussion under Agenda item 8.

It cancels and replaces the version issued on 6 June to correct a numbering problem.

Contact: Nadim AHMAD, Head of Trade and Competitiveness Statistics Division, SDD (nadim.ahmad@oecd.org)

Use of business statistics: business statistics indicators for the 21st century

1. Introduction

1. There is a growing demand for more granular data on businesses, witnessed by the increase in requests for access to micro data held within statistics offices and, also, by the explosion in the use of unofficial private data sources.

2. Although the regular collection of structural business (and related) statistics by the Statistics and Data Directorate provides an internationally comparable view of the structures of economies at a relatively detailed level of activity (4 digit ISIC/NACE) - also broken down by size class, with additional demographic indicators, such as births, deaths and survival - these data sets also struggle to meet the growing demands of users.

3. Contemporary releases of official business statistics have remained largely unchanged since their inception as inputs into the national accounts and as datasets that can be used to analyse structural changes in the economy. Underpinning their design, and indeed use, has been the view that the aggregation of firms around characteristics such as industry of classification or size of firm was largely, sufficient for analysing structural change and economic growth. These same characteristics have also been central to the construction of core accounting tools, such as supply-use tables, and remain the key elements typically used in survey design and stratification.

4. This view was shaped, at least in part, by the underlying assumption that for any given cohort of size and industry, all firms were by and large representative of their universe, i.e. they had similar production functions, similar productivity, similar exposure to markets (foreign and domestic), similar use of capital and so on.

5. Of course, these assumptions have never been true but until at least a few decades ago, before the current wave of globalisation (and in particular the explosion of global value chains), and a growing knowledge and digitally based economy, they were thought to be reasonably robust.

6. The increasing body of evidence that has emerged in recent years based on aggregations of firms around a broader set of characteristics reveals that this is no longer the case. We know for example that foreign affiliates of firms tend to be more capital intensive than their domestic counterparts, and have higher productivity and higher levels of integration in global value chains (Figure 1). We also know that, because they are part of a global production network, foreign affiliates also have greater scope to engage in fiscal optimisation.

Figure 1. Foreign-owned firms trade more than domestically-owned firms

Source: OECD Trade by Enterprise Characteristics

7.   In addition, within a given firm cohort based on industry and size, there is significant heterogeneity in performance across a wide range of measures – be that productivity levels, productivity growth, output growth, employment growth, carbon dioxide footprints and so on. (Figure 2) for example reveals significant shares of high-growth firms within the ICT sector across most countries pointing clearly to heterogeneity within sectors.
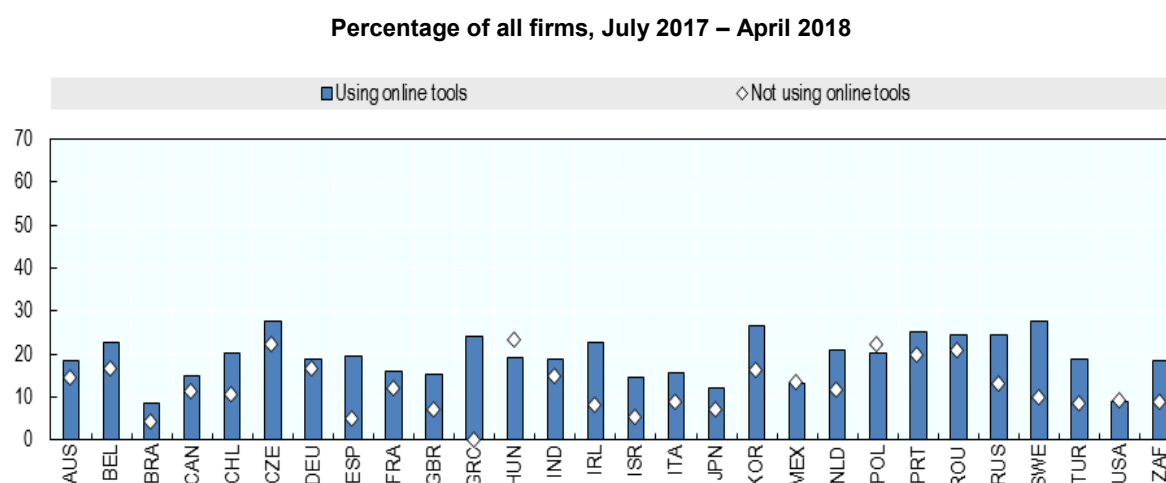
Figure 2. Share of high-growth firms in the ICT sector
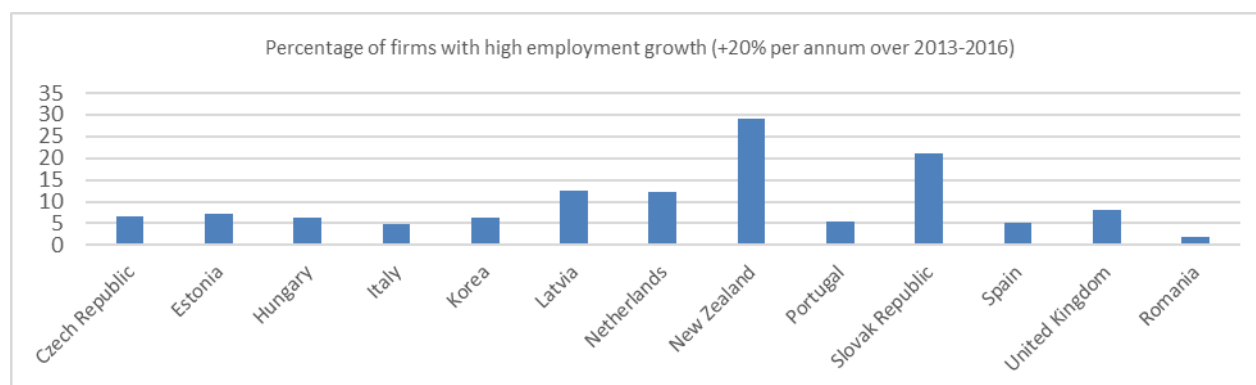
Source: OECD Business Demography Statistics

