UNECE

Guidelines on the use of statistical business registers for business demography and entrepreneurship statistics
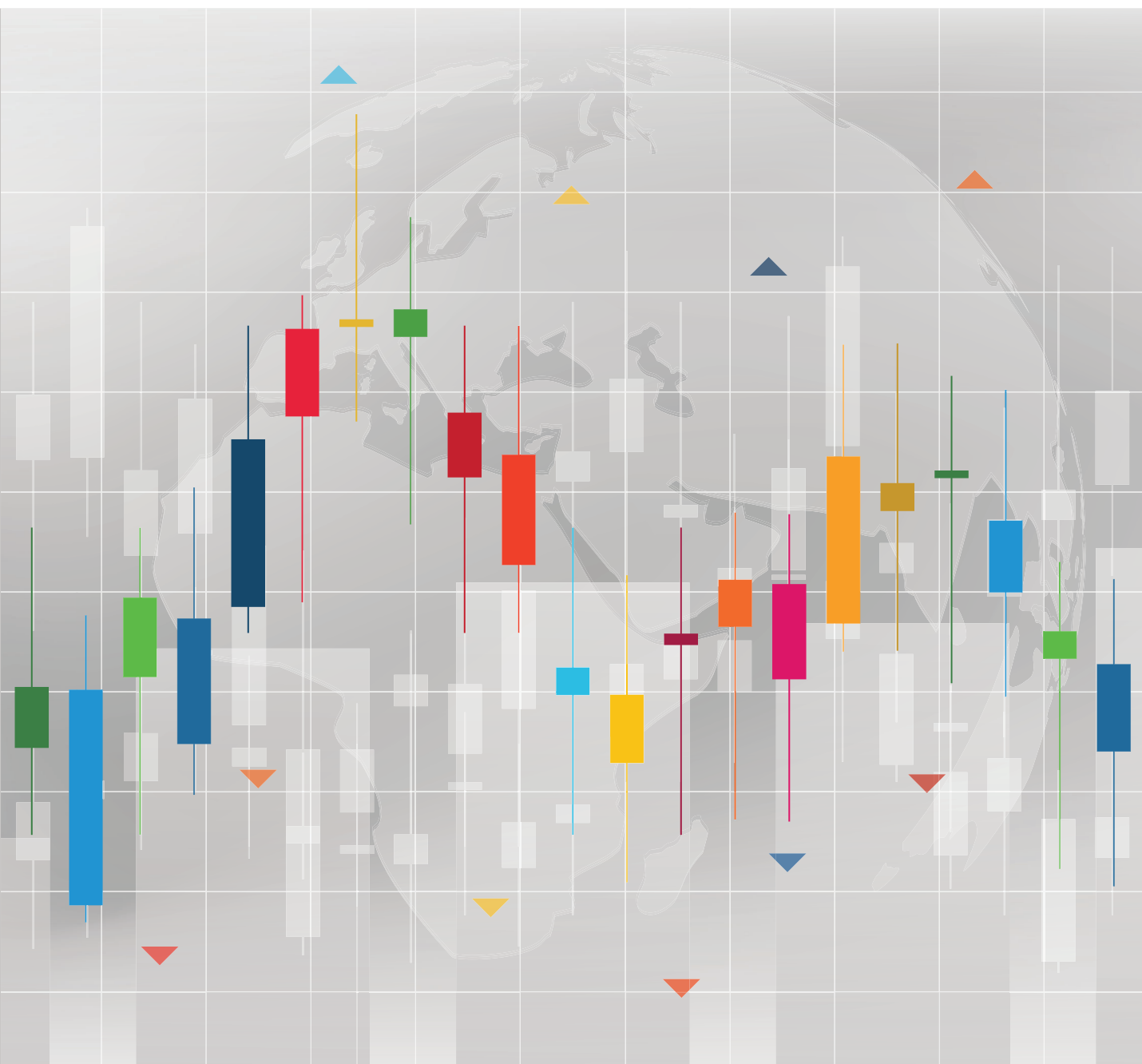
UNITED NATIONS

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

Guidelines on the
use of statistical business registers for business demography and entrepreneurship statistics

United Nations

New York and Geneva, 2018

NOTE

The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning delimitation of its frontiers or boundaries.

This volume is issued in English only.

Preface

Business demography and entrepreneurship statistics provide information about the creation, survival and dissolution of businesses and about the characteristics and activities of entrepreneurs. Over the previous decades, interest in these types of statistics has increased as they are sources of information about innovation, competitiveness, economic growth and job creation. Such information is used to inform business policies, for economic planning, and for analytical and research purposes.

The production of business demography and entrepreneurship statistics of suitable quality requires good data sources. The statistical business register, which plays a central role in the production of economic statistics, can provide the statistical infrastructure for producing business demography and entrepreneurship statistics in the form of structured data on business units and their characteristics and activities. The statistical business register can be the sole source of information or may be combined with data from other sources. Use of the statistical business register in addition to, or instead of, other sources, provides an opportunity to reduce costs, increase efficiency and improve the quality of the resulting statistics.

The Guidelines provide practical guidance on how to develop and maintain statistical business registers to support the production of business demography and entrepreneurship statistics. The Guidelines contain definitions and compilation guidance for key demographic events and propose business demography indicators that can be produced directly from the register. They also describe the requirements for the statistical business register to produce business demography indicators and give guidance on linking the register with other data sources.

The Guidelines were developed by a UNECE Task Force established by the Bureau of the Conference of European Statisticians. It is hoped that statistical offices will find the Guidelines useful.

Background and acknowledgements

In 2015, the Conference of European Statisticians (CES) endorsed the UNECE Guidelines on Statistical Business Registers (SBRs), which provide practical guidance and recommendations on the establishment and maintenance of SBRs. The CES recognized the central role of SBRs in the production of economic statistics and the potential for using SBRs to produce business demography and entrepreneurship statistics. The CES also recognized a growing interest in entrepreneurship statistics as an area where new statistics could be developed based on the SBR and other data sets that national statistical offices have already assembled.

The growing importance of entrepreneurship statistics was first noted by the Bureau of the CES in 2013, when the future work on entrepreneurship statistics was discussed, based on an in-depth review prepared by OECD and Eurostat. The review highlighted areas where further work was needed to improve the production of business demography and entrepreneurship statistics, for example, by developing linkages between datasets, and by using longitudinal data sets.

With this background, in July 2016, the Bureau of the CES established a Task Force to develop guidelines for producing business demography and entrepreneurship statistics. Following the terms of reference of the Task Force, the Guidelines describe the statistical infrastructure required to support the production of business demography and entrepreneurship statistics. More specifically, they provide:

a) Guidance on how to develop SBRs to support the production of business demography and entrepreneurships statistics, including:

   - Definitions and compilation guidelines for key demographic events.
   - Suggestions of demographic statistics that could be produced directly from SBR.
   - Recommendations on the requirements for the SBR to produce business demography information. longitudinally, annually and sub-annually, and by region within country.

b) Guidance on linking the SBR with information from other data sources to produce business demography and entrepreneurship statistics.

The Guidelines take into account existing relevant material, in particular the *Guidelines on Statistical Business Registers* (UNECE, 2015), the *Manual on Business Demography Statistics* (Eurostat & OECD, 2007) and the *Business Registers Recommendations Manual* (Eurostat, 2010) and make use of the definitions contained in these documents. The target audiences for the Guidelines are SBR statisticians in national statistics offices, experts in the production of business demography and entrepreneurship statistics, and users of these statistics.

During the drafting the Guidelines over the period 2016 to 2018, the Task Force worked mainly through the exchange of e-mails and audio conferences. The Task Force had also two face-to–face meetings. Drafts of chapters and relevant materials were shared on a common, designated UNECE wiki page. During the first half-year, the Task Force agreed on the time plan of activities and worked on a detailed outline for each chapter of the Guidelines. The structure and content of the draft chapters was discussed in detail during the first face-to–face meeting that took place in Geneva on 13-14 June 2017.

An outline of the draft Guidelines was presented to a broader audience over the period 27-29 September 2017 during the joint UNECE, Eurostat and OECD biennial meeting of the Group of Experts on Business Registers. Meeting participants were given the opportunity to submit written comments and suggestions on the draft chapters. The comments received were considered in subsequent versions of the Guidelines.

The second face-to-face meeting of the Task Force took place on 27 September 2017 in Paris, back-to-back with the meeting of the Group of Experts on Business Registers. This meeting discussed the draft

chapters and the time plan for remaining work. From September 2017 to January 2018, the Task Force continued to work on the draft chapters and clarify outstanding issues. A draft version of the Guidelines was finalised in March 2018 and circulated to all members of the CES for written consultation. The Guidelines were endorsed by the CES plenary session in June 2018. Comments and suggestions received from CES member countries and organisations were incorporated in the Guidelines. The final editing and formatting of the Guidelines took place from July to November 2018.

Acknowledgements

Contents

Chapter 1: Introduction

1.1   Importance of measuring entrepreneurship

1.1   The role of entrepreneurs in stimulating economic growth has received renewed attention in the past two decades, when the extraordinary development of information and communication technologies and other advanced technologies fostered new breeds of start-ups, as well as a new attitude toward entrepreneurial activity. *Entrepreneurs*, in these Guidelines, are meant as persons who own an enterprise and seek to generate value through the creation or expansion of economic activity, by identifying and exploiting new products, processes or markets. *Entrepreneurship* indicates the phenomena associated with entrepreneurial activity.

1.2   The interest in entrepreneurship was further stimulated by the economic recession that followed the global financial crisis in 2007/2008. The notion that entrepreneurship is a crucial driver of economic growth has become part of the policy narrative on how to ensure growth and prosperity across the world, though the theoretical foundations for such an assumption are, to some extent, still under development. While the work of Joseph A. Schumpeter (Schumpeter, 1912) at the beginning of the 20th century had already identified entrepreneurship as a driving force for innovation and an engine for economic development, it was only in the past two decades that researchers formally proposed considering entrepreneurship as an additional factor in explaining economic growth. The assumption is that entrepreneurship is distinct from human capital and from research and development (R&D), the two drivers of growth already put forward by endogenous growth theories, and that it constitutes the missing link between investment in new knowledge and economic growth. Entrepreneurs, via business creation, exploit opportunities provided by new knowledge and ideas that have not yet been discovered or commercialised by incumbent firms.

1.3   A second mechanism is also at work. The process of "creative destruction" first described by Schumpeter is itself a driver of economic growth. New firms entering the market displace obsolete firms, and the business dynamics of entry and exit contribute to productivity dynamics and eventually to economic growth.

1.4   Against this backdrop, the literature on entrepreneurship also acknowledges that entrepreneurs are a heterogeneous group, encompassing very different types. Innovative "Schumpeterian" entrepreneurs, who exploit market opportunities or innovative ideas, coexist with "necessity" entrepreneurs (also called "survival" entrepreneurs) who start a business out of necessity as they do not have another means of generating income. While necessity entrepreneurs may eventually become successful, not all entrepreneurial activities have the same, positive impact on productivity and growth. Entrepreneurship is conducive to knowledge creation and diffusion, and economic growth, depending on how and where it occurs, for example in which sectors and/or locations.

1.5   The impact of entrepreneurship has also been associated with poverty reduction and social inclusion. Stimulating entrepreneurial activities among individuals who belong to disadvantaged or marginalised groups (because of their age, gender, ethnic characteristics, lack of education and skills, or the geographical area where they live) can be an important vehicle for social inclusion, and it can contribute to poverty reduction by bringing into a country's active labour force segments of the population previously excluded.

1.6   To advance understanding of the role of entrepreneurship in economic growth and social inclusion, sound evidence on the entrepreneurial phenomenon, and its determinants and impacts is critical. However, empirical research on entrepreneurship continues to rely to a large extent on private

data sources and much less on official statistics[1]. Indeed, suitable and comprehensive data for the analysis of entrepreneurship are not always found in national statistical offices (NSOs). The main challenge is that entrepreneurship data span a multitude of domains. They cross the traditional boundary between economic and social statistics by encompassing topics such as the demography and performance of businesses, the profiles of individuals who create enterprises, the attitudes toward entrepreneurship, and the regulatory environment for setting up new businesses.

1.7   The Guidelines focus on economic statistics and describe concepts and methods to produce business demography statistics to complement the wealth of business data already produced by NSOs, in particular structural business statistics and foreign affiliate statistics (FATS). Structural business statistics have existed for many years and describe the structural characteristics of the business population, for example, business counts, employment, value added, and turnover. However, they cannot adequately account for the entrepreneurial dynamism as reflected by the creation of new businesses and their survival, growth or death. Business demography statistics, on the other hand, capture these fundamental aspects of the phenomenon of entrepreneurship. Over time, the demography of a business population follows the populations used to produce statistics such as structural business statistics and FATS. In this way, the developments in business demography statistics can be linked to the developments described in economic statistics.

1.8   The Guidelines are meant to assist NSOs and other statistical institutions in developing the production infrastructure, in particular the statistical business register (SBR), for compiling statistics on business demography and entrepreneurship by linking business demography data with other business and social statistics data. While the primary purpose of the Guidelines is to provide advice on how to develop the SBR to facilitate the production of business demography statistics and related statistics, the Guidelines acknowledge that there are other approaches. Not having a fully developed SBR should not be considered an insuperable obstacle to production of business demography statistics.

1.2   Statistics for measuring entrepreneurship

1.9   The topic of entrepreneurship was included as a statistical area in the Classification of Statistical Activities (CSA) only recently, in 2010.[2] Entrepreneurship is categorised under *Domain 3: Environment and multi-domain statistics*. Entrepreneurship statistics (Statistical Area 3.3.7) cover "the measurement of the determinants, performance and impact of entrepreneurial activities of people and organizations". This classification acknowledges that entrepreneurship statistics are a multi-domain statistical area. For example, data on self-employment and business demography on the one hand, and data on access to finance for new and young businesses on the other, are both key statistics supporting the study of entrepreneurial phenomena. The classification has been inspired by the conceptual framework developed by the OECD-Eurostat Entrepreneurship Indicators Programme[3], which characterises as separate categories the *indicators of entrepreneurial determinants, performance and impact*. As shown

---

[1] At the international level, three collections of entrepreneurship data developed outside NSOs have gained popularity among analysts: the *Global Entrepreneurship Monitor (GEM)*, the *Eurobarometer on Entrepreneurship* of the European Commission, and the World Bank's *Entrepreneurship Databases*. The first two collections rely on ad-hoc surveys of individuals, while the third mainly consists of administrative data on business registrations.

[2] The Classification of Statistical Activities is an international standard for describing official statistical work by domain. See https://unstats.un.org/unsd/iiss/Classification-of-International-Statistical-Activities.ashx

[3] The OECD-Eurostat Entrepreneurship Indicators Programme (EIP), launched in 2006, develops policy-relevant and internationally-comparable indicators. The Programme produced methodological tools to structure the development and collection of indicators of entrepreneurship: the Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007), and the Working Papers "Defining Entrepreneurial Activity: Definitions Supporting Frameworks for Data Collection" (Ahmad & Seymour, 2008) and "A Framework for Addressing and Measuring Entrepreneurship" (Ahmad & Hoffmann, 2008)

in Figure 1.1, *entrepreneurial performance* refers to the manifestation of entrepreneurial activity, e.g., the creation of new businesses, their growth and survival, or the innovative ability of start-up and young firms.

Figure 1.1 **Conceptual framework of the OECD-Eurostat Entrepreneurship Indicators Programme**

Determinants						Entrepreneurial performance	Impact
Regulatory framework	Market conditions	Access to finance	Knowledge creation and diffusion	Entrepreneurial capabilities	Culture	Firm based	Job creation
Administrative burdens for entry	Anti-trust laws	Access to debt financing	R&D investment	Training and experience of entrepreneurs	Risk attitude in society	Employment based	Economic growth
Administrative burdens for growth	Competition	Business angels	University/industry interface	Business and entrepreneurship education (skills)	Attitudes towards entrepreneurs	Wealth	Poverty reduction
Bankruptcy regulation	Access to the domestic market	Venture Capital	Technological co-operation between firms	Entrepreneurship infrastructure	Desire for business ownership		Formalising the informal sector
Safety, health and environmental regulations	Access to foreign markets	Access to other types of equity	Technology diffusion	Immigration	Entrepreneurship education (mindset)		
Product legislation	Degree of public involvement	Stock markets	Broadband access				
Labour market regulation	Public procurement						
Court and legal framework							
Social and health security							
Income taxes: wealth/bequest taxes							
Business and capital taxes	Patent system standards						

Firms	Employment	Wealth
Enterprise birth rates	Share of high growth firms (by employment)	Share of high growth firms (by turnover)
Enterprise death rates	Share of gazelles (by employment)	Share of gazelles (by turnover)
Business churn	Ownership rate start-ups	Value added, young or small firms
Net business population growth	Ownership rates business population	Productivity contribution, young or small firms
Survival rates	Employment in young firms	Innovation performance, young or small firms
Proportion of young firms	Average firm size after x years of survival	Export performance, young or small firms

Source: OECD (2017)

1.10   The Guidelines are concerned with, and provide guidance for, the production of indicators of entrepreneurial performance based on business data, notably the two types of statistics described below:

a.   *Statistics that focus on businesses*. These primarily consist of business demography statistics measuring the births, deaths, survival and growth of businesses as well as the employment changes associated with these events. Also, by linking business demography data with other business statistics, important indicators can be generated to describe the performance of new or young enterprises versus that of mature enterprises in domains such as trade or innovation. Examples of linked business data are statistics on exports and innovation activities of young businesses.

b.   *Statistics profiling businesses jointly with the business owners*. These measures of entrepreneurship rely on statistics linking data on the individuals who are business owners with data on the businesses they own. Such statistics provide the profile of an entrepreneur together with the performance of his/her business. Progress in this area has been slower reflecting the greater compilation challenges involved. Two examples are (1) statistics on business births and

deaths by gender of the business owner, and (2) innovation activity of young firms by educational attainment and employment history of the business owner.

1.11   The Guidelines do not cover a third category of measures of entrepreneurial performance that is characterised by having the *individual* (actual or potential entrepreneur) and not the *enterprise*, as the statistical unit. Entrepreneurship indicators on individuals include statistics on self-employment from labour force surveys and population censuses, and statistics from ad-hoc surveys of individuals on attitudes toward entrepreneurship and involvement in entrepreneurial activities. Examples are: statistics on self-employed, disaggregated by the gender and age of the self-employed; individuals' preference for self-employment over wage employment; attitudes toward the risk of business failure; and perceptions on the social and economic role of entrepreneurs. The limitations of these measures of entrepreneurial performance are discussed below.

Entrepreneurship and self-employment

1.12   The approach to measuring entrepreneurial activity has evolved significantly in the past two decades. Until the late 1990s, self-employment was the indicator most frequently used as a proxy for entrepreneurship. In virtually all countries, information on self-employment is collected on a regular basis by labour force surveys and population censuses. This makes use of self-employment data an easy solution for analysts in search of quantitative evidence on entrepreneurship.

1.13   In the 2008 System of National Accounts (2008 SNA), self-employed persons are defined as persons who are the sole or joint owners of the unincorporated enterprises in which they work (2008 SNA, paragraph 19.25). This definition excludes owners of incorporated enterprises, and unincorporated quasi-corporations. The exclusion of this large group of owner-operators of incorporated enterprises and quasi-corporations is a major problem in statistics on entrepreneurship and for a realistic estimation of the number of entrepreneurs. Also, not all the self-employed are in fact entrepreneurs; this is for instance the case of *unpaid contributing family workers*. In addition, some individuals might have a self-employment status for reasons of convenience or – a rising phenomenon - might be *gig workers*. This term describes workers engaged in flexible employment arrangements, or *gigs*. While certain professions, notably in the entertainment industry, have always relied on gigs as an important source of income, the rapid popularisation of gigs has been fuelled by technology and is largely associated with the rise of online platforms such as Uber and TaskRabbit that connect buyers and sellers for one-time transactions. Many gig workers use online platforms to find small jobs, sometimes completed immediately after request (essentially, on-demand), in the context of a contractual relation where any notion of entrepreneurial risk for the gig worker is absent (OECD, 2017).

1.14   The resolution concerning statistics on work relationships adopted by the 20th International Conference of Labour Statisticians (ICLS) in October 2018 addresses the limitations of the SNA-based definition for the identification of entrepreneurs. The resolution provides a revised version of the International Classification of Status in Employment (ICSE-18) that replaces the previous version of the classification from 1993. ICSE-18 moves away from the term self-employed due to the term being interpreted differently in SNA and labour market statistics and introduces the category of *independent workers*, who are those persons who own and control an economic unit for which they work, whether it is incorporated or not. The category of *independent workers* in the classification of status in employment would provide the best starting point for the identification and compilation of statistics on entrepreneurs (see ILO, 2018).

1.15   Independently of the problems of measuring entrepreneurs, there are limitations in the ability of self-employment data to represent entrepreneurship in a comprehensive way. Self-employment data or, in future, data on independent workers, can provide information on the population of entrepreneurs, for example, their number, their characteristics in terms of gender, age, education or work experience. They cannot, however, offer any insight into the performance of the businesses being created, for example, their survival or growth. Considering these limitations, over the past decade the international statistical community has made considerable efforts to conceptualise and develop additional indicators

for measuring entrepreneurial activity, in particular indicators that concentrate on businesses as the observation unit rather than individuals.

1.3 Business demography and international comparability

1.16 Business demography statistics refer to statistics on "events like births and other creations of units, deaths and other cessations of units, and their ratios to the business population. This includes following units over time, thus gaining information on their survival or discontinuity. It also covers development over time of certain characteristics, like size, thus gaining information on the growth of individual units, or a cohort of units, by type of activity" (Eurostat & OECD, 2007). In the economic literature the term "*business dynamics*" is used to refer to the analysis of the demography of businesses and their impact on employment and productivity.[4] In these Guidelines the term business demography is used.