8.   And as illustrated by Figure 3, firms that make greater use of digital tools perform better and are better able to tap into and penetrate foreign markets.

Figure 3. Firms using online tools trade more internationally

Percentage of all firms, July 2017 – April 2018

Source: Facebook-OECD-World Bank Future of Business Survey.

9.   However, whilst there is a growing recognition that firm heterogeneity is not only a function of size and industrial classification, there has been little systematic change in the way we compile our business statistics nor indeed select stratification variables in business survey design.

10.   This is somewhat at odds with the accounting systems that depend on business statistics where many changes have been made in recent years to try to account for changes in firm behaviour, and so, in turn, heterogeneity. For example the 2008 SNA and BPM6 changes for goods for processing transactions were in part about capturing processing firms, whilst the Guide to Measuring Global Production[1] included significant commentary on factoryless producers. More recently this has also been picked up in the design of Extended Supply-Use Tables, which include aggregations for trading status and ownership structures. Moreover, many more changes are in the pipeline, for example aggregations of Digital Intermediation Platforms[2] in the Handbook on Measuring Digital Trade and in Digital Supply Use Tables[3].

11.   That being said, a number of countries have attempted to respond to calls for greater granularity through the creation of longitudinal and linked (e.g. employer-enterprise) datasets. The OECD itself has responded to user needs through the creation of distributed network approaches such as Dynemp[4] and MultiProd[5], generating a range of indicators not currently disseminated as official business statistics by national statistics offices.

12.   However, beyond these initiatives, and isolated examples, very little has changed in the core compilation and dissemination of official business statistics. Perhaps the only major coordinated initiatives launched over the last two decades have been the development

---

[1] UNECE Guide to Measuring Global Production, (2015)

[2] SDD/CSSP/WPTGS(2019)4

[3] SDD/CSSP/WPNA(2019)1

[4] https://www.oecd.org/sti/dynemp.htm

[5] https://www.oecd.org/sti/ind/multiprod.htm

of Business Demography Indicators and more recently Trade by Enterprise Characteristics (TEC) indicators.

13.   An important commonality of both of these datasets, and central to the proposal presented in this paper, is that both have been created through exploiting existing data. In the case of Business Demography Indicators, through <u>linking over time</u> to create a coherent view of individual firms over time, and in the case of TEC by <u>linking across datasets</u> (trade and business registers).

14.   In what follows, the paper presents and considers a series of potential new indicators that could be created (with varying degrees of difficulty/feasibility). In this respect, it is important to note that the proposals take as a starting point the existence of the necessary data in most countries (i.e. they don't imply higher respondent burdens).

15.   The goal, supported by a workshop, is to converge on a set of indicators that could form the basis of a new standardised dataset of business statistics. Selecting the indicators will need to consider a number of factors, not least burdens in national statistics offices but it is hoped that this could be marginal. However reductions in burdens may also be achieved, as the outcome may lead to fewer ad-hoc requests (e.g. through microdata labs) which add to burdens. Equally, by disseminating internationally comparable and frequently requested aggregations burdens for users could also fall and quality enhanced, especially for users currently relying on private sources of data (which often involve a monetary cost).

16.   The proposal therefore is to create an Informal Group, supported by a workshop, to identify a set of new business statistics indicators, and make a proposal for CSSP consideration.

2. New statistics drawn from conventional Structural Business Statistics

17.   Currently, the core dissemination of structural business statistics (SBS) is provided through the prism of size and industrial sector (typically at a highly disaggregated level – 4 digit ISIC category or equivalent), with information, typically on numbers of firms (enterprises/establishments), total turnover, value-added, gross operating surplus, output, compensation of employees, wages, employment (to name the most important variables). For any given aggregation (cohort) of firms these data provide a measure of the overall importance of that cohort but they also provide a measure of the average firm (statistical unit).

18.   Given the often significant heterogeneity of firms within these cohorts, additional statistical information that could be provided include **medians, variance/standard deviations, coefficients of variation, measures of skewness etc.**, but also, subject to confidentiality being preserved, more detailed information on the distribution of each of the cohorts – for example, by presenting breakdowns that show the average firm within deciles/quartiles etc. of firms within the cohort, or by including measures such as P25, P75 etc. in addition to the median.

19.   For countries that use censuses, this could, in theory, be achieved with few empirical challenges (notwithstanding the challenges of confidentiality presented by distributional information). However, for countries that use survey type approaches, there would necessarily need to be some reflection on how the information could be compiled. In the simplest case, this could be by presenting medians, variances/standard deviations etc. only for those firms included in the survey (i.e. not grossing up to the entire population for a given cohort). However some grossing mechanism would be needed for higher-level

aggregations (e.g. at the 2-digit level rather than the 4-digit level), especially for statistics on distributions (deciles/quartiles etc.).

20. Applying these new measures to composite statistics that are frequently derived from SBS data would also be better able to meet a number of user-demands, for example on:

- **Labour productivity;**

- **Value-added to output ratios;**

- **Value-added to turnover ratios;**

- **Gross operating surplus to turnover ratios.**

21. Certainly, especially given the focus on the recent productivity slowdown in most OECD countries and the fact that productivity gaps may have emerged between frontier firms and laggards, information on productivity distributions (even if only presented as annual snapshots, i.e. not fully longitudinal) could provide important insights on these trends. (Figure 4) below produced by ISTAT as part of a collaboration with the Statistics and Data Directorate illustrates the potential (and feasibility) for these new indicators.