1.17 Business demography statistics can be compiled from a range of data sources, in particular SBRs, business censuses and business surveys. Research conducted in the mid-2000s on how to improve cross-country comparability of data on the creation of new businesses identified the following advantages and challenges associated with these different sources of data on business demography:[5]

- SBRs usually provide comprehensive coverage of the population of interest and are a reliable source. In fact, data from a comprehensive, frequently updated SBR are likely to be more reliable than those from small scale surveys or studies on enterprise creation. However, the scope of the SBR, and of specific threshold restrictions, may involve exclusions from the population coverage that are different from those desirable from a business demography perspective.

- Census data can be as good as SBR data, and sometimes better if they have less scope restrictions. However, the cost of running a census of businesses every year makes this approach unrealistic in any country. Data from less frequent censuses can be used, but the statistics become out of date, and the comparability of statistics across countries is difficult when their periodicities differ.

- Survey data have also been used by some countries, most notably in the Eurostat project "*Demography of Small and Medium-sized Enterprises (DOSME)*" conducted in twelve countries of Central and Eastern Europe at the time of their transition to the market economy. This approach is useful when SBRs are not sufficiently developed. It allows collection of more information on the profiles of entrepreneurs than is available from other sources. It allows collection of information on the informal economy. However, it suffers from the usual constraints set by sample size when detailed data breakdowns are required. Also, it is not possible to identify business deaths through surveys.[6]

1.18 The advantages and challenges listed above point to the use of SBRs as a most convenient source for compiling internationally comparable business demography statistics, especially if the SBRs are already subject to some harmonisation, i.e., if basic requirements of the SBRs are harmonised across countries. Today, an increasing number of countries compile business demography statistics on a regular basis, although with differences in the scope of the data collection, and mostly based on country specific approaches. Box 1.1 provides an overview of country practices that highlights this.

---

[4] Business scholars use the same term to indicate modelling methods to test the effectiveness of different policies on business outcomes and for the analysis of business strategy; for example, see Sterman (2000).

[5] Vale (2006)

[6] UNECE (2014)

Box 1.1 **Overview of country practices**

In 2013, the United Nations Statistics Division (UNSD) conducted an "SBR Global Assessment Questionnaire Survey" aimed at collecting information on the status of the business register in all countries. The Global Survey Questionnaire was applicable not only to NSOs that operate a comprehensive SBR but also to those that maintain and update one or more lists of enterprises, economic census frames, or annual enterprise surveys. The list(s) of enterprises could be compiled based on multiple surveys or could be built by combining survey data with administrative data. For the purposes of the survey, all such lists were referred to as a "business register".

The analysis of the questionnaire findings highlighted the following.

- In most countries, compilation of business demography statistics from SBRs started in the early 2000s or more recently. France, where the production dates to the 1970s, is an exception.

- Business demography statistics produced by countries may consist simply of data on registrations and de-registrations of firms with the relevant authority (as in Singapore, for example).

- On the other hand, business demography statistics may be the result of implementing a methodology for identifying and handling demographic events. This is the case in OECD countries, EU member states and other developed and emerging economies where:

  i. a *country specific* methodology is implemented, for instance, in Australia, Costa Rica, and Tunisia;

  ii. an *internationally harmonized* methodology is implemented, typically based on the *Eurostat-OECD Manual on Business Demography Statistics* (Eurostat & OECD, 2007), as in EU member states, Brazil, Canada, Israel, New Zealand, Norway and Switzerland.

It is worth noting that since the 2013 survey there have been developments in some countries; for instance, at the time of the survey, the Russian Federation had a country specific methodology but since 2017 has implemented the approach recommended by the Eurostat-OECD Manual (2007).

1.19   In contrast to a country specific approach is the coordinated effort conducted by the countries within the framework of the OECD-Eurostat Entrepreneurship Indicators Programme (EIP), launched in 2007. Countries that are member states of the European Union (EU) are required to comply with mandatory regulations and to produce business demography statistics following the methodology and definitions recommended by the *Eurostat-OECD Manual on Business Demography Statistics* (Eurostat & OECD, 2007). As a result, Eurostat and the OECD have created databases of internationally comparable business demography statistics that were previously missing from official statistics.

1.4   Aims of the Guidelines

1.20   The main purpose of the Guidelines is to support the development of the SBR as the statistical production infrastructure for the compilation of business demography statistics and related indicators of entrepreneurship. Notably they address the requirements for the SBR to support the production of business demography statistics. The Guidelines build on the *Guidelines on Statistical Business Registers* (UNECE, 2015) and take these forward to provide guidance on developing a longitudinal SBR. They also expand the methodological explanations and provide illustrative examples of country practices for compilation of business demography statistics from SBRs.

1.21   The Guidelines also build on the *Eurostat-OECD Manual on Business Demography Statistics* (Eurostat & OECD, 2007), which provides definitions of business demographic events (to serve as recommended international standards) and proposes a methodology to operationalise the definitions

and compile business demography statistics. This Manual is supplemented but not superseded by the Guidelines. The definitions and concepts of business demography statistics in the Guidelines are fully consistent with those in the Manual. The SBR concepts used in the Manual are themselves based on the Eurostat Business Register Recommendations Manual.

1.22   The Guidelines complement and extend the Manual by further developing methodological aspects in relation to the compilation of statistics by linking SBR data with other data sources. The Guidelines also provide an initial discussion of issues concerning business demography at the sub-national level and concerning business demography using statistical units other than the enterprises (i.e., establishments and local units) and the need to adapt concepts accordingly. Finally, the Guidelines emphasise the effects of the compilation of business demography statistics in supporting improvements in the quality of economic statistics.

1.23   In summary, the Guidelines:

a. Deliver guidance on how to develop SBRs to support the production of business demography and entrepreneurships statistics, including:

- Definitions of key demographic events, including births and deaths.

- Suggestions for basic demographic statistics that can be produced directly from the SBR.

- Guidance regarding methods to produce business demography statistics.

- Guidance on methods for linking SBRs to other data sources (administrative/social registers and other statistical registers) to support analysis of enterprises and entrepreneurs, including gender entrepreneurship.

- Guidance on methods for linking SBRs with statistics on international activities of enterprises, such as statistics on trade, foreign-owned enterprises, and foreign affiliates.

b. Provide examples of other approaches not based on the SBR to produce business demography statistics.

c. Give an overview of entrepreneurship indicators that can be produced by NSOs.

1.5   Overview of the Guidelines

1.24   The Guidelines are structured into seven chapters, a Glossary and three annexes. The following paragraphs provide short descriptions of the topics and contents of each chapter.

Chapter 2: **Definitions and key concepts of business demography statistics**

1.25   Chapter 2 presents the definitions of the key business demography concepts and related variables, based on the Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007) and the UNECE Guidelines on Statistical Business Registers (UNECE, 2015). The chapter provides definitions and explanations of continuity rules, and identification of entrepreneurs, including the treatment of self-employed persons. It discusses definitions of high-growth enterprises, including so-called gazelles. The chapter also includes explanations on regional and sub-annual business demography, as well as the use of establishments and local units as statistical units, which is not dealt in the Eurostat-OECD Manual.

Chapter 3: SBR requirements to support the production of business demography statistics

1.26   Chapter 3 focuses on how to develop the SBR to facilitate production of business demography statistics. The chapter aims at building the bridge between the concepts of business demography statistics and the SBR infrastructure required to support the production of business demography statistics. There are no guidelines on this topic elsewhere. Development and maintenance of longitudinal business data play a key role in the chapter, including methods to identify demographic

events and to determine whether a business has been born, survived or died. Other forms of SBR requirements are also dealt with, such as coverage of the SBR, updating frequency and the variables needed to describe demographic events.

Chapter 4: Linking the SBR with data from other sources

1.27   Linking business demography data with other existing data sources is an efficient way of producing statistics and increases the analytical power of the demography dataset. As SBRs cover only a limited set of variables on the businesses and usually no information on entrepreneurs as (natural) persons, this chapter provides methods for and guidance on linking SBR information with other data sources. Other sources may include, for example, data on the success of enterprises (e.g., growth of employment, turnover, profit), foreign trade data, and data about the entrepreneur (e.g., sex, age, nationality, (former) employment status and education) that may be obtained from surveys, administrative registers or other statistical registers.

Chapter 5: Other approaches to produce business demography statistics

1.28   While the Guidelines focus on the production of business demography statistics based on the SBR, it is recognised that, in the short run, this may not be possible in all countries, so that the production of the business demography statistics must be based on other information. Thus, the chapter provides examples of production using data from surveys, censuses or administrative sources, or based on a combination of sources.

Chapter 6: Examples of business demography and related entrepreneurship indicators

1.29   The chapter outlines the minimum set of business demography indicators recommended to support analysis of entrepreneurship. Examples are presented to illustrate country practices in production and dissemination of statistics on business demography.

Chapter 7: Topics for further work and research

1.30   The chapter includes topics for possible further work and research, based on the discussions of the Task Force. The research agenda includes: development of longitudinal databases for producing business demography statistics; linking and integrating the SBR with information from other data sources, including administrative sources; development of a statistical framework for business demography statistics; the use of different statistical units in business demography; international comparability of business demography statistics; and sub-annual (quarterly) and regional indicators.

1.31   The Guidelines also include a Glossary with definitions and explanations of key terms and concepts, Annex 1: Developing and using longitudinal businesses data in Canada presenting the development and use of longitudinal businesses data in Statistics Canada, Annex 2: Profiling entrepreneurs by linking business and social statistics in Denmark illustrating the linking of business and social statistics in Denmark to profile the entrepreneurs, and Annex 3: Family businesses in the Netherlands on family businesses in the Netherlands.

Chapter 2: Definitions and key concepts of business demography statistics

2.1 Introduction

2.1 This chapter aims at providing the definitions and key concepts of business demography statistics and of statistics on high-growth enterprises. The chapter is mainly based on the Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007). The concepts and definitions of annual business demography based on enterprises presented in this chapter are fully in line with the Manual. In addition, the chapter describes the use of statistical units other than the enterprise, and regional and sub-annual business demography statistics, which are not covered in the Manual. The chapter also draws material from the Eurostat Business Register Recommendations Manual (Eurostat, 2010) and the UNECE Guidelines on Statistical Business Registers (UNECE, 2015). These manuals complement each other.

2.2 Chapter 2 is organised as follows: In Section 2.2, the basic concepts of business demography are introduced, followed in Section 2.3 by a detailed explanation of the key concepts, especially concerning births, survivals and deaths. The definitions of the variables and characteristics of business demography are explained in Section 2.4. Section 2.5 provides the key concepts of high-growth enterprises including "gazelles".

2.2 Basic concepts of business demography

2.2.1 Subject of business demography statistics

2.3 As explained in Chapter 1, business demography statistics provide data on the development of the business population over time and the core variables that describe this development. The focus of business demography statistics is on the birth of new businesses, the survival of newly created businesses and the death of businesses. Based on the numbers of births, survivals and deaths, various indicators can be derived, such as:

- Birth rate: newly born businesses in *t* as percentage of all active businesses in *t*.

- Survival rates: percentages of businesses born in *t* that are still active in *t+1*, *t+2*, *t+3*, etc.

- Death rate: businesses closed in *t* as percentage of all active businesses in *t*.

2.4 These indicators are usually broken down by economic activity according to the International Standard Industrial Classification (ISIC), legal form, size class (employment or turnover) and/or region. Additional characteristics may be obtained by linking business demography units with other business statistics databases, such as foreign trade databases, R&D statistics and innovation statistics (as further described in Chapter 4). In case of sole proprietorship businesses and owner-operators of incorporated enterprises or quasi-corporations, the data can be further enlarged by including characteristics of the entrepreneur, such as age, or sex, or other personal characteristics relevant for analysis of entrepreneurship. These additional characteristics may be obtained by micro-data linking with social statistics databases. Further information on the structure of the businesses is provided by indicators for the entire population of the active businesses. An example is data on the age distribution of businesses.

2.5 The two most important variables used to describe the size and development of the businesses are turnover and employment. It is of special political interest to provide data on the number of jobs that have been created by newly born businesses. Of course, the opposite also needs to be considered, i.e., the loss of jobs through cessation of businesses.

2.6   Other aspects of business demography are indicators of the growth rates of businesses. Within the entire population of active businesses, it is of particular interest to identify those that are growing very fast and thus creating considerable numbers of new jobs. Conceptually, *high-growth* businesses are defined as businesses over a certain size threshold that have an increase of employment or turnover above a certain threshold, over a three years period (see Section 2.5). "Gazelles" are a special sub-group of high-growth businesses, namely those that are young.

2.7   In principle, the opposite of high-growth businesses should also be considered to provide a full picture of the business development: i.e., businesses with large decreases in employment or turnover. However, international data collections by the EU and OECD cover only high-growth businesses.

2.8   Business demography statistics can be based on various statistical units and populations. In international data collections by the EU and OECD the statistical unit is the *enterprise*. Discussion of other units is provided in the following sections.

2.2.2   Populations for business demography statistics

2.9   There are three different populations that can be the subject of business demography statistics.

- The first is the population of *all businesses that are active* (i.e. having employment or turnover) at some time during the reference period. This population includes businesses that have no employees, termed *self-employed businesses*, which may be sole proprietorships or partnerships. This population is the basis for "business demography" and is usually the starting point for the compilation of business demographic statistics.
- The second is a subset of this population comprising *those businesses that have at least one employee*. This is the basis for "*employer* business demography" (referred to "employer enterprise demography" in Eurostat & OECD (2007).
- The third is a subset of the second population comprising *those business that have at least two employees*. It is the basis for "*economic* business demography" (referred to as "economic enterprise demography' in Eurostat & OECD (2007).

2.10   There is invariably a significant difference between the number of businesses in the overall population and in the two subpopulations, due to the many self-employed businesses. The definitions of births, survivals and deaths in the three populations are also different, as further discussed in Section 2.3 below.

2.11   *Business demography* provides the most comprehensive picture of the business births, survivals and deaths as it covers all businesses, including self-employed businesses (i.e., sole proprietorships and unincorporated partnerships).

2.12   As not all countries have data on self-employed businesses, excluding them increases international comparability. This is the reason for using *employer business demography*.

2.13   The reason for using *economic business demography* is of a more conceptual nature. If an entrepreneur creates a business and decides to choose the legal form of a limited liability company (or any other form of an incorporated business), he/she himself may be regarded as an employee (rather than as self-employed) by a legal/administrative source on which the SBR is based. Thus, such a business may be regarded as having an employee and included in employer business demography even though the business is a self-employed business from an operational perspective. Setting a threshold of two employees ensures such a business is excluded from *economic business demography*. If such a business was to be counted as a self-employed *for the purposes of official statistics*, even though the entrepreneur was regarded as employee in an administrative source, the distinction between employer business demography and economic business demography would be unnecessary (for more information on the concept of entrepreneurs, see Section 4.4.1).

2.14   In international data collection by the EU and OECD, business demography statistics on all enterprises (including both employers and non-employers) and separately on employer businesses (including only enterprises with at least one employee) are produced.

2.2.3   Standard statistical unit types

2.15   There are four standard types of statistical units that can be used in describing a business: *enterprise, kind-of-activity unit (KAU), establishment* and *local unit*. They are defined in the *International Standard Industrial Classification of All Economic Activities (ISIC) Rev.4* and in the *2008 SNA*. Box 2.1 presents these definitions. Further information is provided in *Guidelines on Statistical Business Registers*, chapter 4 (UNECE, 2015).

2.16   The existence of an international standard does not mean that for all statistical domains and in all countries the same statistical unit(s) is (are) used. In general, the choice of statistical unit depends upon the domain, the geographic coverage, and the reference period. For instance, for structural statistics the enterprise is typically the unit, as all kinds of business data can be collected from enterprises. However, an enterprise may have two or more locations in different regions. Thus, for regional statistics, the establishment or the local unit is a more appropriate statistical unit.

Box 2.1 **Definitions of enterprise, kind-of-activity unit, local unit and establishment**

Definition of an enterprise:

An enterprise is an economic transactor with autonomy in respect of financial and investment decision making, as well as authority and responsibility for allocating resources for the production of goods and services. It may be engaged in one or more productive activities and may have one or more locations.

Definition of kind-of-activity unit:

A kind-of activity unit (KAU) is an enterprise or part of an enterprise that engages in only one activity or in which the principal productive activity accounts for most of the value added. – This definition refers to the kind of activity carried out; it does not refer to the location where the activities are carried out.

Definition of a local unit:

A local unit is defined as an enterprise or a part of an enterprise which engages in productive activity at or from one location. - This definition refers only to the geographic dimension; it does not refer to the kind of activity that carried out.

Definition of an establishment:

The establishment is defined as an enterprise or part of an enterprise that is situated in a single location and in which only a single (non-ancillary) productive activity is carried out or in which the principal productive activity accounts for most of the value added. The establishment is also referred to as local kind-of-activity unit.

Source: ISIC, rev. 4 (United Nations, 2008)

2.17   Business demography statistics can be based on any of these statistical units. In the European Union the enterprise is mostly used. In non-European countries the establishment is often used. International data collections by the EU and OECD use the enterprise and do not produce demographic data for other types of units.

2.18   In addition to the four units listed in the box above there is also *enterprise group*. Of course, business demography data could be established for this unit too. However, due to the global character of enterprise groups and the impossibility to cover a group completely by national SBRs, there is yet no

empirical application anywhere. They are also not considered in these Guidelines. Chapter 22 of the Eurostat Business Register Recommendations Manual (Eurostat, 2010) deals with the demographic changes and continuity rules for the enterprise groups.

2.19   Demographic events and statistics depend on the choice of the statistical unit. For example, if a new establishment of an existing enterprise is created, then in establishment-based business demography there has been a birth whereas in enterprise-based business demography there has not been a birth. Likewise, if a new local unit of an existing enterprise is created, then in local unit-based business demography there has been a birth whereas in enterprise-based business demography there has not been a birth.

2.20   A specific complexity in identifying statistical units is that a business might consist of more than one legal unit. Legal units are legal persons whose existence is recognised by law independently of the individuals or institutions which may own them or are members of them, as well as natural persons who are engaged in an economic activity in their own right. In most cases, a legal unit with business activities is an enterprise. However, under certain circumstances, a legal unit may not be an enterprise as it does not have the core characteristics of an enterprise – it does not have autonomy with respect to financial and investment decision making, or it does not have authority or responsibility for allocating resources to produce goods and services. Such a legal unit can only occur in the context of an *enterprise group*, defined in the SBR Guidelines (UNECE, 2015) as *a set of legal units bound together by legal and/or financial links under the same control*. In summary, in the great majority of cases, a legal unit is an enterprise, an enterprise comprises only one legal unit, and an enterprise birth comprises just one legal unit.

2.21   Given the possibility that a newly created legal unit may not be an enterprise, it must be checked to see whether it forms a new enterprise or should be allocated to an existing enterprise. Identification of legal units that are *special purpose entities* (devices to raise capital or hold assets or liabilities) is straightforward as they usually have neither employment nor turnover. Such legal units are not regarded as active enterprises and are combined with the other legal units to form enterprises collectively.

2.22   There are circumstances in which legal units that have employment and turnover are still not regarded as separate enterprises. Such cases can arise from vertical integration or where production factors have been allocated to separate legal units, for example, one legal unit for buildings and machinery, another for the employment, etc. These legal units are combined with others to form an enterprise collectively.

2.2.4   Regional business demography

2.23   The focus of *regional* business demography is to provide data on the development of the business population for a specific region. An appropriate choice of statistical unit is essential. Regional statistics should not be limited to a view of national statistics by geography. The approach to regional business demography should take account of the regional aspects. This refers mainly to the statistical unit that is used. The enterprise, which can have locations in more than one region, is not the ideal statistical unit for regional statistics of any type, including regional business demography statistics. Whilst most enterprise births are likely to be very small units that are active in just one region, this is not the case for enterprise deaths. Furthermore, indicators "birth rate" and "death rate" by region based on enterprises are biased as enterprises are counted only in the one region where they have their headquarters and/or their main production.

2.24   Regional business demography should describe the development of a regional economy through growth and decline of employment/turnover in the respective region caused by births and deaths of businesses. In this sense, job creations by an additional establishment or local unit of already existing enterprises is as important as job creation by an enterprise birth in that region. In summary, the best unit for regional business demography is the establishment or the local unit, not the enterprise.

2.25   Even though enterprises are conceptually not ideal units for regional business demography, international data collections by the EU and OECD are currently based on enterprises, with the result that national business demography statistics are simply subdivided by region. The totals of births and deaths over the regions are the same as the national totals.

2.26   The main reason why the enterprise is used as the statistical unit is data availability. Data at the level of local unit or establishment are not available in many countries.

2.2.5   Sub-annual business demography

2.27   Sub-annual business demography provides data on births and deaths on a typically quarterly basis with the aim of providing data that are timelier than annual data. There is, in principle, no conceptual difference between annual and sub-annual business demographic events and the resulting statistics. However, lack of availability of sub-annual data limits implementation of all the concepts applied to annual data. For example, a unit seems to be dormant and information indicating whether it is just temporarily inactive and might be reactivated in some near future or whether it is dead may not be available sub-annually.

2.28   Therefore, for sub-annual business demography statistics less precise definitions may be appropriate. For example, instead of identifying real business births, legal unit registration data from a company register or other administrative register may be used. Thus, sub-annual business demography data tend to provide proxies and short-term indicators for business dynamics rather than conceptually precise demographics. In view of these weaknesses, sub-annual data are normally shown in the form of time series indicators only, rather than in absolute values.

2.2.6   High-growth enterprises and gazelles

2.29   The creation of jobs through new enterprises on the one hand, and through fast growing, already existing enterprises on the other, are the two main anchors of entrepreneurship statistics. *High-growth* enterprises are defined as enterprises over a certain size that have a high increase in turnover or employment over a certain period of years. It is not relevant whether these enterprises were recently created or not (see Section 2.5.1).

2.30   High-growth enterprises that were born in recent years are known as *gazelles*. They are described in Section 2.5.2. Data on high-growth enterprises, including gazelles, are typically broken down by economic activity and size classes.

2.3   Key concepts and definitions

2.3.1   Demographic events and continuity rules

2.31   Demographic events are events with an impact on the existence of statistical units and/or the links between them. The mere growth or decline of a statistical unit is not a demographic event. Demographic events can be split into two types:

- *existential change,* i.e., a change involving the emergence or disappearance of the combination of production factors constituting a statistical unit; and

- *distributional* change, i.e., a change in the distribution of production factors between units.