Figure 4. Differences in Labour Productivity Growth and Skewness in Levels (Italy, 2012-2016)
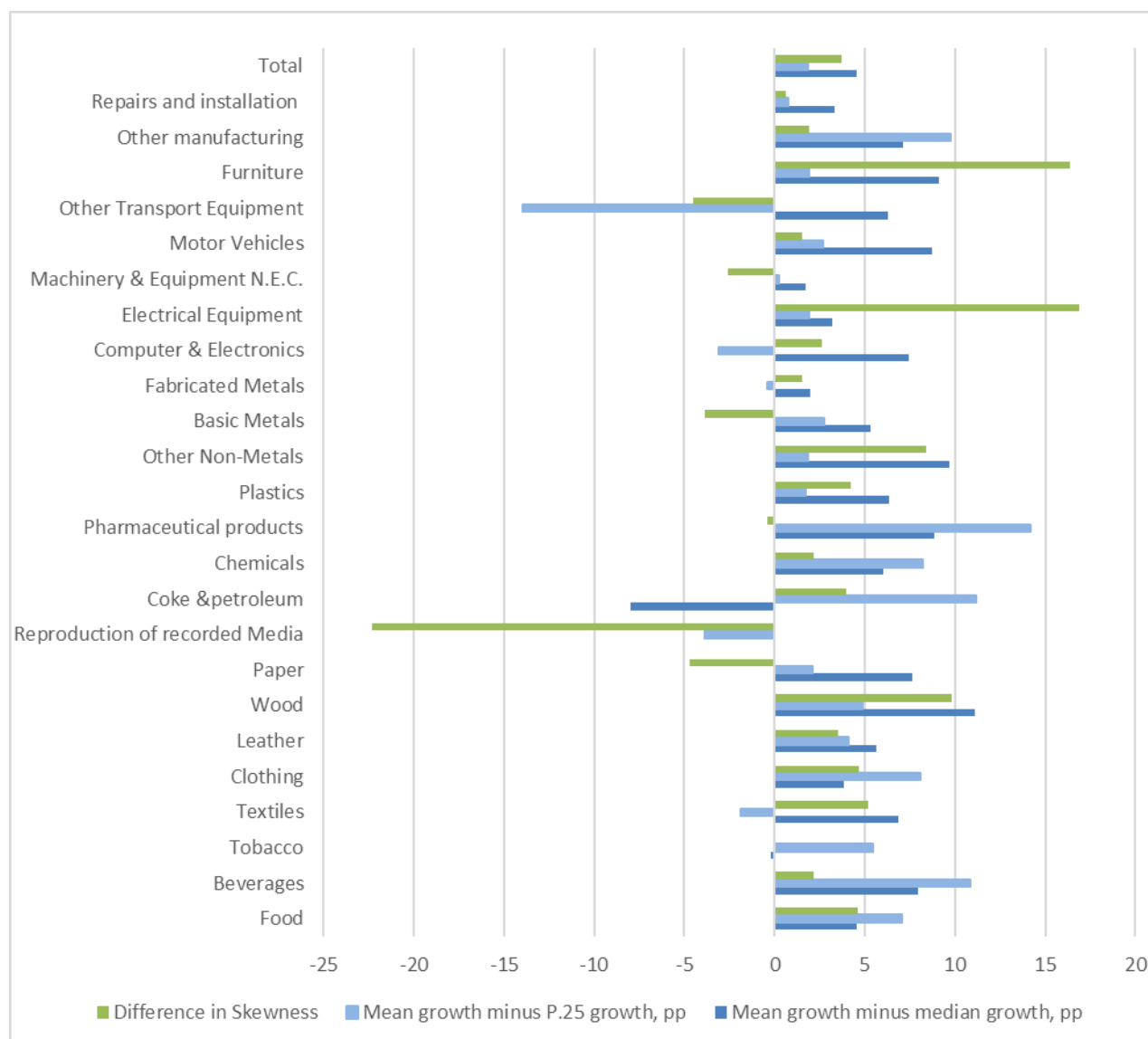
Source: Based on ISTAT data

22.   (Figure 4) reveals that whilst average productivity growth within a sector outpaced the growth of firms with median productivity in nearly all sectors over the period 2012-2016, this cannot always be interpreted as a sign of growing gaps between frontier firms and laggards. In the other transport equipment sector for example the productivity growth of the average firm outpaced that of the median firm by 6 percentage points but P.25 firms (on the 25% quartile) saw productivity grow 14 percentage points higher than the average, and overall measures of skewness in labour productivity levels in the sector fell from 7.7 in 2012 to 3.2 in 2016, pointing to smaller gaps between frontier firms and laggards. Similarly whilst productivity gaps increased between the average firm and the median and P.25 firms in the pharmaceuticals sector, measures of skewness declined.

23.   In addition, to respond to growing concerns around winners taking all/most, simple measures of **concentration,** such as the Herfindahl-Hirschman index[6] or shares of total turnover, value-added, employment, profits generated by the top 10% or top 10, 20, 30% etc. of firms could be produced.

3. New aggregations of Structural Business Statistics

24.   New aggregations of firms around firm-characteristics, such as ownership structures, could also help to meet user needs. Efforts are now underway in different fora (such as in the CSSP Working Group on Extended Supply-Use Tables) to develop these aggregations. Mainstreaming their compilation within the framework of SBS could help to solidify these efforts, whilst also providing scope for new statistical releases.

25.   There are a number of options that could be explored in this area (noting that it will remain important to retain some aspect of the current core characteristics – size and industry). Some of these reflect core firm-based characteristics that could be captured in most conventional statistical business registers (and in turn help to improve the quality of business survey design) such as aggregations by:

- **Ownership structures** (foreign owned, domestically owned with no affiliates abroad, domestically owned with affiliates abroad);

- **Dependencies**: Enterprises/Establishments that are part of larger groups (e.g. independent vs dependent SMEs);

- **Age of firms.**

26.   Other options that could be explored, albeit with higher levels of complexity as the characteristics are more dynamic and may also require some form of data-linking,   include the following:

- **Trading Status:** (exporting, non-exporting, two-way trader);

- **Business model**: Processing firm; Factoryless Producer; Digital Intermediation Platform; E-tailer; Firms broken down by number of customers (e.g. one vs many) to provide new insights on control relationships and the degree of resilience to shocks firms may have

- **Capital Intensity:** (e.g. with breakdowns by level of capital intensities - capital stock to output ratios/capital stock to value-added ratios – e.g. high/medium/low);

- **Skills intensity**: Breaking firms down into levels or distributions of skills intensity (e.g. shares of high-skilled/medium skilled/low skilled employees; quartiles of intensities etc.);

- **Productivity levels** (when combined with other firm characteristics such as, trading status, ownership etc.).