2.32   A *production factor* is any good or service used to produce output. Production factors are normally grouped into categories such as employment, machines and equipment, land, buildings, management, and intangible assets. The continuity of a statistical unit is determined by the continuity of its production factors.

2.33   Existential changes mean the emergence or disappearance (births and deaths) of statistical units. Due to existential changes, the number of statistical units may increase or decrease. Distributional changes may result in a decrease of the number of statistical units, an increase, or no change.

2.34   Demographic events can be defined for all types of statistical units (enterprise group, enterprise, kind-of-activity unit, local unit, and establishment). The enterprise is the central statistical unit as all other types are defined with respect to it. A kind-of-activity unit results from the division of an enterprise by economic activity, a local unit results from a division by geography, and an establishment results from division by economic activity and geography. An enterprise group is a combination of enterprises bound together by legal and/or financial links.

Demographic events for enterprises

2.35   Table 2.1 provides an overview of the typology of demographic events for enterprises, drawn from the Eurostat-OECD Business Demography Statistics Manual (Eurostat & OECD, 2007). The explanations of the resulting unit structures caused by the demographic events listed in Table 2.1 below are as follows.

Birth: independent event affecting only one enterprise and involving the creation of a new combination of factors of production.

Reactivation: enterprise is dormant for a period of less than two years and then recommences activity in a way that complies with the definition of continuity.

Death: independent event affecting only one enterprise and involving the dissolution of the combination of factors of production.

Change of ownership: restructuring of the legal basis of an enterprise; the enterprise remains alive and active; it is said to *continue* (or equivalently, to *survive*); there is no impact on the demographic variables.

Merger: opposite of a break-up; consolidation of the production factors of two or more enterprises into a single new enterprise; no enterprise survives, but the closures of the original enterprises are not considered as deaths; similarly, the new enterprise is not considered as a birth.

Take-over: opposite of a split-off; the enterprise that does the take-over survives; the enterprise(s) taken over is (are) closed, i.e., they do not survive, but not considered to be death(s).

Break-up: opposite of merger; division of the production factors of an enterprise into two or more new enterprises; the original enterprise is closed, but its closure is not considered a death; similarly, the new enterprises are not considered to be births.

Split-off: similar to a break-up except that the original enterprise does survive; there is no death; one or more new enterprises are created but these are not considered to be births.

Creation of a joint venture: two or more independent enterprises agree to commit some of their resources to work together on a common project towards a common goal; none of the independent enterprises have control over the joint venture; a joint venture is considered a birth if it involves the creation of a new combination of production factors.

Cessation of a joint venture: opposite of a creation of a joint venture; considered as death if less than half of the employment is transferred back to the participating enterprises.

Restructuring within an enterprise: involves only one enterprise and has no impact on the demographic variables.

Restructuring within an enterprise group: involves the creation or cessation of one or more enterprises under common ownership: does not result into enterprise births or deaths.

Change of group: similar to the change of ownership; the enterprise remains alive and active; there is no impact on the demographic variables.

Complex restructure: similar to a restructure within an enterprise group but involves two or more enterprise groups; depending on the kind of the restructuring, the outcomes in terms of demographic events may be quite different from one another.

Table 2.1 **Typology of demographic events for enterprises**

Demographic event	Numbers of enterprises		Enterprises in SBR	
	Before the event	After the event	Number of creations	Number of deletions
Birth	-	1	1	-
Reactivation	-	1	-	-
Death	1	-	-	1
Change of ownership	1	1	-	-
Merger	n	1	1	n
Take-over	n	1	-	n-1
Break-up	1	n	n	1
Split-off	1	n	n-1	-
Creation of a joint venture	n	n+1	1	-
Cessation of a joint venture	n	n-1	-	1
Restructuring within an enterprise	1	1	-	-
Restructuring within an enterprise group	n	n	0 or more	0 or more
Change of group	1	1	-	-
Complex restructuring	n	n	0 or more	0 or more

Source: Based on Eurostat & OECD (2007). (Note: n = 2 or more)

2.36   A key message from these definitions is that the *creation (entry)* of an enterprise does not necessarily mean the *birth* of an enterprise for business demography as defined above. This may be confusing for SBR staff who quite likely refer to creation of an enterprise in the SBR as a "birth". Likewise, the cessation (closure, deletion) of an enterprise does not necessarily mean a death in the business demography context. Again, this may well be confusing for SBR staff who quite likely refer to closure of an enterprise in the SBR as a "death". However, where only one enterprise is involved, the SBR and business demography definitions of birth coincide, as do the definitions of death.

Continuity rules for enterprises

2.37   The core conceptual basis for treatment of demographic events in the SBR and in business demography in the case of an enterprise are the *continuity rules*, i.e. the conditions for keeping or changing an enterprise identity number. Various enterprise characteristics are considered in defining

continuity (also called survival), depending on the use of the resulting data. In the SBR and in business demography, continuity of an enterprise is defined in terms of the continuity of its production factors. If its factors of production are considered to have continued, the enterprise is considered to have continued. It is discontinued if its factors of production are discontinued. In practice, all the production factors cannot be analysed in detail, and more practical rules are used.

2.38   Considering what is available in the SBR and the information supplied by administrative and other sources, the following three criteria have proved to be practical in defining continuity rules:

Controlling legal unit. The controlling legal unit controls the production factors of the enterprise. The continuity of the management of the enterprise may be assumed to be linked to the continuity of the controlling legal unit. Thus, the convention is that continuity of controlling legal unit suggests continuity of the enterprise. However, a change of the controlling legal unit is not a sufficient reason to delete the existing enterprise record in the SBR and create a new one.

Economic activity. It is assumed that a change in the economic activity does not imply a change in production factors, at least not abruptly. The convention is that a change of the principal activity is not sufficient reason to delete an existing enterprise record and create a new one in the SBR.

Location. The continuity of the locations where the activities are carried out is closely linked to the continuity of the land and buildings used by the enterprise. In the case of an enterprise ceasing its activities at its main location and resuming its activities at another location within the national territory, the answer to the question whether there is continuity or not, is not obvious. If the activities do not move far, the probability is high that the production factors (other than land and buildings) are largely continued. However, a long-distance move might mean that other production factors have also changed and thus continuity is not certain. The general rule for a change of the main location is that continuity is assumed. In other words, change of main location is not a reason in itself to delete an existing enterprise record and create a new one in the SBR.

2.39   From the above it can be concluded that a change in only one of these three criteria does not generally imply loss of continuity. On the other hand, if all three criteria change, this certainly is a loss of continuity. A practical rule is that in case of a change in two of the three criteria, continuity shall not be assumed. However, even this rule should be applied with care. Analysis of individual circumstances, especially in the case of big enterprises and complex changes, is advisable. The following table summarizes the general continuity rules for the enterprise:

Table 2.2 **Continuity rules for an enterprise**

Change of controlling legal unit	No	Yes	No	No	Yes	No	Yes	Yes
Change of principal activity	No	No	No	Yes	No	Yes	Yes	Yes
Change of main location	No	No	Yes	No	Yes	Yes	No	Yes
Continuity of enterprise?	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**	**No**	**No**

Source: Eurostat & OECD (2007)

Demographic events and continuity rules for local units and establishments

2.40   In addition to demographic events and continuity rules for enterprises, analogous typologies and rules are required for the other types of statistical units. Typologies for local units and establishments are discussed in the following paragraphs.

2.41   As previously noted, all types of statistical units are defined in terms of the enterprise. If an enterprise is created, a local unit and an establishment are created at the same time. When an enterprise

is ceased, also its local units and its establishments are ceased. Thus, demographic events at enterprise level have effects for the associated local units and establishments.

2.42   In defining continuity rules for local units, location, production factors and principal activity come into play. The production factors are usually assumed represented by the employment of the unit, as this is relatively easy to observe. A practical convention is that the local unit is deemed to continue if 50% or more of the persons employed by the local unit continue to work at or from the same location.

2.43   Another case to consider is a change of the enterprise to which the local unit belongs. A local unit might be transferred from one enterprise to another. If nothing else changes, the local unit is deemed to have continued.

2.44   Concerning change of location, the general rule is that a local unit loses its continuity, irrespective of whether other criteria (enterprise, principal activity, and employment) change or not. However, if the move is over a short distance without changes in any other criteria, the local unit is assumed to continue. There is no harmonised definition of what should be understood as "short distance". A possible approach is to base the rule on the regional classification in use in the country and state that moving out of a region results in a loss of continuity. The "regions" to which this rule is applied should be the smallest administrative areas, for example, municipalities.

2.45   The following table summarizes the continuity rules for local unit under the assumption that there is no change of the location of the unit:

Table 2.3 **Continuity rules for a local unit that does not change location**

Change of enterprise	No	Yes	No	No	Yes	No	Yes	Yes
Change of principal activity	No	No	No	Yes	No	Yes	Yes	Yes
Change of employment	No	No	Yes	No	Yes	Yes	No	Yes
Continuity of the local unit?	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**	**No**	**No**

2.46   In discussing demographic events for local units, it is useful to distinguish demographic events *below* the enterprise level from those *at* the enterprise level. In the first case, as the enterprise itself is unchanged, only events at the level of the local unit are defined. In the second case, events affecting the enterprise itself are involved and these may cause changes at the local unit level.

2.47   The birth of an enterprise entails the birth of a local unit; the death of an enterprise means the death of the local unit(s). For all the other kinds of demographic events of enterprises (merger, takeover, break-up, split-off, joint venture, restructuring, etc) local units are transferred to an enterprise that exists after the demographic event. Thus, they continue, unless other changes also happen. For example, during such demographic events, an enterprise may close one or more of local units or create one or more new ones. Of course, enterprises may also create or close local units without any demographic events at the enterprise level.

2.48   The continuity rules and the treatment of the demographic events for the establishments are the same as for the local units, thus are not further dealt with here.

2.3.2   Coverage

2.49   Coverage is defined using the enterprise unit. Where business demography statistics are provided for establishments or local units, the units included are those that belong to the enterprises defined to be in scope.

2.50   Business demography statistics aim to provide a picture of the development of the business population by data on newly created businesses and the number of jobs created through these new

businesses on the one hand, and by data on business closures and the jobs that are lost by these closures on the other. In principle, business demography should include *all market-oriented enterprises*. However, in the EU, the agricultural sector is excluded, and this is quite a common exclusion.

2.51   As business demography focusses on market producers, government institutions as well as non-profit institutions are excluded. Of course, government and non-profit institutions could be included in business demography statistics, but, if they were to be, they would need to be presented in separate tables.

2.52   The coverage of a business statistics domain is often defined by referencing the ISIC Rev. 4 activities that should be included. However, ISIC Rev. 4 does not distinguish between market and non-market activities. The delineation of the coverage of business demography statistics should thus be based on the classification of institutional sectors according to 2008 SNA and the European System of National and Regional Accounts, ESA 2010 (Eurostat, 2013). Included in business demography are the sectors:

- S 11   Non-financial corporation
- S 12   Financial corporations
- S 14.1   Household sector: Employers
- S 14.2   Household sector: Own account workers

2.53   For details on the sector classifications and its application see 2008 SNA and ESA 2010 as well as the SBR Guidelines (UNECE, 2015). According to EU obligations, the classification of institutional sectors is incorporated into the SBRs and thus all institutional units in the register are coded by this classification. By defining the coverage in terms of the sector classification, additional coverage definitions by economic activities are not needed, except where certain economic activities should explicitly be excluded, such as the exclusion of ISIC section A Agriculture, forestry and fishing, which is widely applied.

2.3.3   Population of active enterprises

2.54   The starting point for derivation of business demography events, where the statistical unit is the enterprise, is the definition of the population of *active enterprises*. For (overall) business demography it comprises all enterprises that were active in the whole year t or in a part of year t. Being active is defined as producing goods and services (for the market) and thus having employment (at least one self-employed) or turnover. Also, units without actual turnover, but with investment activities should be considered as being active.

2.55   In case of the employer business demography, or economic business demography, the criteria of being active is measured by the number of employees, which, in the former case, should be at least one and in the latter case at least two. For the employment criterion, the data should be measured on an *annual average basis* (see Subsection 2.4.2).

2.56   The number of enterprises in business demography statistics is larger than the number that were active at the end of the year as enterprises that ceased activity during the year are counted as being active in the year.

2.3.4   Enterprise births

2.57   Entries (creations) of enterprises, i.e. enterprises that are present in a given period but were not present in the previous period, are generally easy to identify in the SBR. However, as previously noted, only some of these entries are births for business demography. The main criterion in the definition of a birth is that, in the process of entry, no other enterprise is involved. Enterprises emerging from mergers, split-offs and restructuring are not viewed as births. Also new legal units based on changes in the legal form or on short-distance move of the location are not viewed as enterprise births. Box 2.2, drawn from Eurostat & OECD (2007), provides more explanation.

2.58   An enterprise that principally fulfils the below criteria and is undertaking investment activities, but does not yet have turnover or employment, should also be viewed as being active and thus being an enterprise birth.

2.59   A newly born national or foreign subsidiary should be considered an enterprise birth if:

- it is an enterprise (i.e., legal unit not just local unit or branch) with autonomy of decision making; and

- new production factors are created, rather than transferred from another unit.

Box 2.2 **Definition of an enterprise birth**

> An enterprise birth amounts to the creation of a combination of production factors with the restriction that no other enterprises are involved in the event: Births do not include entries into the populations due to mergers, break-ups, split-offs or restructuring of enterprises. They do not include entries into a sub-population resulting only from a change of activity.
>
> A birth occurs when an enterprise starts from scratch and actually starts activity. An enterprise creation can be considered an enterprise birth if new production factors, in particular new jobs, are created. If a dormant unit is reactivated within two years, this event is not considered a birth.
>
> *Source:* Eurostat & OECD (2007)

2.60   For any newly created enterprise, a set of criteria must be checked before deciding if the new enterprise is a birth or not.

2.61   As three business demography populations are distinguished, the definition of enterprise birth needs to be appropriately adapted to each. In the case of the business demography, as noted above, the focus of the definition is the new creation of a combination of production factors. In case of employer business demography and economic business demography the decisive criterion is the number of employees. More details are provided in Table 2.4.

2.62   The same enterprise may be recorded as a birth in more than one of these populations. A newly created enterprise starting from scratch is counted as a birth when the economic activity started, irrespective of whether the enterprise has employees. If it has at least one employee from the beginning, it is also a birth for employer business demography. Similarly, if it has two or more employees it is a birth for economic business demography. However, taking on employees might not happen in the same year the enterprise is a birth. It may start with no employees and engage its first employee in a later year than the creation year. In this case the enterprise will not be a birth for employer enterprise demography until this later year. Table 2.5 lists the possible cases when an enterprise should be recorded as a birth for the three business demography populations. Related questions are: when does a birth happen from a business demography perspective and how can it be recorded? Three possible solutions are as follows:

- an enterprise may be viewed and recorded as a birth when investments are made, i.e., even before any production activities, or sales to the market;
- an enterprise may be viewed and recorded as a birth when production activities start - an indicator for this is turnover/sales, which is more likely to be available than information on investment;
- an enterprise may be viewed and recorded as a birth when the enterprise engages one or more employees, even if there are no sales.

Table 2.4 **Enterprise births for the three kinds of business demography populations**

Births for business demography	Births of all enterprises, regardless of whether they are employers or not. No general threshold is applied to the size of the enterprise in terms of employment or any other characteristics.
Births for employer business demography	Births of enterprises with at least one employee, comprising (1) births of enterprises that have at least one employee in the birth year and (2) enterprises that existed before the year in consideration, but were below the threshold of one employee (entry by growth).
Births for economic business demography	Births of enterprises with at least two employees, comprising (1) births of enterprises that have at least two employees in the birth year and (2) enterprises that existed before the year in consideration, but were below the threshold of two employees (entry by growth).

2.63   For employer business demography and economic business demography the solution is conceptually easy: such units are recorded as births in the year when the employees' threshold has been reached.

Table 2.5 **Cases of enterprise births**

	Business demography	**Employer business demography**	**Economic business demography**
Enterprise A starts its activities in t as a self-employed person	Birth in t	Out of scope	Out of scope
Enterprise B starts its activities in t with 1 employee from the beginning	Birth in t	Birth in t	Out of scope
Enterprise C starts its activities in t with 2 employees from the beginning	Birth in t	Birth in t	Birth in t
Existing Enterprise A engages one employee in t+1	Not a birth	Birth in t+1 (entry by growth)	Out of scope
Existing Enterprise A engages one further employee in t+2	Not a birth	Not a birth	Birth in t+2 (entry by growth)
Existing Enterprise B engages one further employee in t+1	Not a birth	Not a birth	Birth in t+1 (entry by growth)

2.64   Lastly, it should be noted that an enterprise may consist of more than one legal unit. As explained in paragraph 2.20 ff., it is necessary to check whether a new legal unit forms its own enterprise (in which case there is an enterprise birth), or whether it should be allocated to an existing enterprise (in which case it is not an enterprise birth).

2.3.5   Enterprise survival

2.65   The second focus of business demography is on the survival of newly born enterprises over the subsequent years. This type of analysis is based on data for a cohort of enterprise births in year *t*, which enables follow-up of the development of the newly born enterprises over the years: how many have survived and what is their economic development in form of employment and/or turnover. The time

horizon for this type of analysis is usually three to five years but could also be much longer. For each year, a survival rate is calculated. Box 2.3 contains the definition of enterprise survival.

Box 2.3 **Definition of survival of an enterprise**

Survival of an enterprise occurs if an enterprise is active in terms of employment and/or turnover in the year of birth and the following year(s). Two types of survival can be distinguished:

- An enterprise born in year t or having survived to year t is considered to have survived to year t+1 if it is active in terms of turnover and/or employment in any part of year t+1 (= survival without changes).
- An enterprise is also considered to have survived if the linked legal unit(s) have ceased to be active, but their activity has been taken over by a new legal unit set up specifically to take over the factors of production of that enterprise (= survival by take-over).

Source: Eurostat & OECD (2007)

2.66   As stated in Section 2.3.4 newly created enterprises that have only investment activities, but not yet turnover or employment are also counted as enterprise births. Therefore, such enterprises are also viewed as having survived to year *t+1*, even if they still have only investment activities, without turnover or employment.

2.67   As described in paragraph 2.35, in the case when an enterprise A takes over another enterprise B, then enterprise B does not survive, and enterprise A survives. In case of a merger of two enterprises none of the two is considered to have survived.

2.68   However, if a new enterprise A (not a reactivation) is created in *t+1* to take over the production factors of another enterprise B, which was active in *t*, then the taken over enterprise B is considered as survived in *t+1*. In this case the newly created enterprise A is not considered as an enterprise birth.

2.69   Table 2.6 summarizes the definitions of survival for the three kinds of business demography.

Table 2.6 **Enterprise survival for the three business demography populations**

Business demography	An enterprise born in year t or having survived to year t has survived to year t+1 if it is active in terms of turnover or employment in year t+1.
Employer business demography	An employer enterprise born in year t or having survived to year t has survived to year t+1 if it has at least one employee in year t+1.
Economic business demography	An economic enterprise born in year t or having survived to year t has survived to year t+1 if it has at least two employees in year t+1.

2.70   The rules for survival of enterprises are illustrated in Table 2.7.

2.71   Identification of units that have survived should always be done on a year-to-year basis. An enterprise born in year *t* can survive to year *t+2* only if it has survived to year *t+1*. The methodology for the identification of survivals is the same for all consecutive years starting from the year of birth. The criteria of survival are, of course, different depending on the population that is being considered.

2.72   Survival rates decrease over the years. After five years there could be cases where only about half of the enterprises are likely to be still active. In other words, half of the enterprises have ceased their activity within five years. However, survival rates differ between countries, economic activities and size of the enterprise. Of course, business demography statistics based simply on employment and/or turnover do not indicate the reasons for cessation. Additional information is needed that may be available only directly from the enterprises.

Table 2.7 **Cases of enterprise survival**

	Business demography	Employer business demography	Economic business demography
Enterprise A continues its activities in t+1 with only a self-employed person	Survival to t+1	Out of scope	Out of scope
Enterprise B continues its activities in t+1 with 1 employee	Survival to t+1	Survival to t+1	Out of scope
Enterprise C continues its activities in t+1 with 2 employees	Survival to t+1	Survival to t+1	Survival to t+1

2.3.6   Enterprise deaths

2.73   Like enterprise births, enterprise deaths are identified based on the entire business population. Data on enterprise deaths thus refer to enterprises that were born in any year, not necessarily just in the preceding years. The definition of an enterprise death is provided in Box 2.4.

Box 2.4 **Definition of an enterprise death**

A death amounts to the dissolution of a combination of production factors with the restriction that no other enterprises are involved in the event. Deaths do not include exits from the population due to mergers, take-overs, break-ups or restructuring of a set of enterprises. It does not include exits from a sub-population resulting only from changes of activity.

Source: Eurostat & OECD (2007)

2.74   The definitions of deaths for the three different business demography populations are given in Table 2.8.

Table 2.8 **Enterprise deaths for the three business demography populations**

Business demography	Enterprise deaths cover all enterprises, regardless of whether they are employers or not. No general threshold is applied to the size of the enterprise in terms of employment or any other characteristic.
Employer business demography	An employer enterprise death occurs either as an enterprise death with the enterprise having at least one employee in the year of death, or as an exit by decline, the enterprise moving below the threshold of one employee.
Economic business demography	An economic enterprise death occurs either as an enterprise death with the enterprise having at least two employees in the year of death or as an exit by decline, the enterprise moving below the threshold of two employees.

2.75   Like in the case of births also in the case of deaths, the same unit may be recorded as a death in more than one of these populations. For instance, an enterprise may move below the threshold of one employee in year *t+1* and thus will be counted as employer enterprise death. It may cease all its activity in *t+2*, which means an enterprise death in *t+2*. Table 2.9 provides some examples.

2.76   The definitions of deaths should be applied in similar ways to those for births and survivals so that there is consistency between birth, survival and death data. This means that the following enterprise closures should not be counted as enterprise deaths:

- An enterprise that closes due to merging or breaking-up of production factors.