---

[6] The Herfindahl-Hirschman Index (HHI), often used by competition authorities and regulators, is calculated by squaring the market share of each firm in a market and then summing the resulting numbers (bounded between zero and 10,000).

4. New variables – through data linking - included in Structural Business Statistics

27.   New variables could also be included through data linking with:

- **Tax data**: to provide insights on effective tax rates of different categories of firms (including for example by ownership structures);

- **Trade intensities:** e.g. share of output/turnover destined for export.

28.   Falling under this umbrella is the suite of variables identified by the CSSP Informal Advisory Group on Measuring GDP in the Digital Economy, and in the Handbook on Measuring Digital Trade, which advocate breakdowns of transactions by their digital nature i.e. whether they were digitally ordered and/or digitally delivered. Through linking SBS to e-commerce surveys (in the case of digital ordering), SBS data could be enhanced to reveal the share of turnover that was:

- **digitally ordered;**

- **digitally delivered;**

- **digitally delivered and ordered;**

- **digitally ordered** through the firm's own website

- **digitally ordered** via an intermediation platforms

5. New aggregations of Business Demography Statistics

29.   Business Demography (BD) data dissemination is currently grounded on the same firm-level characteristics that govern SBS (i.e. size and industry). Similar new aggregations to those proposed above could also be considered for inclusion in BD statistics. Perhaps the most relevant include aggregations by:

- **Ownership:** providing a means to differentiate between, for example, start-ups, high-growth firms, gazelles etc. that are foreign owned (where the underlying financing model, level of dependence, and level of entrepreneurship typically differ from pure *in vitro* start-ups) from domestically owned firms.

- **Dependencies**: where the ability to differentiate between dependent and independent firms would bring similar benefits to those described above.

- **Trading status:** (importer, exporter, two-way trader), providing a means to better understand 'born-global' firms but also the importance of trade channels for high-growth and survival rates.

6. New variables in Business Demography Statistics

30.   Considerable attention has been focused in recent years, for example through efforts such as the OECD's Dynemp exercise, to extract more granular insights on employment and turnover dynamics of firms, especially broken down by firm age.

31.   Wider international take-up of the most important indicators in these efforts could be helped through their inclusion in official disseminations of Business Demography Statistics. Core indicators that could be included are:

- **Average turnover/employment growth** between two periods of time (t and t-1, t and t-5 etc.) for distinct cohorts of firms (new entries, expanding, contracting , stable) broken down by industry and size class but also including a new characteristic **of age**; helping to provide important insights on, among others, the prevalence of Zombie Firms[7].

- **Distributions of turnover/employment growth** between two periods of time (t and t-1, t and t-5 etc.), as a full distribution or through percentiles.

- **Number of jobs created/destroyed** between two periods of time (t and t-1, t and t-5 etc.) for distinct cohorts (new entries, existing, expanding, contracting) broken down by industry and size class but also by **age**.

- **High-growth enterprises and Gazelles by age.**

- **Scale ups**: There are a number of efforts underway, notably in Nordic Economies[8], and now also given high priority within the EC, for new indicators on Scale-Ups, which share characteristics with indicators on high-growth enterprises (Figure 5).

Figure 5. Number of scale-up enterprises in the non-financial business economy in the Nordic countries
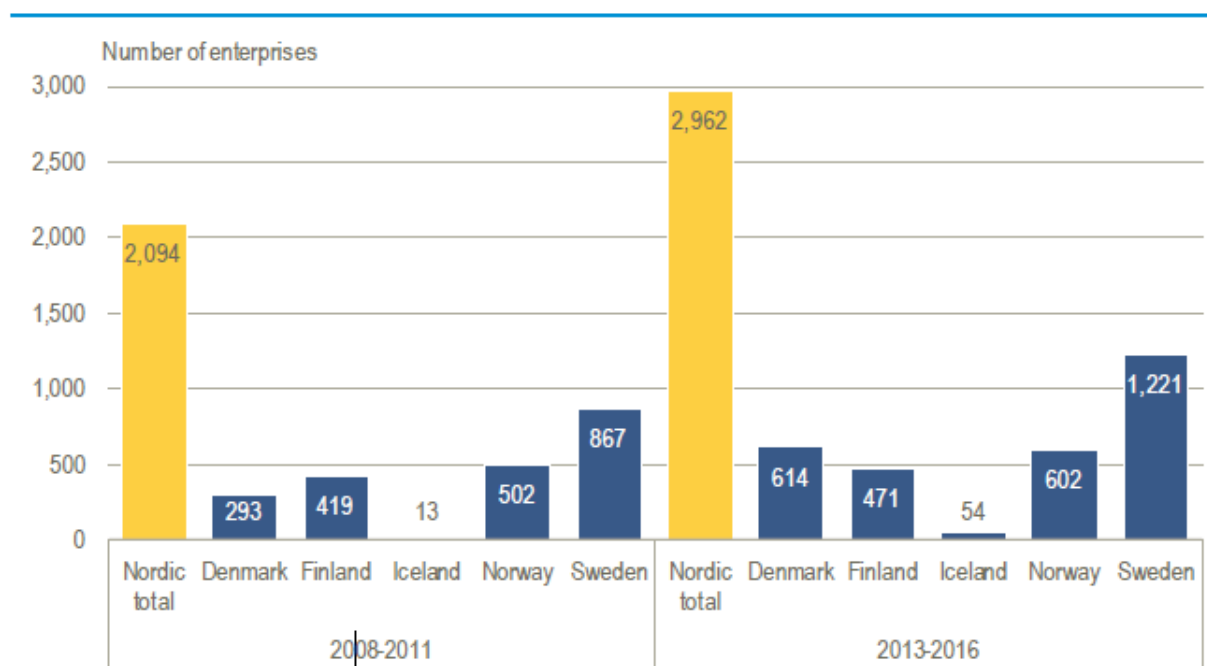
Note: Figures reflect the number of scale-ups at the end of the three year period
Source: Scale-ups in the Nordics − Statistical Portrait 2008-2016 (2019).

---

[7] The Walking Dead? Zombie Firms and Productivity Performance in OECD Countries, OECD Economics Department Working Paper, 1372 (2017)

[8] Scale-ups in the Nordics – Statistical Portrait 2008-2016 (2019).

7. New variables through longitudinal Structural Business Statistics

32.   A growing number of countries are now producing SBS datasets that link firms over time (often through links to other datasets, for example employment data, to create Linked Employer-Employee Databases, LEED). Whilst these are being increasingly exploited as tools by analysts (including within the OECD, for example through MultiProd) they remain largely under-utilised as vehicles for delivering new standardised official indicators, including:

- **Distributions of growth** (over two periods of time t and t-1, t and t-5 etc.) in:
  - o **Labour productivity** – providing important insights on the breakdown in diffusion narrative[9].
  - o **Wages**
  - o **Turnover** (see also above under BD statistics)
  - o **Output**
  - o **Value-added**
  - o **Employment** (see also above under BD statistics**)**

- **Progressions across cohorts:** Number of firms moving from and into new size classes (and other groupings, e.g. from domestic to foreign owned etc.) over two periods.

33.   In addition to the above, there is a range of additional datasets that could also be linked to SBS data including for example on occupations, educational attainment etc., and other data sources such as R&D and ICT intensities.

8. Moving Forward

34.   It is not expected that all of the indicators described above can be included in any future national releases and not only because of burdens. The issue of confidentiality also needs to be carefully considered, especially as the higher the number of indicators based on additional firm-level characteristics, the greater the risk of secondary disclosure.

35.   In this respect, subject to CSSP approval, the next steps would be for the OECD Secretariat to develop a detailed template/questionnaire that could be circulated to countries who would be asked to explore the feasibility of developing experimental statistics in all **or only some** of the fields, described above.

36.   This exercise could form the basis of a two-three day workshop hosted at the OECD to explore feasibility and to flesh out challenges and solutions in implementation, with a view to sharing national experiences (and identifying best practice). In turn, the workshop would help to determine which indicators could/should be in scope for any potential new standardised questionnaire and national dissemination strategies. The workshop would also be an opportunity for countries to share national practices in the construction of linked datasets, and to consider potential implications for future business survey design.

37.   **If this proposal is approved, the Secretariat would move to develop an experimental Questionnaire that could be circulated in Q3/Q4 of this year with a view to organising a workshop in 2020, and preparing a set of conclusions/recommendations for CSSP's consideration.**

---

[9] The Future of Productivity, OECD (2016)