- An enterprise whose activity is taken over by another enterprise.

- An enterprise that is deleted due to a change of legal form, for example, a sole proprietor changes legal form to a limited company.

- A reactivated enterprise that restarts activity within 2 calendar years.

Table 2.9 **Cases of enterprise deaths**

	Business demography	**Employer business demography**	**Economic business demography**
Enterprise A with no employees ceases its activities in t+1	Death in t+1	Out of scope	Out of scope
Enterprise B with one employee ceases its activities in t+1	Death in t+1	Death in t+1	Out of scope
Enterprise C with two employees ceases its activities in t+1	Death in t+1	Death in t+1	Death in t+1
Enterprise B reduces its employment from one employee to zero employees in t+1	Survival in t+1	Death in t+1 (exit by decline)	Out of scope
Enterprise C reduces its employment from two employees to one employee in t+1	Survival in t+1	Survival in t+1	Death in t+1 (exit by decline)
Enterprise C with one employee ceases its activities in year t+2	Death in t+2	Death in t+2	Out of scope

2.77 Enterprise deaths should be counted in that period where the enterprise has ceased its production activities. There may be some employment or turnover during an ongoing liquidation process, so commencement of liquidation does not necessarily mean closure. Accordingly, the Eurostat/OECD Manual (Eurostat & OECD, 2007) recommends that a unit in liquidation should still be viewed as active, so not as a death. Also, even if production has ceased, the legal unit associated with the enterprise may continue to be recorded as an active unit in an administrative database.

2.78 In practice, recording enterprise deaths is more difficult than recording enterprise births. One reason is that information on the data of cessation may not be available in a timely manner from administrative sources and/or these sources may reflect only administrative deaths. Thus, additional data, such as information on turnover or employment may have to be used. However, this information may not be available in a timely manner, especially turnover data from taxation.

2.79 A further difficulty is that an enterprise can be temporarily inactive. If the enterprise restarts its activity within two years it should not be counted as a death. Hence, whether an inactive enterprise has died or not, can only be determined at the end of a two-year period. Thus, data on enterprise deaths may be regularly reviewed and updated in later years, which is normally not the case with birth data.

2.80 In the case of an enterprise comprising more than one legal unit, the rules concerning its possible death when one of its legal units dies are the mirror image of the rules concerning enterprise births (see paragraph 2.64).

2.4   Variables and characteristics of business demography

2.81 This section discusses the variables and characteristics that form the core of business demography statistics. In case of micro-data linking of business demography units with other databases,

additional variables and characteristics will emerge. Such variables may, for example, be engagement in international trade, or in R&D, as further discussed in Chapter 4. They are not discussed in this section.

2.4.1 Activity classification

2.82   For presentation of business demography data, activity classification ISIC Rev. 4 should be used. In most countries, the ISIC Rev. 4 is applied in the SBR and thus activity codes are available. In the European Union, a more detailed, but fully compatible activity classification (NACE Rev. 2) is obligatory by EU regulation. Using ISIC Rev. 4 (or NACE Rev. 2) ensures international comparability. Any national adaptations of ISIC Rev. 4 for national purposes should be in line with the hierarchical structure of the classification so that international comparability is ensured. Further information concerning the application of activity classifications is available in these documents and in UNECE (2015).

2.83   As concerns the appropriate level of detail, use of the detailed classification levels of ISIC Rev. 4 is recommended to provide a detailed picture of the development of businesses. This refers especially to activity areas with a high number of enterprise births or deaths. Of course, the level of detail might be influenced by confidentiality constraints. As most newly born enterprises are active in services industries, it is recommended that a more detailed breakdown of services be applied for services than for other economic sections.

2.84   There is one specific aspect in the use of activity codes for survival time series that must be considered. Changes in economic activity that take place are recorded in the SBR according to the appropriate SBR rules. However, changes in the activity codes of newly born units over time make it difficult to follow a birth cohort over subsequent years by economic activity. Thus, for business demography, the activity code as recorded at the time of birth is needed. This means that the longitudinal database that needs to be created for business demography statistics should include the activity codes at the year of birth, as further discussed in Chapter 3.

2.85   There is a further complication in case of a high-growth enterprise as its economic activity code should be observed at the end of the growth period. If there is a change in activity during the growth period, the activity into which the enterprise grew may be considered more relevant than the activity out of which it grew. So, survivals are reported using the classification code of the year of birth, while high-growth enterprises, including gazelles, are reported using the current year activity code. Enterprise deaths are also reported using the classification code at the time of the death.

2.86   The approaches described above are the current conventions (Eurostat & OECD, 2007) and are recommended here. In addition, survival rates and data on high growth enterprises and gazelles could also be provided based on activity codes at birth. However, such data are not part of the international data outputs.

2.4.2 Employment

2.87   The most important variable in business demography is employment, comprising self-employed and employees. It is the main size class criterion (see 2.4.4). It is necessary to define the employer business and economic business populations. Employment data are typically presented for the entire population of active businesses, with breakdown of:

- employment changes of the business population

- employment growth created by newly born businesses

- employment loss by closure of businesses

- employment growth (or loss) of newly born units over subsequent years

2.88   The definition of employment in business demography should be the same as in other business statistics. The Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007) provides the basic definitions based on 2008 SNA and ESA 2010, as indicated in Box 2.5.

Box 2.5 **Definition of employees and self-employed**

Employees are defined as all persons who, by agreement, work for another resident institutional unit and receive compensation in the form of wages, salaries, fees, gratuities, piecework pay or remuneration in kind. An employer-employee relationship exists when there is an agreement, which may be formal or informal, between an enterprise and a person, normally entered voluntarily by both parties, whereby the person works for the enterprise in return or remuneration in cash or in kind.

The term "employees" includes part-time workers, seasonal workers and persons on strike or on short-term leave but excludes persons on long-term leave. Employees do not include voluntary workers.

A person is considered wage or salary earner of a unit if he or she receives a wage or salary from the unit regardless of where the work is done (inside or outside the production unit). A worker from a temporary employment agency is considered an employee of the temporary employment agency and not of the unit (customer) in which he or she is working.

Self-employed persons are defined as persons who are the sole owner, or joint owners, of the unincorporated enterprises in which they work, excluding those unincorporated enterprises that are classified as quasi-corporations. Self-employed persons are classified here if they are not also in paid employment which constitutes their principal activity: in the latter case they are classified as employees. Self-employed persons also include the following categories: unpaid family workers, outworkers and workers engaged in production undertaken entirely for their own final consumption or capital formation, either individually or collectively.

Source: Based on 2008 SNA and ESA 2010

2.89   As noted in Chapter 1, the 2008 SNA definition of self-employed excludes owners of incorporated enterprises and unincorporated quasi corporations. Self-employed according to SNA is not representing the whole group of entrepreneurs as the owner-operators of incorporated enterprises and quasi-corporations are classified as employees. Due to the significant share of owner-operators who have chosen for their enterprise another legal form than that of a sole proprietorship, the count of self-employed persons based on the SNA results in a major underestimation of entrepreneurs.

2.90   The International Classification of Status on Employment (ICSE-18) adopted by the 20th International Conference of Labour Statistician in October 2018 (ILO, 2018) gets around the limitations of the SNA-based definition for the identification of entrepreneurs. ICSE-18 avoids the term self-employed and defines *independent workers* as those who own and control an economic unit for which they work, whether it is incorporated or not. The category of *independent workers* provides the best starting point for the identification and compilation of statistics on entrepreneurs (see Section 4.4.1).

2.91   ICSE-18 also introduces the category of *workers in employment for profit* and *workers in employment for pay* based primarily on the nature of the remuneration for a particular job. Self-employed persons according to SNA corresponds more closely to the concept of workers in employment for profit than independent workers, except for quasi-corporations which is not separately identified in ICSE-18.

2.92   A person managing and owning an incorporated enterprise or an unincorporated quasi-corporation should, from an economic point of view, be regarded as self-employed rather than as an employee in business statistics, including business demography statistics. However, current international definitions in economic and business statistics follow SNA and treat a person managing and owning a limited liability company as an employee of that company. In future business demography and entrepreneurship statistics, a definition of such owner-operators without employees based on the 20th ICLS standards should be targeted. However, such a conceptual change would need to be introduced simultaneously in all business statistics domains to avoid inconsistencies.

2.93   In economic statistics and in national accounts, employment measured in *full time equivalent* or, even better, in *hours worked*, is the preferred measurement concept. However, data on full time equivalent employment or hours worked are not available in many countries. Thus, for reasons of international comparability in business demography statistics, *number of persons employed* has been chosen as the measurement unit, irrespective of their hours of work. For annual business statistics, employment should be calculated on an *annual average basis*. Thus, in principle, for each business, more than one employment measurement per year should be available. Ideally, monthly employment data should be available.

2.94   In the case of monthly data, the annual average should be calculated by dividing the aggregated total employment by 12, and in the case of quarterly data by 4. Of course, if there is only one employment figure available, for instance at the end of the year, an annual average cannot be calculated. In this case, if there is reason to suppose this figure does not closely represent an annual average, additional estimates might be derived in some way. Differences between an employment figure by the end of the year and the (unknown) annual average are particularly likely in case of seasonal enterprises.

2.95   Where an enterprise is not active for the whole year, being born or closed during the year, the question is whether an *annual average* should be applied or an *average that is based on the length of the actual operating period*. The annual average is the main employment concept in economic statistics, as jobs are viewed as a labour input indicator to the economic production process. In business demography statistics – where the main interest is in the development of the number of jobs - it could be argued that an average over the operating period is preferable. For businesses, operating the whole year there is clearly no difference between the results of these methods. However, for businesses that have an operating period less than a year, the average over the operating period results in higher employment figures than using an annual average. For the entire population of active businesses, the total effect is quite small as most enterprises operate over the whole year.

2.96   Assume that an enterprise was born in October of year t and has 4 employees in each of the remaining months of year t. The operating period of that unit is thus three months, so the average employment based on the operating period is 4, whereas in case of the annual average the average employment is 1. The average employment at the end of year t (compared to year t-1) has increased by 4 by the first method and by 1 by the second. The newly born enterprise has created four jobs, but not for the whole year. However, if it is assumed that the 4 jobs at the end of year t are sustainable, the business demography message is that employment has increased by 4.

2.97   Following Eurostat & OECD (2007) the number of persons employed and the number of employees should be calculated as an *annual average over the operating period* of the enterprise. This is also recommended here.

2.4.3   Turnover

2.98   Turnover is one of the output measures of economic activity. Turnover is defined in Box 2.6. Strictly speaking, turnover (sales to the market) is a variable that can be observed only for enterprises. In case of establishments, the output measure is the total output, which also includes deliveries (internal sales) between the establishments of an enterprise.

2.99   In the case of business demography statistics, turnover does not play a leading role. In the obligatory data delivery programme of the European Union no turnover data are requested. One of the reasons is that turnover data – especially when taken from administrative sources – tends to be untimely. Another reason is the focus in business demography statistics on job creation related to enterprise births. A further reason is that, due to the differing economic development of the countries, the different currencies and purchasing power parities, monetary data on turnover are not easy to compare internationally. However, difficulties in interpreting monetary data over countries do not necessarily reduce the significance of turnover data for national purposes.

Box 2.6 **Definition of turnover**

> Turnover comprises the totals invoiced by the business during the reference period, and it corresponds to market sales of goods and services supplied to third parties.
>
> Turnover includes all duties and taxes on the goods or services invoiced by the unit with the exception of the value added tax (VAT) invoiced by the unit vis-a vis its customer, and other deductible taxes directly linked to turnover.
>
> It also includes all other charges (transport, packages, etc.) passed on to the customer. Reduction in prices, rebates and discounts as well as the value of returned packaging must be deducted.
>
> *Source:* Eurostat & OECD (2007)

2.4.4 Size classes

2.100   There are two main indicators of the size of businesses: turnover and employment. As noted above, turnover data do not play a significant role in international business demography data collections and turnover is not used as an indicator of enterprise size. The core size indicator in business demography is based on the number of employees. For example, in the European Union four employee-based size classes are applied:

- 0 employees
- 1 to 4 employees
- 5 to 9 employees
- 10 employees or more

2.101   The size class structure has the focus on enterprises with a low number of employees and does not subdivide enterprises with higher numbers of employees. The highest class has the size band of 10 or more employees. The reason for this focus on smaller enterprises is clearly that enterprise births usually have only a few employees, or even no employees. This is also the reason why there is a size class of zero employees. These size classes are part of the international data collections by the EU and OECD and are recommended for other countries too.

Table 2.10 **Boundaries for number of employees**

Employee size class	Boundary in average number of employees
0 employees	0 employees
1 to 4 employees	>0 to < 4.5
5 – 9 employees	≥4.5 to <9.5
10 or more employees	≥9.5 employees

2.102   The size class structure may be adapted for national purposes. In business statistics, data are usually presented in more differentiated size classes for the medium and larger enterprises. Thus, additional breakdowns of the size class 10 or more employees may be introduced. However, any national adaptation should be done in such a way that the resulting data can be transformed to the international standard.

2.103   Data by size class are not only important for the population of active enterprises but also for enterprise births and deaths. An enterprise may grow after its birth and may belong to another size class in the surviving year(s) than in the birth year. However, for analysis purposes, surviving enterprises

should stay in their initial size classes. This is analogous to retaining the initial activity codes of surviving enterprises. As the size class allocation is determined based on *average* employment data, and these values may not be integer, conventions are needed for the boundaries, as shown in Table 2.10 above.

2.4.5 Legal form

2.104   Another variable used in business demography statistics is *legal form*. Legal form can be applied only for enterprises and not for the other types of statistical units. Even in case of an enterprise, the appropriate legal form is not obvious where the enterprise consists of more than one legal unit and two or more of the legal units have different legal forms. For such cases, there is currently no agreed international approach. One option is to classify the enterprise according to the legal form of the legal unit that is the main producing unit. Another option is to use the legal form of the legal unit that has largest employment or value added.

2.105   The choice of a legal form made by a newly created enterprise depends on various criteria. The most relevant are the kind of liability that the entrepreneur wishes to face and taxation and other administrative criteria. These criteria may change over time and thus choice of legal form may vary over time. Furthermore, there might be political or taxation incentives in favour of a specific legal form.

2.106   Another critical issue is that the legal forms may differ considerably between countries, thus reducing international comparability to some extent. For comparability reasons it is therefore advisable to use a classification of legal forms that is quite aggregated. For instance, in European business statistics only three kinds of legal forms are distinguished:

- Personally owned and no limit to personal liability (= sole proprietorship).

- Private or publicly quoted joint stock companies, with limited liability for those owning shares (=limited liability companies and other corporations).

- Personally owned limited and unlimited liability partnerships, including legal forms such as cooperatives, associations, etc. (=partnerships and other legal forms).

2.107   These three categories of legal form are used in the international data collections by the EU and OECD. The Guidelines recommend using them whenever possible. They are not concrete and specific kinds of legal form, but rather constructs based on characteristics of actual legal forms. The mapping of the legal forms in a country to these three types must be done by the NSO itself.

2.5   Key concepts and definitions of high-growth enterprises and gazelles

2.5.1   High-growth enterprises

2.108   In addition to birth, survival and death analysis, a further focus of business demography is the identification of enterprises that have relatively high growth and are therefore intensively contributing to the growth of overall employment. Identification of high-growth enterprises is based on the entire population of active enterprises. High-growth enterprises are defined as the enterprises that (i) satisfy some predetermined threshold that distinguishes them as high-growth over a certain period and (ii) are above a certain size to mitigate any small enterprise growth bias. Growth can, in principle, be measured both in terms of employment (number of employees) and in terms of turnover. Box 2.7 provides a definition adapted from the Eurostat/OECD Business Demography Manual (Eurostat & OECD, 2007). However, in practice, because of lack of international comparability of data in different currencies and purchasing powers, growth is usually measured by the number of employees.

2.109   In Eurostat & OECD (2007), the threshold growth rate is 20%. However, based on results of pilot studies the threshold in international data reporting has been reduced to 10% (for the EU definition see Commission Regulation 439/2014). The main reason for this reduction was that the number of high-growth enterprises is relatively small. This also results in confidentiality problems. The 10% ratio is also recommended in these guidelines.

2.110   This definition thus refers to an average annualised growth rate of 10%, or more, over a period of three years. However, use of a percentage growth rate alone can be misleading because, for example, an enterprise that grows from one to two employees qualifies as high-growth enterprise. Thus, in order not to distort the resulting data set with enterprises that have a small number of employees, a minimum size threshold should be used. In the European implementation, this size threshold is currently 10 employees at the beginning of the growth period. For international comparability, the same threshold is recommended in these Guidelines.

Box 2.7 **Definition of high-growth enterprises**

High-growth enterprises are enterprises having at least 10 employees in the beginning of the high-growth period with an average annualised growth greater than 10% per annum over a three-year period. Growth can be measured by the number of employees or by turnover.

If growth in the number of employees or turnover was due to mergers or take-overs, the enterprise in question should not be considered a high-growth enterprise.

Source: adapted from Eurostat & OECD (2007) to reflect the threshold applied in international data collections.

2.111   It follows from the definition that enterprises born in year *t* should not be included in the group of high growth enterprises before *t+4*. Note also that it is not required that an annual growth of 10% or more occurs in each year of the three-year period; an average annualised growth of 10% over three years is sufficient.

2.112   As stated in section 2.4.1, the activity code of the final period should be used for stratifying high-growth enterprises, as it can be assumed that this activity is the one that produced the growth rather than the activity at the beginning of the observation period.

2.113   An enterprise may be included in the population of high-growth enterprises for several successive years. From a policy point of view, it is interesting to split the high-growth enterprises that are newly entered in the high growth population from those that were high-growth already in an earlier observation period. As indicated in the definition, growth by mergers and takeovers should not be considered as real growth. This means that the calculation of high-growth enterprises must be based on the same longitudinal database of enterprises that was created for business demography statistics.

2.5.2   Gazelles

2.114   Gazelles are a subset of high-growth enterprises, namely those that are up to five years old, as defined in Box 2.8.

Box 2.8 **Definition of gazelles**

Gazelles are enterprises up to 5 years old with an average annualised growth greater than 10% per annum over a three-year period. Growth can be measured by the number of employees or by turnover.

Source: adapted from Eurostat & OECD (2007) to reflect the threshold applied in international data collections.

The basic concept of gazelles is the same as of the high-growth enterprises. In Eurostat & OECD (2007) the threshold of the growth rate for gazelles was 20%, as for high-growth enterprises. Based on results of pilot studies the threshold in international data reporting was reduced to 10% (for the EU definition see Commission Regulation 439/2014) which is also the threshold recommended in these guidelines.

Chapter 3: SBR requirements to support the production of business demography statistics

3.1 Introduction

3.1   In describing the SBR role in supporting the production of business demographic statistics, this chapter aims to build a bridge between the concepts of business demography and the SBR infrastructure. It focuses on how to develop an SBR to facilitate the production of business demography statistics. It describes, recommends and advises on refinements to the SBR Guidelines (UNECE, 2015). The general assumption in Chapter 3 is that an SBR is available, or is in development, and that the SBR is aligned with international recommendations, such as the SBR Guidelines (UNECE, 2015) .

3.2   An SBR that follows the SBR Guidelines (UNECE, 2015) holds the information needed to support the production of business demography statistics. Some of the SBR components, decision rules and concepts - be it the live register, snapshots and historical frame, journal tables, statistical units, demographic events, continuity/survival rules, SBR maintenance and update procedures - are directly linked to the creation of a database to support business demography. However, the structure of an SBR designed primarily to serve business survey programs does not automatically allow easy and efficient generation of business demography statistics. As recording and keeping track of historical events may not be a priority, the SBR may not be organized in a way that provides *longitudinal units*, i.e., units that continue over two or more consecutive reference periods.

3.3   Basically, an SBR supporting business demography must allow (1) retroactive updates to business characteristics – classification, size and statuses, and (2) easy tracking of longitudinal units through time (continuation/survival). A continuing (surviving) longitudinal unit may be linked to more than one statistical unit in the SBR over the course of its history (see Section 2.3.1). Therefore, SBR statistical identifiers cannot, on their own, be used to track longitudinal units.

3.4   The path toward business demography statistics production from the SBR may involve adding a longitudinal component to an existing SBR or including a longitudinal component in the development of an SBR. Countries that do not have a complex SBR are encouraged to develop their SBR with a view to producing statistics, which is a natural extension of the primary role of the SBR.

3.5   Section 3.2 highlights the role of the SBR as it relates to business demography. Section 3.3 presents the concept of longitudinal (business) data in the context of the SBR. The characteristics of longitudinal data are detailed in Section 3.4. Section 3.5 supplements the information on data sources, coverage, statistical units and population that was provided in Chapter 2. Section 3.6 presents the system requirements associated with longitudinal data, including maintenance and update schedule, corrections and changes, business rules, governance and IT considerations.

3.2 SBR and production of business demography statistics

3.2.1 SBR roles – production of economic statistics

3.6   Section 2.1 of the SBR Guidelines (UNECE, 2015) identifies eight roles for the SBR. Role number 5 (SBR based statistics, described in Section 2.6 of the SBR Guidelines) states that the SBR can "play a role in providing more information at less cost" on its own or in combination with other administrative registers or statistical data sources. The SBR holds information that can satisfy business demography

requirements. Indeed, all or most of the information of interest are already found in the SBR. In addition, the SBR offers record linkage facilities enabling linkage with other statistical and administrative data files.

3.7   The SBR Guidelines (UNECE, 2015) recognize that the SBR should be the backbone for business demography statistics production by NSOs. Paragraph 2.51 in these Guidelines states that the SBR is "...*an ideal source in terms of coverage and costs for the production and dissemination of enterprise business demographic statistics.*" Furthermore, throughout countries that have developed their SBRs in accordance to the SBR Guidelines (UNECE, 2015), SBR based business demography statistics will be consistent and comparable.

3.8   However, an SBR developed to serve primarily as a survey frame may not include all the features and information needed to easily and readily generate coherent business demography statistics.

3.2.2   SBR shortcomings in relation to business demography

3.9   The SBR Guidelines (UNECE, 2015) in Chapter 7 provides direction on dealing with demographic events. In the SBR, the outcomes of demographic events (birth, death or survival) are guided by the SBR main roles, which are to serve as a frame and to support NSOs' survey programs. The continuity of statistical units in the context of the production and maintenance of survey frames does not always align with the definition of survival for business demography purposes (as detailed in Chapter 2 above).

3.10   There are likely to be statistical units that are *ceased (death)* or *created (birth)* in the SBR that are not regarded as *deaths or births* in the business demography context. For example, in the case of a merger, the enterprise that arises is a birth for the SBR but not for business demography purposes. In the case of a break-up, the enterprise that is broken-up is a death for the SBR but not for business demography purposes. As such, the SBR population of statistical units must be "cleaned" of births and deaths that are "false" or "spurious" from a business demography perspective. This is an important step to the production of business demography statistics from the SBR.

3.11   It is essential to reassess the outcomes of demographic events in the SBR population from the business demography perspective. Identification of survivals where the SBR recorded births or deaths can be achieved through record linkage, taking into account continuity of production factors. For example, where there is a death on the SBR but a survival from the business demography perspective, linkage between the ceased SBR unit and the SBR unit that is deemed to continue the activity must be recorded in some way. This can be done by creating a table listing the linked units or by assigning longitudinal business identifiers to the units.

3.12   Where an SBR is being developed for the first time, or is being substantially redeveloped, consideration should be given to including components (linkage tables or longitudinal identifiers) that would enable recording and monitoring units that are births or deaths in the SBR but are *longitudinal* units, i.e., continuing from a business demography perspective.

3.13   Other limitations or shortcomings of the SBR should also be considered when developing the infrastructure for the production of business demography statistics. For example, SBR population under-coverage in countries with a large informal sector impacts the quality of business statistics. As such, most factors negatively impacting the quality of the SBR may also negatively impact the quality of business demography statistics created from longitudinal data derived from the SBR.

3.3   Longitudinal business information and a survey oriented SBR

3.14   A longitudinal representation of statistical units' data from the SBR allows identification of longitudinal, i.e., surviving units versus births and deaths, from a business demographic statistics perspective. It also permits the production of business demography statistics over time, through consecutive reference periods, and at two non-consecutive points in time. It enables individual

longitudinal units to be followed through time. Over time, a single longitudinal unit may link to multiple SBR statistical units. As such, this calls for the creation of a new longitudinal units' identification or tracking scheme.

3.15 A business demography database must provide characteristics and size information by reference period from the date of creation to the time of closure of the longitudinal units. During its life, a unit may go through events that modify one or more of its production factors. In the SBR these changes in production factors translate into unit birth, death or survival. They may also reveal fast growing units (gazelles and other high-growth firms), and allow for the identification of fast growing, or declining, sectors or regions.

3.16 In the context of a survey based SBR, the history of statistical units may be created by linking information from a series of tables (live register, snapshots or frozen frames) and journal tables (i.e., records of live register updates) as illustrated in Figure 3.1. Note that Figure 3.1 is a representation of one of several possible structures of a SBR. Assembling longitudinal data from multiple tables, i.e., creating historical views of statistical units, considering impacts of demographic events, may require the use of record linkage methods.

Figure 3.1 **SBR (survey oriented) based demography statistics**
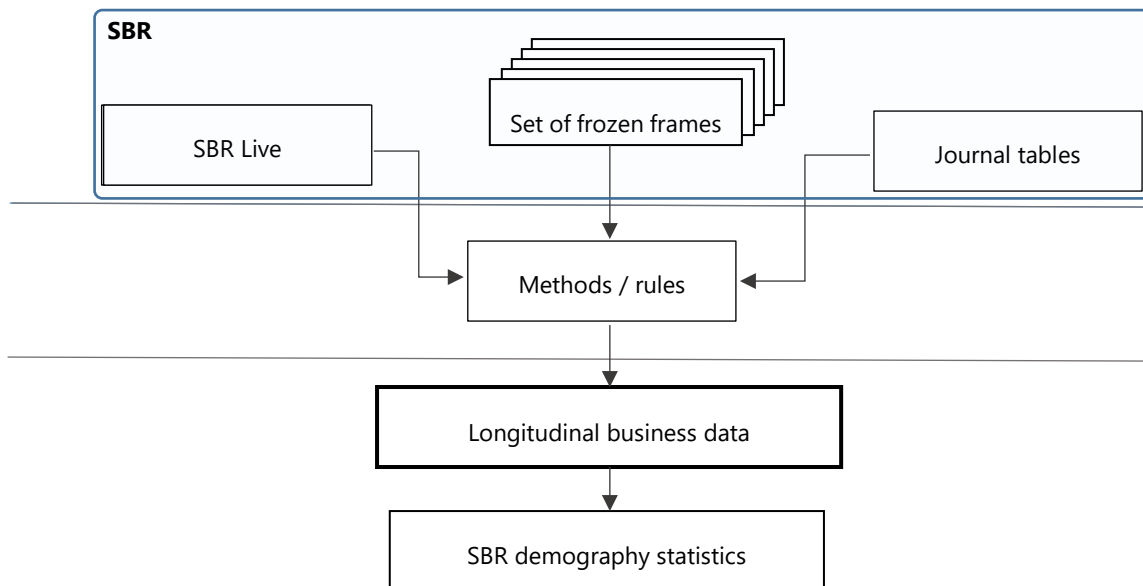
3.17 Morphing different pieces of information from the SBR into a continuous image of the longitudinal units[7] is a complex and time-consuming task. The process cannot be run each time business demography statistics are to be created. Once assembled, the images of the longitudinal units must be stored in business demography tables and updated on a regular basis.

3.18 There are several possible approaches to storing longitudinal data. The preferred approach depends on several factors, including the structure and nature of the existing SBR, the level of development of the SBR, the level of integration of longitudinal unit data within the SBR, and the ease at which SBR demographic events can be translated into birth, death or survival from a business demography perspective.

3.19 The method of storing longitudinal data, be it in a (satellite) longitudinal database or a longitudinal component added to an existing SBR, should be designed with sufficient flexibility to allow

---

[7] Note that not all snapshots are required for each and every statistical unit. There are cases where units keep the same basic information through time.

for the addition of new information, to adapt to new and emerging methods for identification of business survivals (for example, *labour tracking* or use of legal information on predecessors and successors), and to accommodate changes in the SBR components and methods resulting from SBR redesign cycles.

3.20   Some statistical offices (e.g. Statistics New Zealand) have opted to develop a "longitudinal-friendly" business register data structure that essentially eliminates the need to produce and maintain ongoing snapshots or journal tables separately from the live register itself. Under this approach it is no longer essential to create snapshots and use the frozen images to get a complete longitudinal image of a given statistical unit. The live business register itself allows accurate and at will creation of snapshots at any point back in time. These snapshots are not prone to the usual limitations attributed to a combination of journal tables and frozen frames.

3.3.1   Live SBR, snapshots and journal tables

3.21   If the SBR holds a live register, a series of snapshots and journal tables, the initial creation of the longitudinal units calls for the use of all these SBR data, and possibly historical administrative data sources. Statistical units' information included in these data tables need to be linked and organized to construct continuous, from past to current, longitudinal images of the units.

Live SBR

3.22   The live register contains the most up to date image of the SBR population. As stated in paragraph 2.52 of the SBR Guidelines (UNECE, 2015), the live register is refreshed on a continuous basis with the latest information available. The live register includes size and characteristics information for the most recent reference period. Size variables (turnover, employment, assets) may cover a year (fiscal or calendar year), a quarter or a month. The most recent reference period is specific to each administrative data source and to individual statistical units. The Timing of the update to the classification variables (industry, geography, SNA sector, etc.) in the SBR also varies by type of variable and statistical unit.

Snapshots

3.23   SBR snapshots are a series of frozen frames/images of the live register at specific points in time. Snapshots are created on a regular basis (typically monthly, quarterly, or annual) and are kept for a definite period, or permanently. The retention period may vary by data table/type/source, etc.

3.24   The aim of business demography statistics is to describe the developments between different frozen frames in time. There are different approaches to the use of frozen frames in statistical production. Commonly, these images are completely frozen in time, i.e. they are never updated. However, NSOs may choose to update or create versions of frozen frames to correct frame errors, like industry classification for example, or to correct for demographic events like births, deaths, mergers, etc. Each strategy has its consequence and may add or reduce complexity in the creation of the longitudinal database.

3.25   The snapshots are of use in creating the first representation of the longitudinal database. The business demography population, births, deaths and survivals can be obtained by comparing the snapshot for a given reference period with that of the next and previous reference periods.

3.26   Regular information updates, correction of errors, changes in SBR content and coverage, integration of new data sources, and all other SBR changes that impact the live SBR, also related historical SBR information, are not used to update the snapshots. They are frozen. Therefore, the content of the snapshots (coverage, variables included, etc.) differs from that of the live register.

3.27   Thus, while being an important input to the construction of longitudinal data, the frozen aspect of the snapshots limits its usefulness for the construction of the early years of the longitudinal data. Snapshots can, however, serve as the main input to subsequent regular and periodic update of the longitudinal data, as each newly created snapshot is a current image of the live SBR.

3.28   To construct the full and complete longitudinal representation of the SBR units, a third source of information is required, namely journal tables.

Journal tables

3.29   Every change and update to a live register is recorded, with a time stamp, in journal tables. Journal tables are the record keeper of all changes ever made in the SBR. While later values supersede earlier ones in the live SBR, earlier values are not overwritten in the journal tables. A view of the SBR at any point in time can, in principle, be generated from journal tables. For example, for every update to the SBR, journal tables hold, amongst other items, the new and old values, source of the update, reasons for the change, date of correction (in case of correction), and effective date of change[8]. Information stored in journal tables are essential in constructing a historical image of the statistical units.

3.30   Journal tables have limitations as well. New information on a statistical unit for a past period that should have been, but was not, reflected in the snapshot for that period, may not have resulted in an update to the live register and as such will be missing from the journal tables as well. The same may apply to errors and their corrections.

3.31   Therefore, all the information required to create a full and complete continuous image of statistical units may not be available in the SBR, i.e., it is not available from live register, snapshots or journal tables. Depending on the significance of the missing information for business demography, a decision is required whether to accept the information gaps, to make use of other longitudinal data sources, like administrative data files, or to impute missing values based on system rules and accepted assumptions. Decisions to fill, or not, missing information should be well documented. The alternative data sources, rules, and methods used to create synthetic data should be documented.

3.3.2   Creation of longitudinal data

3.32   Once a longitudinal representation of the statistical units has been created (i.e., inception of the longitudinal database) the approach chosen to keep this information up to date is a function of design, method and frequency of updates. The approach also dictates which of the SBR component(s) will serve as input, be it the SBR live register and journal tables, the SBR snapshots and/or administrative data sources.

3.33   There is a clear distinction between establishing a first longitudinal representation and the subsequent maintenance of the longitudinal data. While past/historical snapshots and journal tables are needed to construct the early years of the longitudinal data, decisions on the content, the scope and characteristics of this information should be future oriented, mainly driven by (a) the current and future state of the SBR live register, and (b) the desired business demography statistics.

3.34   Some information included in the live register might not be available for past periods. The information available for past periods may vary between snapshots due, for example, to the addition of new data sources, changes in methods, or following SBR redesign cycles. Historical images may be readily available or easily derivable from data included in the SBR, and this information may go back many years for some variables and types of information. Other data may be available only for a limited number of years, or may have been subject to significant changes in content or format at certain times in the past. These inconsistencies in the content or structure of the snapshots may lead to limitations in the construction of the past.

3.35   The major consideration in establishing a longitudinal database should be the completeness and historical comparability of the population in the SBR snapshots. When developing the first longitudinal data, the statistical office should plan to include enough years (history) to support the production of *flagship* business demography statistics (and entrepreneurship indicators) at an

---

[8] In the context of business demography and the construction of continuous image of statistical units, the preferred date of change shall be the effective date, i.e. the date at which the change occurred in the "real world", not on the SBR database. If the effective date is not available, then a best approximate date can be used.

acceptable quality level. As such, the statistics should cover, at a minimum, a period of six years, i.e. the current reference period and the five previous years. This six-year span allows to produce five-year survival rates, one of the key measures of business demography. A statistical office starting longitudinal data without historical information cannot produce complete business demography statistical output but can generate birth and death rates prior to reaching the six-year span.

3.36   At the inception of the longitudinal data, if a minimum and sufficient number of years cannot be included, the alternative to creating a flagship production is to start with an *experimental* production covering a shorter period.

3.37   Reference period start dates vary by administrative data sources, some going back many years and others becoming available only recently. Therefore, the availability, or otherwise, of historical information for all administrative data sources should not be a major consideration.

3.38   Once established, the longitudinal data will continue to grow. There may be plans to add new data sources, for example through supplemental record linkage to these sources.

3.39   Box 3.1 describes Statistics Finland steps in building business demography data.

3.3.3   Identification of longitudinal units

3.40   The SBR Guidelines (UNECE, 2015) discuss the need for record linkage methods in the SBR. They outline types of matching, name and address standardization, string matching, blocking, etc. Record linkage is primarily used in the context of updating the live register or adding new sources of information to the SBR. Record linkage allows the association of newly received information (e.g., administrative data) with the SBR and transformation of this information, resulting in creation of new statistical units (births) or updates of existing units.

Record linkage methods

3.41   Sound record linkage methods are a prerequisite for building a longitudinal database from the SBR. Data from SBR tables (live register, journals and frozen frames) must be assembled and combined using record linkage techniques and methods to create the longitudinal history of units. Record linkage is used to relate and compare populations of units from subsequent periods. The linkage will identify pairs of units that are common to both periods (survival), units only present in the previous period (death), and units only found in the latest period (birth).

3.42   The recommendations in the SBR Guidelines (UNECE, 2015) on record linkage – based on existing identifiers (e.g., from administrative data sources), or on units' names and addresses – are equally applicable to linking records in the creation of longitudinal units.

3.43   The success rate of a linkage method impacts the quality of the longitudinal data and the resulting business demography statistics. It is advisable to adopt any new methods, techniques or approaches that can improve record linkage in the context of the longitudinal data. For example, legal predecessor/successor data, or employer-employee based record linkages, may help identifying new relations in the population of statistical units[9]. Under the employer-employee linkage approach, a unit may be deemed to continue if a given proportion of its workforce moves from a ceased to a newly created unit in the SBR. Phone numbers or geographic coordinates can be used in record linkage.

---

[9] See Geurts (2016) and Rollin (2013)

Box 3.1 **Three steps from SBR to business demography - example from Finland**

The first step in producing business demography data is to determine enterprise births, deaths and reactivations from annual information on enterprises in the SBR (annual snapshot). This step reads the turnover and employment information for five years prior to the reference year, the reference year and information available for reference year+1. Each unit is considered active in a given year if it has turnover or employment. For each enterprise, years of birth and death are stored in the SBR's business demography table. Reactivations are also detected in this step and are updated to the business demography table by simply removing the previous year of death.

The second step is to detect real enterprise births and deaths. Also, to find cases of survival by take-over, a predecessor or successor may be identified for some enterprise births or deaths. Step 2 builds a panel of data which lists, for each unit by reference year, data on turnover, employment, year of birth and death, other annual information (name, address etc.) and national administrative data.

This panel is used to match newly born units with active and dead units using matching methods specified in the OECD – Eurostat business demography Manual (Eurostat & OECD, 2007) (the same methods are conducted for newly identified deaths). This allows for the identification of births possibly matching with already existing units. In addition, administrative data on reorganization of legal units is used to produce possible matches. Enterprise group register and types of start/cessation from tax sources are used to verify births and deaths.

The results of the matching methods are summarized into distinct tables for births and deaths. The tables essentially list newly born (or dead) units per reference year with all possible matches to other units. The tables also include indicators of the source of the match (administrative, address match, etc.). The most difficult task is inferring which of the matches are believed to indicate continuity, as one new/dead enterprise may match with many enterprises. It checks the administrative data, enterprise group relationships, and leaves out most complex cases. The inferring process is overly strict by choice, and it could be improved by applying more advanced algorithms and new matching methods, such as transfers of local units and employment.

Each match is scored and if considered good, it may be chosen for updating. Administrative data is believed to be the most reliable source and a match from this source is chosen first. A match by address and activity alone rarely indicates continuity. An enterprise with no known direct predecessor and born in an enterprise group may be recorded as 'not real'. The most difficult cases include newborn enterprises in shopping malls, restaurants, barbershops and enterprises in office buildings. The largest enterprise births and deaths are verified and edited manually using the SBR microdata editing tool, which has a separate view for editing data in the business demography table.

The results of the second step are updated into the business demography table in the SBR. Step 2 produces indicators of real births, deaths and ID for predecessor and successor. For each enterprise birth, the birth indicator is set to 0 for other creations, or 1 for "real" births to be included in business demography statistics. The indicator for enterprise deaths is updated similarly. In cases where a newly born enterprise is determined to have a predecessor (i.e. not be a birth for business demography purposes), the enterprise ID of the potential predecessor is recorded. For enterprise deaths, the ID of the successor is recorded similarly.

The third and final step is to create a longitudinal data for reporting. The business demography table is merged with annual information from SBR to link enterprise records. Where an enterprise has been taken over by a "continuator" enterprise, the enterprise ID is set to be that of the continuator, and year of birth is set to be that of the predecessor. The original enterprise ID is recorded as a secondary ID. The resulting table contains annual size variables (turnover, employment), and classification variables (NACE, legal form, 'home municipality in taxation', etc.) for each enterprise. Indicators for survival and employer enterprises are calculated. This reporting table is then used as a basis for all business demography statistics.

3.44   The result of a record linkage may be multiple potential linkages to a single statistical unit. In this case, the linkage process should generate a score, or probability of true match, for each potential linkage. Each record linkage method (based on administrative information, on name, on address etc.) may have its own scoring mechanism, and precedence may be given to one or other method. Inferring which potential linkage is the best is sometimes difficult. It is usually based on a matching score threshold above which a potential link is deemed to be a true match, and below which it is rejected. The decisional process should be fully automated, limiting manual review of potential linkages to significant units and only a limited number of these.

3.45   If significant improvements are made to a record linkage method, the NSO should consider relinking using this new method to identify previously missed linkages or wrongly created linkages. Any improvement to record linkage will lead to a better longitudinal data.

Normalization of the SBR population (dealing with duplication)

3.46   The base population used to construct the first longitudinal database is the SBR live register and/or a snapshot. To identify potential duplication, a prerequisite to building longitudinal data is to scan the base population of statistical units for duplication, if this is not already and systematically done in the SBR. Duplication of units in the SBR can result, for example, from loading information for a given unit from multiple data sources that do not carry the same unit identifier. If good record linkage methods are not in place, information from a second data source about an existing SBR unit may not be associated with the unit, and thus result in creation of a new unit that duplicates the existing unit.

3.47   The *normalization* process (i.e., *removal of duplicates*) is done through a record linking/matching process, where all units in the live SBR or given snapshot are matched with one other. This search for and elimination of duplicated units ensures that newly created units are true births and ceasing units are true deaths. This is essential in constructing longitudinal data. Skipping this step may lead to the duplication of statistical units, creation of spurious births, and inappropriate deactivation of units, and thus biases in the resulting business demography statistics. Identified duplicates should be listed in a table that is used as an input to the identification of business births, deaths and survivals. Possible treatments include the dissolution of one of the duplicated units, or the creation of partnership or family business. If desired, the SBR live register or snapshot populations can be corrected with the results of the record matching.

Steps in identifying births, deaths and survivals

3.48   The concepts of birth, death and survival, and units' continuity rules are defined in Chapter 2, and are mainly based on the Eurostat/OECD Manual on Business Demography (Eurostat & OECD, 2007). Deaths and survivals in a given reference period *t* are identified by matching the population of units in this period *t* to the population of the previous reference period *t-1*. Units are linked based on statistical or other identifiers, unit name(s), addresses and relevant linkage characteristics, using record linkage methods. Pairs of units, present in both periods, correspond to continuing units. Ceased units are those present in the previous reference period *t-1* that do not find a match in the next period (SBR deaths), and the opposite represents the creation of units (SBR births).

3.49   As indicated, there are instances where a birth or death differs between SBR and business demography purposes. For example, a unit recorded as a death in the SBR because of a change in ownership (legal status) may be deemed to be a longitudinal unit, i.e., to continue, for business demography purposes. The process used to update longitudinal data should accept all SBR survivals but check SBR births and deaths as they may be survivals from a business demography perspective.

3.50   If an SBR birth (a new statistical unit) in the current period *t* is linked to a unit in the population for the previous reference period *t-1* it is a potential "false" or "spurious" birth from the business demography perspective. Similarly, if a statistical unit considered an SBR death at the end of period *t-1* is linked to a unit in the population for reference period *t*, it is a potential spurious death from the business demography perspective. The outcome of matching SBR births and deaths to the populations

for the previous and next reference periods respectively can be summarized in one longitudinal table, or in several longitudinal tables.

3.51   A new numbering scheme may be created to more easily identify longitudinal (continuing) units created from multiple SBR statistical units. Whenever possible, the type of demographic event leading to births, deaths, and survivals should be categorized and recorded in the longitudinal tables, along with the best estimate of the date of occurrence. Data on potential links identified during the linkage process not confirmed as true links should also be recorded. Such information may be used to validate the reliability and quality of linkage processes and for future reference, as new information may confirm links that were previously rejected.

3.3.4   Tracking units through time

3.52   For business demography purposes, a longitudinal unit must be tracked over time for as long as it is deemed to continue - from the moment of its creation (birth) to the point where it is deemed to have ceased (death) from a business demography perspective. During its life, a longitudinal unit may be involved in demographic events with other units, and still be considered as a continuing unit (see Section 2.3.1). As previously noted in several places, and again emphasised here, a single longitudinal unit may correspond to more than one statistical unit in the SBR (see Section 2.3.1). Thus, the creation or deletion of an SBR unit identifier does not necessarily correspond to the birth or death of a longitudinal unit. Therefore, SBR statistical identifiers cannot by themselves track longitudinal units.

3.53   Two options for creating and maintaining longitudinal units are presented below.[10]

- The first method involves use of a sequence of concordance tables, one for each reference period, that contain *longitudinal identifiers*, enabling the construction of longitudinal units.
- The second method is based on the use of a new longitudinal unit identifier.

3.54   Longitudinal identifiers, along with relevant business demography information, can be stored in a set of *reference period specific* tables, or into a *single longitudinal data* table.

Reference period specific tables

3.55   In the approach based on a sequence of reference period specific concordance tables, identification of longitudinal units is achieved through the series tables containing period to period pairs of statistical units using the SBR statistical identifiers. Thus, with this approach there are as many tables as there are reference periods included in the database. Each table includes all current to previous and current to current reference period linked statistical units, i.e., continuing from a business demography perspective. A longitudinal unit is therefore represented by a series of linked statistical unit identifiers for consecutive reference periods. To create a longitudinal image of a unit requires that information from the complete series of tables be put together.

3.56   Two such tables are shown in Table 3.1. Each pair of linked units in reference period *t-1* and reference *t*, corresponds to a continuing business, a survivor. New units and ceased units are also included. For a unit that ceased (death) in period *t-1* there is no link to a unit in period *t*, while a new unit (birth) in period *t* will not be linked to a unit in period *t-1*. In other words, each table relates the statistical units at the end of the reference period to those at the start of the period. If there is no start unit, the end unit is a birth. If there is no end unit, the start unit is a death.

3.57   NSOs may decide to assign a longitudinal identifier to the SBR statistical units. A longitudinal identifier may help in the tracking of units through time. In Table 3.1, S1 is a survivor (continuing unit), S2 shows a birth in period *t-1*, S3 a death in period *t-1*, and S4-S6 a continuing unit with a change of ownership in period *t*. S8, S9 and S10 represent the case of a merger at period *t* (S8 and S9 become

---

[10] There are other approaches to track units through time, including methods that combine or derives from the listed options.

S10), while S11 and S12 correspond to an acquisition (S11 ceased existing in *t* and S12 continues). The last two examples, S13 to S17, are representations of a break-up and a split-off, both in period *t*.

Table 3.1 **Reference period specific tables**

Reference Period	Longit Id*	Stat Unit ID at t-2	Stat Unit ID at t-1	...	Reference Period	Longit Id*	Stat Unit ID at t-1	Stat Unit ID at t	...
t-1	L1	S1	S1		t	L1	S1	S1	
t-1	L2		S2		t	L2	S2	S2	
t-1	L3	S3	S3		t	L3	S3		
t-1	L4	S4	S4		t	L4	S4	S6	
t-1	L8	S8	S8		t	L8	S8		
t-1	L9	S9	S9		t	L9	S9		
t-1	L10				t	L10		S10	
t-1	L11	S11	S11		t	L11	S11		
t-1	L12	S12	S12		t	L12	S12	S12	
t-1	L13	S13	S13		t	L13	S13		
t-1					t	L14		S14	
t-1					t	L15		S15	
t-1	L16	S16	S16		t	L16	S16	S16	
t-1					t	L17		S17	
...	...	...	...		...	...	...	...	

* The Longit Id column is optional.

3.58   To facilitate tracking of longitudinal units, especially in the case of split-offs, break-ups, take-overs or mergers, information on the demographic event, which includes the successor of the ceased statistical units, the predecessor of newly created statistical units, the date of the event, and the event type should be recorded. The demographic event information can be store in a distinct table (see paragraph 3.73) or within the identifier table (Table 3.1). This is helpful in identifying the longitudinal unit resulting from the demographic event. The benefit of this approach is that a table can be created for each of the snapshots in the SBR, maximizing the use of the current structure of the SBR, while accepting that the snapshots are not retroactively updated with recent methodological changes and new information on statistical units.

3.59   It is not essential to assign a longitudinal identifier. In the example from Statistics Netherlands (see Box 3.2) a second set of inter-year tables substitutes for a longitudinal identifier.

3.60   Statistics can only be produced for the reference period covered by the tables. A shorter reference period, say monthly, will offer more flexibility than a longer reference period, say, annual. Therefore, the decision regarding the length of the reference period is critical.

Longitudinal multi-year tables

3.61   In an alternative approach, tracking of longitudinal units depends upon a *longitudinal* unit identifier. This longitudinal identifier – a completely new identifier, distinct from SBR statistical unit identifier – is assigned to each statistical unit in the SBR that is associated with a unit in the longitudinal data. The longitudinal identifier enables cross-sectional identification of matching or associated units, and identification of surviving units.

Box 3.2 **Statistics Netherlands' method of tracking business linkages**

This box outlines a practical solution for cases in which the SBR, for reasons of design or administrative updating mechanisms, for example, does not store linking information relating to the survival of statistical units. In this case, a statistical unit that is regarded as continuing longitudinal unit is renumbered between two SBR snapshots.

A source for the linking information can be the results of the matching steps outlined in Chapter 5 of the Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007). These steps are used to discriminate between 'real' births and administrative births caused by administrative renumbering in the SBR: those enterprises that can be linked to a previous enterprise in the SBR are not real births. This information is also useful for establishing a link between two SBR units which are identical but with different identification numbers.

If stored properly, this linking information can be valuable for business demography analysis. Firstly, it allows accurate survival analysis. Secondly, it allows assigning a correct yearly total of turnover or exports to an enterprise that during the year has two or more different enterprise IDs. Assigning correct yearly totals is important for the analysis of growth of individual enterprises.

The approach in Statistics Netherlands is to create two linking tables. The first one is for intra-annual renumbering. The second one is for inter-annual renumbering. The following example shows a continuing longitudinal unit associated with enterprise statistical units A, B, C and D in years 2014 and 2015:

2014	→	2014	→	2014	→	2015
A	→	**B**	→	**C**	→	**D**

This chain can be split up into two parts:
- Intra-annual events (2014):   **A → B → C**
- Inter-annual events (2014 – 2015): **C → D**

To reconstruct this chain, two link tables are needed:
- One within the same year *(for 2014)*
- One between two different years *(transition 2014 – 2015)*

ID_Start *(2014)*	ID *(2014)*	ID_Finish *(2014)*
A	A	C
A	B	C
A	C	C

ID_Start *(2014)*	ID_Finish *(2015)*
C	D

Moving forward or backward in the chain is accomplished by using the enterprise ID_start and enterprise ID_finish columns: Enterprise A ultimately becomes enterprise C (in 2014). Totals, turnover or export value, for example, for all the renumbered enterprises in a year can be obtained by summation of the values for the enterprise ID column (the middle column) and grouping by either the first or last column.

The chain is elongated by the inter-annual table, by coupling the enterprise ID_finish columns of both tables. The complete chain from enterprise A to enterprise D is now restored. If appropriate, further elongation of the chain is straightforward by repeating this cycle, starting with the intra-annual table again.

3.62   Longitudinal identifiers can be stored in a single "*longitudinal identifiers*" table, as illustrated in Table 3.2. Each row in the table associates a longitudinal identifier with an SBR statistical unit identifier and includes, at a minimum, the reference period start date and the reference period end date for which the association is valid. Information about whether a statistical unit's birth/death translate into a business demography birth/death or business continuity may not be conveyed by the longitudinal identifier alone. To this end, other significant business demographic information may have to be added to the table, for example, demographic event with effective date, or included into a distinct table.

3.63   A longitudinal identifier may appear in more than one row in the table. In fact, it appears in a separate row for each SBR statistical unit with which it is associated. With this approach, the creation of the continuous image of a longitudinal unit is performed by extracting and ordering all rows in which the longitudinal unit is listed.

3.64   Table 3.2 shows the same examples as in Table 3.1 but with possible period start and end dates added. L1/S1 is a survivor (continuing unit), L2/S2 shows a birth in January 2013, L3/S3 a death in December 2015, L4/S4-S6 would be cases of continuing units with a change of ownership in January 2016. L8/S8, L9/S9 and L10/S10 represent the case of a merger in December 2016 (L8 and L9 become L10), while L11/S11 and L12/S12 correspond to an acquisition (L11 ceased existing in December 2017 and L12 continues). The last two examples, L13/S13 and L16/S16, are representations of a break-up and a split-off, in January 2017.

3.65   In this simple representation, the longitudinal identifiers seem to serve a limited purpose as there is a direct and one-to-one link between most longitudinal and statistical identifiers. Real life data is more complex and longitudinal identifiers become very valuable in the identification of longitudinal business going through multiple demographic events. Note that an infinite future date is used to represent business continuity, or live longitudinal businesses.

Table 3.2 **Single multi-year longitudinal identifier table**

Longit Id	Stat Unit	Period Start	Period End	...
L1	S1	2008/01/01	9999/12/31	
L2	S2	2013/01/01	9999/12/31	
L3	S3	2008/04/01	2015/12/31	
L4	S4	2008/04/01	2015/12/31	
L4	S6	2016/01/01	9999/12/31	
L8	S8	2008/04/01	2016/12/31	
L9	S9	2010/06/01	2016/12/31	
L10	S10	2017/01/01	9999/12/31	
L11	S11	2009/01/01	2017/12/31	
L12	S12	2007/01/01	9999/12/31	
L13	S13	2008/04/01	2016/12/31	
L14	S14	2017/01/01	9999/12/31	
L15	S15	2017/01/01	9999/12/31	
L16	S16	2009/01/01	9999/12/31	
L17	S17	2017/01/01	9999/12/31	
...				

3.66   Similar to the previous approach, a single table approach should also record the statistical unit identifier of the successor of the ceased statistical unit, predecessor of the newly created statistical unit, date of the event and event type in order to facilitate tracking of units, especially in the case of split-offs, break-ups, take-overs or mergers. The demographic event information can be stored in a distinct table or within the longitudinal identifier table (Table 3.2). The information is helpful to identify the longitudinal unit resulting from the demographic event.

3.67   Assigning a new longitudinal identifier and listing the identifiers in a single multi-year table reduces the number of tables needed to store longitudinal data and offers the flexibility required to

produce business demography statistics for any reference period (month, quarter or year). The assignment of longitudinal identifiers can follow the method used to assign SBR unit identifiers. A longitudinal identifier should be associated with a single and unique longitudinal unit, which in turn can be linked to more than one statistical unit. This applies equally within a reference period and over all reference periods.

Event characteristics table

3.68   As indicated in paragraph 3.63, additional information needs to be recorded to characterize events and associated outcome. Columns may be added in the tracking tables (reference period specific or longitudinal) to indicate whether the outcome of an event is a "true death" or a "true birth". This information can be used to identify business continuity in the context of business demography. Other tables, not listed here, may also be developed. For example, an *event* table listing predecessors and successors, type of event (e.g. merger), effective date, etc. Table 3.3 above is an example of an event table such as shown in Table 3.1 and Table 3.2. Another example would be a table listing longitudinal IDs that have been discarded (fully removed from the business demography population). The longitudinal ID may be removed from the population if, at a later point in time, the statistical unit linked to the longitudinal ID is deemed to be the successor/predecessor of another statistical (and longitudinal) unit.

Table 3.3 **Events table**

Longit Id	Stat Unit	Event	Event Date	BDS_Outcome	RelatedStat Unit	...
L1	S1	Birth	2008/01/01	Birth		
L2	S2	Birth	2013/01/01	Birth		
L3	S3	Birth	2008/04/01	Birth		
L3	S3	Death	2015/12/31	Death		
L4	S4	Birth	2008/04/01	Birth		
L4	S4	Death	2015/12/31	Continuity	S6	
L4	S6	Birth	2016/01/01	Continuity	S4	
L8	S8	Birth	2008/04/01	Birth		
L8	S8	Death	2016/12/31	Death – Merger	S10	
L9	S9	Birth	2010/06/01	Birth		
L9	S9	Death	2016/12/31	Death – Merger	S10	
L10	S10	Birth	2017/01/01	Birth – Merger	S8, S9	
L11	S11	Birth	2009/01/01	Birth		
L11	S11	Death	2017/12/31	Death - acquisition	S12	
L12	S12	Birth	2007/01/01	Birth	S11	
L13	S13	Birth	2008/04/01	Birth		
L13	S13	Death	2016/12/31	Death – Break-up	S14, S15	
L14	S14	Birth	2017/01/01	Birth – Break-up	S13	
L15	S15	Birth	2017/01/01	Birth – Break-up	S13	
L16	S16	Birth	2009/01/01	Birth	S17	
L17	S17	Birth	2017/01/01	Birth – Split-off	S16	
...						

3.69   No matter the approach chosen (a single table as just described, or a series of tables as previously described), a table of longitudinal identifiers should include basic information on the status

of the units (e.g. activity status), on the demographic event/type leading to the creation of a unit or cessation of activity, and all other units' characteristics deemed necessary (for example, operating or legal addresses and names). This is further discussed in Section 3.4.

3.4   Characteristics of longitudinal data tables

3.70   This section details how longitudinal data can be organized and stored in data tables, the location of the tables, their relationship with the SBR, and the need to create summary analysis tables.

3.4.1   Structure of longitudinal data tables

3.71   Two broad types of structures are used in construction of longitudinal data tables from the SBR:

- a series of *reference period specific* data tables; and
- a series of *data specific linkable and longitudinal* data tables.

3.72   The type of structure selected for the longitudinal data tables is guided, among other considerations, by the type of longitudinal identification scheme - reference period specific, or longitudinal.

Reference period specific tables

3.73   Reference period specific data tables align well with a reference period specific identification scheme. Under this approach, reference period longitudinal identifier tables are augmented to include all relevant information available on the longitudinal units for the reference period, see example in Table 3.4. This includes unit identification information such as longitudinal identifier, SBR unit identifier(s), auxiliary variables, names and addresses if required, classification variables (industry, geography…), and size variables (employment, turnover, assets mainly from administrative data sources). To construct the history of a unit or a population, the user must link tables for all reference periods of interest.

Table 3.4 **Reference period specific data tables**

Period	Longit Id*	Stat Unit Id at t-2	Stat Unit Id at t-1	Var A	Var B	Var C	Var D	...
t-1	L1	S1	S1					
t-1	L2		S2					
t-1	L3	S3	S3					
...	...							

Period	Longit Id*	Stat Unit Id at t-1	Stat Unit Id at t	Var A	Var B	Var C	Var D	...
t	L1	S1	S1					
t	L2	S2	S2					
t	L3	S3						
...	...							

* The Longit Id column is optional.

3.74   This is probably the easiest structure for users to query information for business demography purposes. However, it limits the shortest reference period for which data are available to the reference

period of the tables. If tables are created annually, it is impossible to produce quarterly or monthly estimates. The number of tables grows with time and can become considerable if the frequency of creation is sub-annual. Adding new information to the database, or performing retroactive updates to the database, implies that the series of reference period specific tables also needs to be updated.

Information specific tables

3.75   Another approach involves a series of information specific tables each covering the full extent of the longitudinal data, or the full period of availability of the information contained. This approach works well with the approach of storing longitudinal identifiers in a single table (as described in Section 3.3.4). Each data table contains information on a specific issue, source or characteristic. There can be as many tables as there are types of variables or data sources included in the longitudinal data. The central table in this approach is a table linking the longitudinal identifiers to the SBR statistical unit identifiers. Other information specific tables contain, for example, individual administrative data, and classification variables (industry, geography, activity).

3.76   Each record in each table is associated with a longitudinal identifier that links and groups information from the tables of interest.

3.77   A row in a table may cover an unlimited number of reference periods or only a single reference period. The two following cases lead to the creation of a single row for a given longitudinal unit in a table:

- a longitudinal unit does not undergo a demographic event over the course of its life; or

- a longitudinal unit has only one type of activity.

3.78   The following examples (also see Table 3.5 and Table 3.6) present cases where multiple rows in a table are created for a given longitudinal unit. The period covered by a row may also vary by row:

- a longitudinal unit involved in a business demographic event, may be represented in two rows, defined by the event start date and end date;

- a longitudinal unit that had one or more changes in economic activity or in geography;

- monthly or annual size variables for a given longitudinal unit have rows by reference period.

3.79   In short, for any given longitudinal unit, a row is inserted only when a demographic event occurs, a change in a variable of interest is observed, a classification changes, or when size information (mainly from administrative data sources) becomes available for a new reference period.

Table 3.5 **Information specific data tables – size variables**

Longit Id	Stat Unit	Revenue	Period		Longit Id	Stat Unit	Employee	Period
L1	S1	$	t-2		L1	S1	#	t-2
L1	S1	$	t-1		L1	S1	#	t-1
L1	S1	$	t		L1	S1	#	T
....			...		....			...

3.80   This type of data structure is more complex to query than data structured in reference period specific tables. However, it offers more flexibility. Under this structure, the number of tables in the database is relatively stable. Adding new information is simpler and does not impact existing tables. Retroactive updates to existing information, including changes in methods or in classifications, are only applied to the impacted content specific tables.

3.81   Also, tables containing information that does not tend to change frequently, such as geography, industry/activity code, etc., tend to be relatively small. Individual rows may cover multiple reference periods, or the entire history, through the use of a period start and an end dates, as shown in Table 3.6.

Table 3.6 **Information specific data tables – characteristics variables**

Longit Id	Stat Unit	Industry	Period Start	Period End
L1	S1	11	2008/01	2010/03
L1	S1	22	2010/04	2015/06
L1	S1	34	2015/07	9999/12
….				

3.82   Furthermore, the structure allows to produce business demography statistics covering various reference periods, provided the variables of interest are available at the desired frequency. To facilitate data extraction, a set of pre-defined output/analysis tables should be automatically generated.

3.4.2   Location of a longitudinal data table in relation to the SBR

3.83   There are several possible scenarios for the location of the longitudinal data tables. Section 2.10 of the SBR Guidelines (UNECE, 2015) states "*responsibility for and control of a satellite are separated from the SBR and usually take place in a different environment*" on the basis that "*more information can be stored and managed without interfering with the basic functions of the SBR*", and a satellite approach would "*ease the workload on SBR staff and reduce the complexity of the SBR system*". In other words, development and maintenance of longitudinal data tables are a significant undertaking, and the associated burden cannot simply be assumed by the SBR team. New and dedicated resources should be assigned.

3.84   However, there are counter arguments that support a database with closer links to the SBR. First, the SBR and the longitudinal data tables share common data sources, similar rules, concepts and methods. Thus, there is a logic to having the longitudinal database close to, or within, the SBR. Second, it is desirable for the resources dedicated to longitudinal data to possess a good knowledge of the SBR.

3.4.3   Relation of the longitudinal data tables with the SBR

3.85   The inception of the longitudinal representation of the businesses is a significant undertaking. A new, permanent set of longitudinal table(s)/satellite database must be created to store the longitudinal units and their characteristics and size data and to serve as the source to produce business demography statistics.

3.86   There is no single right approach to reach this objective. There are several options for locating this infrastructure in relation to an existing SBR. An NSO should analyse the options and identify the approach that best suits the structure of its SBR. One option is for the new database to be a *satellite* to the SBR, a new product with linkages to the SBR. Another option is *transformation* of the SBR by adjusting and adding components (data tables) to the live SBR so that it can hold a longitudinal component.

3.87   More explicitly, the infrastructure can be:

a) *A satellite of the SBR* - physically separated – at a distance – from the SBR; using SBR rules and/or its own set of rules, update processes, population etc. (see Figure 3.2).

b) *Partially integrated with the SBR* (sharing common resources or input tables); a longitudinal database using same/similar update processes to the SBR, sharing same input data tables for size variables etc. (see Figure 3.3).

c) *Fully integrated with the SBR*; the transformation of the live register (SBR) into a longitudinal live register (LSBR) (see Figure 3.4).

3.88   As noted in paragraph 3.88, the main argument supporting the use of a satellite approach to extend the functionality of the SBR is separation of responsibilities and controls for business demography statistics from those of the SBR, to prevent overtaxing SBR resources. However, technical arguments support the development of a business demography database that has stronger linkages to the SBR. For example, the content of the longitudinal database should match the content of the live register, and updates must replicate most of the SBR rules (including for date and outcome of demographic events, corrections, insertion of new data).

3.89   To maximize efficiency and reduce duplications, the processes for updating and maintaining the live register from new administrative data can be modified to update information in the longitudinal data tables. Integration of longitudinal units within the live register reduces the need to develop new processes specific to the longitudinal database. In summary, building a longitudinal database that maximises linkages to and methods with the SBR is recommended.

3.90   The following three figures illustrate the essence of the options for linkage and/or integration of longitudinal information in the SBR.

3.91   Figure 3.2 illustrates the link from the SBR to a satellite longitudinal database. The satellite database is updated from the SBR but resides at a distance from it. The longitudinal data can be updated from the snapshots, by comparing populations (births and deaths) from period *t* with *t-1*, or alternately directly from the live SBR, using journal tables entries. The snapshots are commonly used to feed information – births, deaths and other demographic events – to the longitudinal database (solid line). Administrative data sources contribute historical size information to complement SBR based data.

Figure 3.2 **Satellite approach**
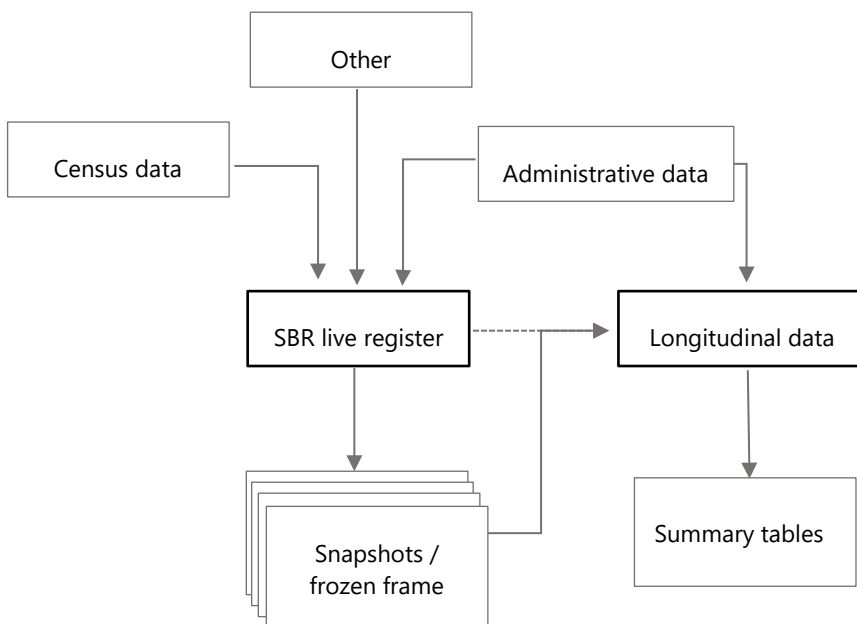
3.92   In Figure 3.3, the longitudinal database is still at a distance from the SBR, but with the difference that some information now flows from the longitudinal database to the SBR. In the illustration, the loading of certain administrative data is channelled first to the longitudinal database, and this information is fed back to the SBR. This is limited to sources where longitudinal database information requirements encompass the SBR data requirements. This approach reduces the number of times the administrative data are loaded. The remainder of the process flow matches exactly Figure 3.2.

Figure 3.3 **Longitudinal data partially integrated with the SBR**
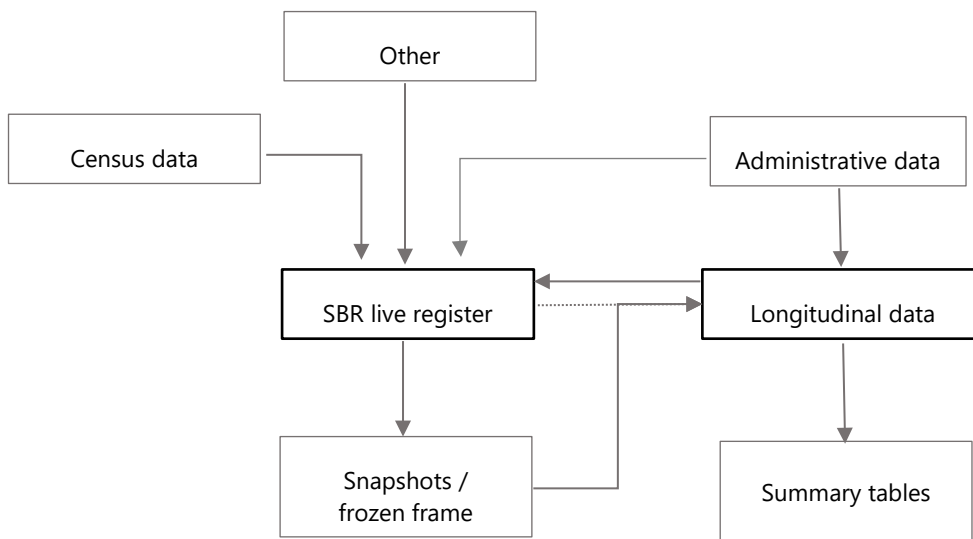
3.93   Full integration of longitudinal data in the SBR (*longitudinal SBR*) could simply mean addition of business demography tables to the SBR, or it may require more significant changes to the SBR. In the presence of a fully functional SBR, the integration of the longitudinal data within the SBR can be done in phases. For example, starting with the development of a functional longitudinal database separate from the SBR, but located in the SBR environment, with a view to eventually share and merge components with the SBR. The level of integration and number of shared components (processes, input tables, etc.) can evolve over time as part of the regular SBR improvement and maintenance work, with a longitudinal SBR as the ultimate goal.

Figure 3.4 **Longitudinal data fully integrated into SBR**
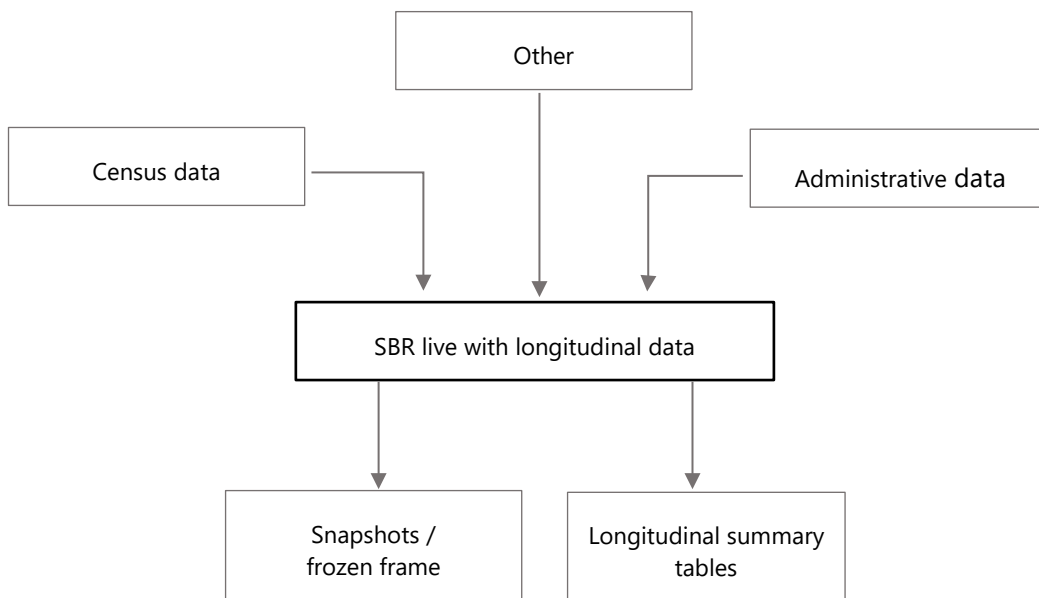
3.94   Box 3.3 illustrates an approach in which longitudinal data are maintained within the SBR itself. No matter the approach selected, the final database must have a longitudinal aspect, align with the SBR Guidelines (UNECE, 2015) and match the content of the live register to the greatest possible extent.

Box 3.3 **Business demography table - example from Finland**

In Finland, business demography statistics are produced from the SBR. A separate longitudinal database is not maintained, as the longitudinal information is integrated into the SBR. Longitudinal microdata and summary reports are produced by using a separate business demography table in the SBR. The table contains variables for:

- Enterprise ID
- Year of birth
- Year of death
- Indicator for enterprise birth
- Indicator for enterprise death
- Enterprise ID of predecessor
- Enterprise ID of successor (continuator)

Therefore, one row in the business demography table represents one period of activity for each enterprise, and consequently one enterprise may have multiple rows in the table. Longitudinal business demography data are produced annually for Eurostat and the business demography table is updated for this purpose. The first two production steps update the business demography table using SBR data as a source. After the updating is complete, a reporting data set is produced. All production steps follow the OECD – Eurostat Business Demography Manual (Eurostat & OECD, 2007).

The annual updating cycle of business demography table corresponds to the most regular user need for longitudinal data. When a longitudinal micro data set is produced, a frozen version of the business demography table is stored for the corresponding business demography. Additional versions can be produced when needed.

3.4.4 Summary tables/snapshots

3.95 The production of business demography statistics from longitudinal data requires linkage of information from multiple tables (reference period specific or information specific). This process can be complex. To help users to produce business demography statistics from the longitudinal database, a set of ready to use summary tables should be created. Their content, coverage, period and rate of refresh should be based on business demography program needs, and international reporting requirements. Users may request copies of these summary tables be saved.

3.96 Longitudinal data should be regularly updated (at least annually). There may be a need to keep snapshots of the longitudinal data and output tables at a set frequency (monthly, quarterly or annual), as well as or instead of historical copies of summary tables. Snapshots consume a significant amount of storage. Restrictions in terms of content, years included, number of copies and duration of conservation of the snapshots/copies of the database should be a function of business demography program needs and the availability and costs of storage.

3.5 Data sources, coverage, statistical units and population

3.97 The appropriate choices in the selection of data sources, coverage, type of statistical units and populations for business demography statistics vary according to the need for international comparability, the administrative and other SBR data sources available, and the need for breakdown by economic activity, geography, etc. Whatever decisions are taken, users must be made fully aware of the criteria used to define the data content and its limitations in producing business demography statistics.

3.5.1 Data sources

3.98   The SBR Guidelines' (UNECE, 2015) recommendations on data sources for the SBR (Chapter 6) also apply to the production of business demography statistics.

3.99   The SBR is continuously updated with the most recent information available from different data sources. Size and other characteristics are refreshed mainly from administrative data from federal/national/regional governments, also from survey feedback and ongoing profiling activities. Information in the SBR live register covers only the most recent reference period for each input data file. Values for previous reference periods may be accessible from the SBR snapshots and journal tables. Information included in the SBR snapshots is frozen. In some instances, for example in the case of errors or batch updates, relying solely on snapshot data may mean using incorrect information.

3.100   To support the business demography, up-to-date historical data are needed from each of the data sources that serve as an input to the SBR or to the production of business demography statistics. Size variables, mostly based on administrative data sources, and characteristics should cover a minimum number of years to be of interest for business demography purposes.

3.101   New and alternative sources, from private companies, financial institutions or web based, may also be investigated with the aim of filling data gaps or complementing existing SBR data sources. Each data source must be evaluated with respect to coverage (current and historical), accuracy, timeliness, acquisition costs, etc. to determine its fitness for use.

3.5.2 Statistical units

3.102   One of the main focuses of business demography is to produce coherent and comparable estimates of birth and death rates, and growth/decline rates by reference period, industry and geography. The feasibility of producing such statistics depends on the availability of characteristics and size variables for the statistical units in the population of interest. As indicated in Chapter 2, a SBR supporting business demography must allow for the production of statistics at the enterprise level. However, the development of the infrastructure to support business demography should take into account the growing demand for sub-national statistics. As such, the aim should be to include all the types of statistical units found in the SBR, even if the statistical unit type(s) at lower level(s) may be missing size variable information.

3.103   Provision for allocation of administrative data at legal unit/enterprise level to lower level units should be included in medium to long term development plans. This allows for the production of business demography statistics at more detailed geographic levels, e.g., for regions and local areas.

3.5.3 Coverage and Population

3.104   In principle, longitudinal data should cover institutional sectors listed in Section 2.3.2 and all economic sectors covered by the SBR. Sector specific shortcomings or limitations, in terms of quality of information and extent of coverage should be clearly stated.

3.105   The same can be said of the population included in the longitudinal data. The SBR population includes all statistical units ever created, that is the units that are alive, temporarily inactive, permanently closed, or created by error, for example in case of duplications or misinterpretation of demographic events. While business demography statistics may focus on active units above a given threshold (for example, at least two employees for economic business demography), there should be no restriction on the initial SBR population included in the longitudinal data.

3.106   A decision may be made to exclude some SBR units, for example, non-employee businesses. However, any exclusion rules should be communicated to, and approved by, the business demography statistics program.

3.6   Longitudinal data system requirements

3.107   As previously noted, longitudinal business data, based on the SBR, should borrow as much as possible from the SBR practices in terms of updating and maintenance procedures, correction of erroneous values in input data sources, business rules and governance, etc.

3.6.1   Regular/scheduled updating and on-going maintenance

3.108   Lack of retroactive updates to snapshots, ignoring of new information for past reference periods, and delay in the information received and recorded are some of the major shortcomings of the SBR in supporting business demography statistics. For business demography purposes, the updating processes of the live register must be expanded to ensure information for the past reference periods is included in the longitudinal data.

3.109   Continuous or daily update of a longitudinal database, for example from an administrative business register event, minimizes the number of records updated each occasion updates are applied and may facilitate the task of categorizing and recording demographic events. However, the longitudinal tables can also be updated on a less frequent basis, to align with the reference period for which demography statistics are created. For example, an NSO producing business demography statistics annually may only update the longitudinal database on an annual basis.

Frequency of updates

3.110   Once created, the longitudinal database must be updated with new information in a similar fashion to the live register, but with the important addition that updates can be retroactive. They can go back to the first reference period of the database or be applied to any reference period from current to past.

3.111   The processes should ensure that *for the most recent reference period* the content of the longitudinal data table is similar to the information found in the live register. Data in the snapshots may differ from longitudinal data because snapshots are frozen while longitudinal data are updated retroactively.

3.112   Depending on the structure and content of the database, and on resources (human and IT), updates can be applied on a continuous basis or periodically, but at least once a year. In fact, for most NSOs, business demography statistics are produced on an annual basis and several of the administrative data files linked to the SBR are also annual files. In such cases, the longitudinal data tables only need to be updated annually.

3.113   If the longitudinal data table must be updated more frequently, it is good practice to plan and organize the updates in alignment with the SBR live register update and maintenance schedule. Efforts should be made, whenever possible, to automate update processes and, where applicable, to utilise SBR processes.

3.6.2   Correction of administrative data and classification variables

3.114   The SBR has quality assurance rules and processes to identify and correct errors and improve the quality of the SBR. Corrections and improvements in the quality of SBR information trickle through to the longitudinal data. Thus, much of what follows in this section is simply good SBR maintenance practice.

3.115   Longitudinal data has its own specific issues not present in the SBR. The correction of errors in longitudinal data can be split by data type, size variables (from administrative and other sources), and classification variables.

Administrative and other data sources

3.116   Treatment and processing of data can include the adjustment of outlier values and imputation for missing records. Treatment and processing of data from administrative or other sources (survey,

census, etc.) are the responsibility of the relevant statistical areas in the NSO. Issues in input data sources identified by the SBR, or other statistical area, through application of quality assessment rules, should be addressed with the data provider and, whenever possible, corrected in the longitudinal data.

Classification variables (industry, geography, …)

3.117   Content, updates and correction of errors in the classification variables (e.g. industry code, geography) may be the responsibility of the SBR. The team in charge of the longitudinal data should address issues in the classification variables that did not get addressed by the SBR, for example issues related to retroactive updates in the data, or unclassified units.

3.118   Classification data required for the construction of the complete history of longitudinal units may not be available for all periods. For example, there may be cases where:

- administrative data are not available for all periods; or

- data gaps exist in the middle of the history of the units.

3.119   It might also be that information in earlier snapshots, covering older reference periods, is based on older versions of classifications. Longitudinal data should always include the most recent version of classifications. When a new version is deployed in the SBR, it should be back casted to the full longitudinal history. In this context assumptions may be required, for example, when an old classification code is split into two or more codes in the new classification.

3.120   Automated rules should be put in place to correct errors and to deal with missing classification values. These rules should be based on agreed principles and methodologies. Data gaps and periods with missing information in the longitudinal data should be filled in with imputed values, whenever there is sufficient information on which to base the imputation and to ensure that it does not create biases in the statistics. New methods based on artificial intelligence/machine learning should be considered for imputing activity codes using, for example, business names, activity description, and administrative data.

3.6.3   Dealing with changes in methods/SBR redesigns

3.121   A mature SBR may have gone through one or several redesign cycles. The redesign(s) or changes in SBR rules, methods, or input data sources may result in changes to the structure or content of snapshots and journal tables. Consequently, a series of snapshots and the live register may hold tables with different formats or populations. Information in journal tables may also be inconsistent over time.

3.122   While it is important to give full consideration on availability and content of snapshots versions, decisions on the content of the longitudinal data (coverage, population, statistical units, etc.) should mostly be guided by the current structure and content of the SBR live register and, where known, any planned or foreseen changes. Since present and future information in the SBR are usually of better quality and richer than in the past, the development of longitudinal data should apply a forward-looking approach, and not be limited by past information.

3.123   Creating longitudinal data with data from tables (including journals) of various formats brings challenges. Changes in SBR methodology, SBR redesign cycles, addition of a new administrative data source, or disappearance of a data source impact the construction of and updates to the SBR. Changes in the content of snapshots or journal tables, availability of new administrative data files, changes in methods or SBR redesign may cause breaks in the longitudinal data series and data inconsistencies. In case of extreme changes, the starting point of the longitudinal data may have to be reset, at the moment of deployment of a major redesign of the SBR, for example.

3.124   NSOs have little control over data collected by taxation offices. Changes to fiscal laws, regulations and taxation requirements can lead to changes in the data included in taxation files. More generally, administrative data may not be available for the whole population or for all reference periods

of the longitudinal database. Changes in the format or in the content of administrative data files may occur from time to time.

3.125   Efforts should be made to minimize breaks in the longitudinal data. Breaks in data series, changes in availability of information, improvements in methods, and any other changes that occur, should be well documented. Users must know the dates or reference periods affected by the changes and the details of these changes, as changes may impact business demography statistics. In summary, documentation of the longitudinal database should: indicate the versions of the classifications used, list any changes in methodology and associated impacts on the data, and provide details on the availability of the data.

3.6.4   Business rules, quality and governance

3.126   A series of business rules should be put in place to ensure that the integrity of the longitudinal data is preserved through time, and errors are not systematically inserted during updates. NSOs should develop rules that are specific to the content and structure of their particular longitudinal data.

3.127   At a minimum, the longitudinal data program should replicate all significant and applicable SBR business rules. However, while SBR rules may help ensure the quality of the longitudinal data they do not suffice. Additional rules are required, for example, to ensure that longitudinal unit identifiers are not duplicated, that dates do not overlap, that the history of units contains no gaps, and that the content of the longitudinal data for the most recent reference period aligns with the content of the SBR.

3.128   On top of business rules, longitudinal data production needs to be supported by a quality assurance program. The quality rules and measures developed for longitudinal data should focus on aspects not covered by an SBR quality assurance program.

3.129   Information included in the governance section (Section 11.3) of the SBR Guidelines (UNECE, 2015) is directly applicable to longitudinal data based on the SBR. The sections of the SBR Guidelines on relationship with users, legislative framework, funding, development, human resources and training can all be applied directly to longitudinal data.

3.6.5   IT considerations

3.130   The SBR Guidelines (UNECE, 2015) in relation to IT considerations also apply to longitudinal data as they share common characteristics with the SBR, including multiple input tables, a live register, a series of snapshots, and journal tables. The IT considerations may however not be as extensive as for the SBR. For example, there may not be a need for a graphical user interface pointing to the longitudinal data.

Chapter 4: Linking the SBR with data from other sources

4.1   Introduction

4.1   The SBR plays a central role in linking data from surveys and administrative data. Data linked at the micro-level can address a wide variety of research questions and policy issues. Analyses can be performed using data already available in the NSO and thus not requiring any additional reporting burden on enterprises. By combining data about businesses with data about entrepreneurs the performance of businesses (for example measured in terms of profit, turnover, investment, and trade volume) can be related to characteristics of the corresponding entrepreneurs.

4.2   An SBR includes data for a limited set of variables, mostly those of direct use for survey frames. The key SBR variables are economic activity, classification of this activity, location and size measures. These characteristics are relevant for sampling applications. The SBR also holds information about demographic events such as mergers, acquisitions, split-offs, changes of ownership, leading to births, deaths and survivals of the longitudinal units that are the basis for business demography statistics.

4.3   Linking SBR data to administrative data and other data sets at the micro level – hereafter referred to as *micro data linking* (MDL) – can provide additional insights into the relationship between business characteristics and business performance. Data from sources with a different unit of observation (for instance international trade data) must be transformed to SBR statistical units. Another dimension can be added by linking to data on entrepreneurs, thus being able to correlate entrepreneurship characteristics with business characteristics and performance. These data can provide insights into life-cycle events of entrepreneurs and businesses which are relevant for policy making.

4.4   The statistical units in the SBR are typically based on (sets of) legal units, which are covered in one of more administrative sources. The SBR provides the links between administrative and statistical information about businesses. However, when administrative data are used for statistical purposes it is important consider the differences between statistical concepts and the concepts used in administrative sources. Administrative data must be transformed into statistical data using an appropriate method. The relationship is illustrated in Figure 4.1 (reproduced from the SBR Guidelines (UNECE, 2015), Chapter 8).

Figure 4.1 **Relationship between statistical and administrative universes**
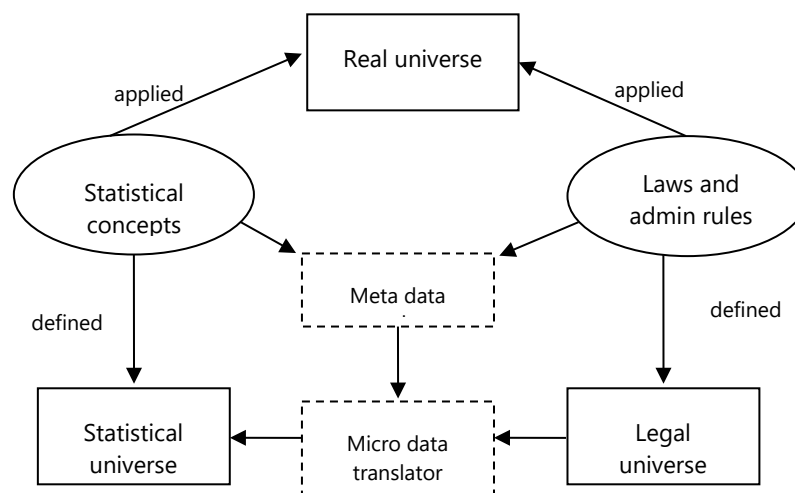
4.5   There is rich documentation[11] prepared by Eurostat describing the overall aims, methodologies and results of the MDL projects completed in the past years and providing indications of work ahead.

4.6   This chapter deals with several aspects of MDL, all focussed on the central position of the SBR in linking data from various sources, including other statistical registers and administrative sources, covering both businesses and persons. A brief review is given of methods for dealing with combining incomplete datasets and ensuring consistency with published data.

4.2   SBR as the core source

4.7   The core source for MDL analysis is the SBR, meaning that only units present in the SBR are part of the analysis. For example, Figure 4.2 shows the pivotal position of the SBR in coupling data on entrepreneurs and business performance.

4.8   The data available on businesses or entrepreneurs may be – at least partly – based on surveys that do not cover the entire population. In particular, micro level data for analysing economic performance (i.e. productivity) is often obtained from surveys that aim to determine macro-economic totals and hence that use stratified random sampling with a focus on larger enterprises. Thus, micro level data are available only for a small percentage of smaller enterprises.

4.9   It is recommended that the best available snapshot of the reference year in the SBR is used in a micro-data linking exercise.

Figure 4.2 **The SBR as the core in linking additional micro-data sets**

---
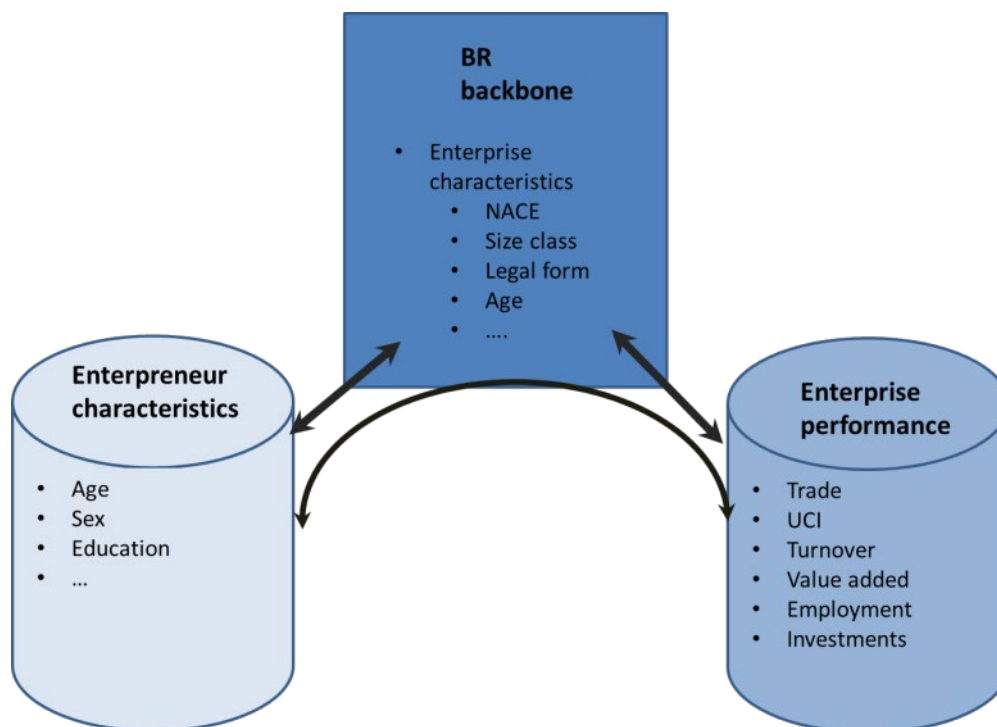
11   See   http://ec.europa.eu/eurostat/statistics-explained/index.php/Microdata_linking_in_business_statistics_-_introduction

4.3   Data on enterprises

4.3.1   Examples of micro data linking

4.10   This section provides an overview of the datasets likely to be available, and the options to be considered for MDL at the enterprise level, which leads to indicators on entrepreneurial performance of enterprises. Examples are provided in Box 4.1 to Box 4.5. The following paragraphs discuss in more detail practices in linking to some of the most relevant enterprise data sources that are likely to be available.

Box 4.1 **Use of structural business statistics**

In the EU, structural business statistics (SBS, see http://ec.europa.eu/eurostat/web/structural-business-statistics/overview) describe the structure, conduct and performance of enterprises from the European Union. The SBS database provides basic information on the enterprise economic performance (e.g. turnover, production value, and value added) as well as information on employment, size class and other specific variables. SBS data are organised according to the NACE Rev. 2 activity classification.

SBS data are typically based on surveys, administrative data (e.g. tax records), or a combination of both. Where SBS are based on surveys, exhaustive micro-data will not be available for the entire population.

The most commonly reported SBS characteristics in an MDL exercise are turnover, value added, employment variables (e.g. number of persons employed, number of employees and/or full-time equivalent), wages and salaries. Other variables can also be used depending on purpose of the study. Purchases of goods and services may also be included (such as in services trade by enterprise characteristics (STEC) statistics, for example).

Box 4.2 **Micro data linking initiatives of Eurostat**

Eurostat in cooperation with member states has been very active in MDL in the recent years. Several MDL projects have been developed:

- Trade in goods by enterprise characteristics (TEC), which started as an MDL project and is currently a regular production in Eurostat.
- Trade in services by enterprise characteristics (STEC).
- Projects linking structural business statistics (SBS) to trade in goods statistics (ITGS) as well as other business statistics.

An overview of past projects is provided in Alajääskö (2015). Future plans include the extension of MDL work to include trade in services as well.

Box 4.3 **Current BDTEC project: Linking Business Demography to Trade by Enterprises Characteristics**

A current Eurostat project involves linking business demography, SBS, and international trade in goods statistics (ITGS) data to provide insight in the relationship and relevance of several business demography characteristics with respect to international trade, turnover, employment and value added.

Business demography breakdowns include: age, high growth status, NACE and size-class. Moreover, from foreign affiliates statistics (FATS) and SBR-data a group status for each enterprise is determined, which indicates if this enterprise is dependent or belongs to a (international) group.

Box 4.4 **Understanding business performance and dynamics: example from Australia**

The Australian Bureau of Statistics (ABS) has created a Business Longitudinal Analysis Data Environment (BLADE - see https://archive.industry.gov.au/Office-of-the-Chief-Economist/Data/Pages/Business-Longitudinal-Analysis-Data-Environment.aspx) that is best described as a series of integrated longitudinal datasets. The BLADE was initially funded by the Department of Industry, Innovation and Science (DIIS) to enable Australia's participation in the OECD's project on employment dynamics (DynEmp). It is now recognised as one of the ABS's high-value enduring statistical assets. The BLADE uses the ABS business register as the integrating spine, and integrates directly collected survey data and transformed administrative data from the Australian Taxation Office and other sources for all active businesses in the Australian economy from 2001-02 to 2014-15. This enables analysis of businesses over time and the micro-economic factors that drive performance, such as innovation, job creation, competitiveness and productivity. The BLADE improves the evidence base for policy development and reform.

Box 4.5 **The Eurostat/OECD entrepreneurship indicator program**

Eurostat in cooperation with the OECD has developed the Entrepreneurship Indicator Program (EIP) (see http://ec.europa.eu/eurostat/web/structural-business-statistics/entrepreneurship/indicators) that collects important indicators on enterprises and their economic performance.

The TEC framework (see Box 4.2 and Section 4.3.2) provides data for the EIP indicator on the export performance of small enterprises. Another indicator, originally identified by the EIP, is the export performance of young enterprises. Data for such an indicator are not yet widely available, but there are ongoing projects (such as the BDTEC project described in Box 4.3) aiming - among other things - to produce such indicators.

4.3.2   Linking trade in goods and services data to the SBR

4.11   International trade in goods and services is a relevant parameter for globalisation research and provides insight in how domestic firms are engaged in the global economy. Trade data are interesting for an MDL exercise, not only from the analytical viewpoint but also from the data source viewpoint. Several aspects of international trade can be addressed, depending on data availability. Interesting aspects of trade are not only total value of foreign trade, but also a breakdown by product and partner country.

4.12   Conventional international trade statistics offer a picture of trade flows between countries, broken down by types of goods and services. While this is an important input for trade analyses, these data do not offer insights into the actors, or the types of businesses, that are engaged in cross-border trade.

EU Programs TEC and STEC

4.13   In the EU, trade in goods data with non-EU countries is comprehensively available (by product and partner country) through customs data. However, for reasons of minimizing administrative burden, there are limitations in the case of intra-EU trade: no breakdown of goods and partner countries is available for intra-EU trade for enterprises with trade value below a threshold, the value of which depends on the specific country.

4.14   Unlike traditional trade statistics, international trade in goods by enterprise characteristics [12] (TEC) aims at describing the structure of trade by characteristics of the trading enterprises, for instance by their economic activities, their size or concentration of trade.

4.15   TEC is based on linking international trade in goods statistics (ITGS) micro data with SBR information, allowing a deeper analysis of international trade by including dimensions such as economic activity, size-class of enterprises, trade concentration, geographical diversification and products traded.

4.16   TEC is an example of an MDL project that matured into regular data production. TEC has been part of Eurostat's regular data collection since 2009. TEC provides very interesting and important statistics on the structure of trade by characteristics of the trading enterprises. As regards entrepreneurship, a particularly interesting TEC statistic is the enterprise size class information, which can be used to provide entrepreneurship indicators of trade performance by size class.

4.17   Eurostat and several countries in partnership have also developed a methodology for services trade by enterprise characteristics (STEC). STEC is an ongoing project that takes forward the development work done in the TEC framework to cover services as well. In November 2017, the Compilers Guide for statistics on Services Trade by Enterprise Characteristics (STEC) (Eurostat & OECD, 2017) was published. The aim of the STEC project is to produce a database on international trade in services by enterprise characteristics. STEC data provide valuable analytical information on the characteristics of enterprises involved in the international trade of services, by breaking down traditional services trade statistics by the trader's size, ownership and main activity. STEC data are compiled by linking the survey frame of services traders' micro-data to the national SBR.

4.18   Together, STEC and TEC will provide key insights on international traders engaged in cross-border trade of goods and services by enterprise characteristics.

4.19   STEC may be quite an important data source for further MDL work. One hypothesis is that young enterprises are more likely to engage in trade in services than trade in goods (the assumption being that - except for resellers of goods - trade in goods necessitates developing a production process, which may be difficult for a young enterprise). Although STEC itself is an MDL project, trade in services data have not been included in other, recent MDL projects. Thus, inclusion of trade in services data in the future MDL projects is noted as an interesting next step forward.

Traders vs. non-traders

4.20   From the point of view of analysing the drivers of entrepreneurship, a goal of any MDL project linking trade data is to be able to split the entire population into (international) *traders*, i.e., enterprises that take part in international trade (of goods and/or services) - being an exporter, importer or two-way trader - and *non-traders*, i.e., enterprises that are active on domestic markets only.

4.21   Exporters (including two-way traders) are of special interest for policy makers because of their potential job creation due to demand from markets abroad. Importers (again including two-way traders) are important since they access themselves (or make available for others) raw materials, intermediate goods, services and technologies or final products that may not be available in a national market, or may be available at better quality/lower price in an international market.

4.22   However, care should be exercised when deciding if an enterprise is a trader or not. If all enterprises that happened to sell (or buy) some goods - regardless of the value - on an international market are included in the "traders" group, then traders will also contain enterprises which are not trading because of their business model but just happened to sell (buy) something by chance, or for specific, one-off reasons (e.g. an one-off sale of used machinery abroad). Allowing such "occasional" traders to be counted as traders may cause difficulties in interpreting the results.

---

[12]   See: http://ec.europa.eu/eurostat/web/international-trade-in-goods/data/focus-on-enterprise-characteristics-tec .

4.23   To be able to focus only on enterprises that trade because of their business model, assumptions must be made and thresholds applied. Further study may be needed to explore and quantify the effects of the chosen thresholds. The country situation and economic field/market should also be considered. To establish all aspects that are relevant for all countries in a way that does not bias results is quite a challenge. In any case, in the context of entrepreneurship indicators, setting a threshold, no matter how experimental, to determine if an enterprise qualifies as a trader or not is preferable to defining all enterprises that happen to buy or sell as traders.

4.24   The above approach means that the group of non-traders will also exhibit some (relatively low) amounts of export (or import) activity in the reference year. This is the sum of the activities of enterprises whose export (import) value was lower than the threshold. If there are trade values which cannot be linked to a known enterprise, then an "unknown" or "non-allocated trade" group can also be created. Thus, consistency with total export (import) values in international trade databases can be ensured as well, i.e. the total trade is the sum of trade by "traders", trade by "non-traders" and non-allocated trade amounts. It is recommended that the concepts of trader vs. occasional trader and non-trader as defined above are explained to the users, for example, when disseminating results on indicators.

Fiscal units and statistical units

4.25   Trade data often refer to *fiscal units*, where a fiscal unit is a combination of one or more legal units whose composition is optimized for fiscal purposes within the fiscal/legal framework at a certain point in time. For further use in MDL analysis these data must be linked and subsequently transformed to the level of an SBR enterprise. Only after these steps can trade characteristics be accurately and correctly linked to other enterprise characteristics.

4.26   In practice, coupling the data from the tax authorities with SBR data is straightforward, with one-to-one or many-to-one linkages in many cases. More complex linkage cases and problems are described in *Compilers guide on European statistics on international trade in goods by enterprise characteristics (TEC)*, Chapter 5 (Eurostat, 2016). These include the following:

Intra-annual business demography changes

- Intra-annual demography events are challenging as monthly available trade data must be assigned to enterprises with characteristics that apply for a full year.

Large and complex businesses

- The linkage may not always provide the expected outcome. Trade flows may be allocated to enterprises whose characteristics seem to be different from economic reality or incoherent with other statistics. These problems are more likely to exist for large and complex businesses.

Incomplete business register data

- A link is established to an SBR unit, but its information is incomplete, for instance lacking economic activity code or size class.

Treatment of estimated trade data

- *Estimated data* refers to non-response, or value below threshold, in trade data. If an estimated value can be coupled to a trader and subsequently to an enterprise this value should be used.

Non-established traders

- Non-established traders are foreign companies that carry out trade transactions in the reporting country and are registered for value added tax (VAT) or have appointed a tax representative. Such entities without a permanent residence in the reporting country may not be included in the SBR.

VAT-groups

- In the EU, under certain circumstances, several VAT declarants may report VAT as a group. In this case, VAT is recorded using only one VAT number. The contents of a VAT group may correspond to more than one enterprise. In these cases, where there is a 1-to-*n* (or even a *m*-to-*n*) relationship between VAT identification numbers and enterprises, the allocation of trade over different enterprises has to be approximated. Suggested approaches include:

  a) allocation of all trade to the dominant enterprise;

  b) allocation of trade associated with specific product codes to the enterprises with corresponding economic activity codes.

4.3.3 Linking group status, size class and foreign ownership

4.27   Enterprise size class information and the status of the enterprise group to which an enterprise belongs are essential as entrepreneurship indicators as they allow comparisons between micro, small or medium, and large enterprises, and between *independent* enterprises and enterprises belonging to an enterprise group (*dependent* enterprises). Currently, in business and trade statistics in the EU, size class information is based on the number of persons employed by the enterprise. There is no information on whether an enterprise belongs (or not) to an enterprise group.

4.28   Group status is very important for analysis as it is well known that enterprises that are part of a group (especially enterprises that are part of a multinational group) may behave in a different way than enterprises that are not part of a group. In the EU, for policy purposes, this is a relevant distinction and is part of Eurostat MDL-projects (e.g. BDTEC project "Linking Business Demography to Trade by Enterprises Characteristics").

4.29   Foreign ownership is an interesting characteristic as foreign controlled enterprises may exhibit specific behaviour, for example different employee characteristics, or higher value added. Foreign ownership of an enterprise group is determined by the country of residency of the ultimate controlling institutional unit (UCI) as outlined in Eurostat (2012). This characteristic is subsequently assigned to all underlying enterprises of the enterprise group. In the absence of information indicating foreign control of a group, it may be assumed that the group is domestically controlled. The results of such an approach obviously depend on the robustness and completeness of the data sources used for the determination of the UCIs.

4.3.4 Linking with R&D and innovation data

4.30   Expenses on investments, R&D or innovation activities may influence enterprise performance. To be able to compete in a dynamic international setting it is vital for enterprises to develop new products and/or more efficient means of production.

4.31   By linking international trade data, foreign ownership data[13], and innovation/R&D/investment data on enterprise performance (for example from SBS), a comprehensive analysis can be performed on the effects of international orientation. Examples of indicators that can be produced include the average R&D expenditure of traders/non-traders and the CIS status of traders/non-traders.

4.32   This data is invariably collected at the level of the enterprise, which allows for straightforward coupling via the SBR. Of course, data are available only for a sample of the survey population. Solutions to this problem are discussed in Section 4.5.

---

[13]   For example Eurostat's R&D and Community Innovation Survey (CIS) datasets: http://ec.europa.eu/eurostat/web/science-technology-innovation/overview

4.4 Data on entrepreneurs

4.33 This section focuses on information about entrepreneurs available from surveys or administrative data sources. It first defines *entrepreneur* and then describes possible data sources, the characteristics that may be collected from these and their population coverage.

4.4.1 Definition of entrepreneur

4.34 All persons who own and control an enterprise with the intention of generating income from it are considered entrepreneurs. This is in line with the resolution concerning statistics on work relationships adopted by the 20th International Conference of Labour Statisticians (ILO, 2018). This resolution introduces the category of "independent workers", who are those who own and control an economic unit for which they work:

> *"Independent workers own the economic unit for which they work and control its activities. They make the important strategic and operational decisions about the economic unit for which their work is performed and the organization of their work, are not accountable to or supervised by other persons, nor are they dependent on a single other economic unit or person for access to the market, raw materials or capital items. They may work on their own account or in partnership with other independent workers and may or may not provide work for others. The category of "independent workers" in the classification of status in employment provides the best starting point for the identification and compilation of statistics on entrepreneurs."*

4.35 The definition of independent workers is therefore not limited to persons who are classified as self-employed in SNA. Thus, owner-operators of the following legal forms are in scope for statistics on entrepreneurs:

- *Sole proprietorship enterprises.*

- *Partnership enterprises;* all owners of a partnership enterprise are considered entrepreneurs.

- *Corporations;* for example, a limited liability company.

4.36 One can argue that the owner-operators of an incorporated enterprise are dependent on the income generated by the enterprise and act in in a similar way to the owners of an unincorporated enterprise. They control the enterprise. In a strict legal sense, they do not own the enterprise as they only own the shares. However, the income generated from the enterprise is, in a similar way, depending on the success of the enterprise. In conclusion, owner-operators of limited liability corporations are not different from owners of unincorporated businesses - they manage and (indirectly) own the enterprise, and their remuneration is dependent on the success of the enterprise. Therefore, this group is considered as entrepreneurs. However, it is recommended to treat them as a separate sub-group so that users of statistical information can select the specific group of interest, depending on their research or policy interests. Box 4.6 describes the situation in the Netherlands concerning the treatment of owner-operators in administrative sources and in business statistics.

4.4.2 Sources of data on entrepreneurs

4.37 The first entry point for data on entrepreneurs is an administrative register or, in some countries, a trade register maintained by a chamber of commerce or by another government institution. One of the goals of a trade register is to create transparency, and thus to identify enterprises and their owners. Coupling data, using an identifying key for data collected in a civil register maintained by the government, enables a wide range of information about an entrepreneur to be brought together. Furthermore, using data from the tax authorities, additional financial information become available. Data available from a labour force survey (LFS) provide insights on several aspects of the positions of individuals in the labour force. However, as an LFS is a sample-based survey, it only provides answers for a part of the population, and it is time dependant: LFS data may not be available for the time period

on which research focusses. Finally, administrative data from the social security database may provide insights into employer history of the entrepreneur (if any).

Table 4.1 **Characteristics available on entrepreneurs by linking administrative sources and surveys**

Subject/Population	Source	Characteristic
Inhabitant	Civil Register	Address, age, sex, marital status, parents, children, nationality, place of birth
Tax-payer	Income Tax Register	Amount of income, source of income (employment, entrepreneurship, capital)
Labour force population	Labour force survey	Educational level, hours worked (as entrepreneur)
Employee	Social security administration	Employee status, employment before, during or after entrepreneurship, Characteristics of: • The employment contract, e.g. permanent, temporarily, working hours, salary • The enterprise of the employer, e.g. size, NACE-code.

4.38   Table 4.1 above provides an overview of possible data sources for entrepreneurs and the characteristics that may be obtained from these data sources. An entrepreneur may be included in one or more of the sources and is not necessarily included in all of them.

Box 4.6 **Owner-operator of a limited liability company in the Netherlands: employee or entrepreneur?**

Statistics Netherlands considers an owner-operator of a limited liability company as an entrepreneur and not as employee even though there may be a labour contract and part of his/her income from the limited liability company is taxed as wages.

For taxation purposes a fictitious wage must be paid by a limited liability company to the owner-operator. This fictitious wage (at least €45000) is taxed as wages and avoids tax arbitrage with respect to the choice of the legal form. In general, it is more profitable to transfer the profits from the enterprise to the owner by means of dividends rather than wages. This is not possible for non-limited companies and thus, without the requirement to pay a fictitious wage, owners of enterprises with a specific legal form would be favoured.

With respect to taxation, owner-operators of limited liability companies are defined as persons who have a decisive vote (with or without their partner or spouse) at a general meeting of shareholders with regard to the continuation of their labour agreement with the company.

For this reason, owner-operators are not entitled to unemployment benefits and do not pay social charges for unemployment.

In conclusion, owner-operators are, for taxation reasons, treated as employees, but are in the social security system not considered as employees. Moreover, in daily practice, owner-operators are considered entrepreneurs just as are owners of non-limited companies. Also, responses to the labour force survey show that owner-operators classify themselves as entrepreneurs, not as employees.

4.39   Business demography events like births, deaths and survivals can be linked to characteristics of the entrepreneur to analyse determinants of business success. In Annex 2: Profiling entrepreneurs by linking business and social statistics in Denmark, a Danish example is presented describing newly established enterprises and their survival, both in relationship to entrepreneurial characteristics.

4.4.3   Populations of entrepreneurs

4.40   For various reasons, the populations of the data sources do not entirely coincide. This section focusses on situations where the population of entrepreneurs from the SBR is different from the populations from the data sources in Table 4.1.

4.41   Reasons why the SBR population is smaller:

- The main factor limiting an entrepreneur population derived from the SBR is that the enterprise must be recorded in the SBR. Differences between SBR based numbers of entrepreneurs and numbers in socio-economic publications are often due to this.

- Numbers of entrepreneurs in socio-economic publications are typically based on a survey with self-assessment of the characteristic *self-employment*.

- Where there is no obligation for an entrepreneur to register in the trade register or other administrative source used for the SBR, the number of entrepreneurs in the SBR will be lower than the number that declare themselves as self-employed.

- No economic activity is observed and therefore an entrepreneur, although registered, is not included in the SBR. This may occur because the enterprise has:

    - no employees; and

    - is exempted from payment of VAT because turnover is below a certain country specific threshold.

- An entrepreneur does not have to be a resident of the country in which his/her enterprise operates. In the situation where an entrepreneur resides in country A, but his/her enterprise is in country B, the enterprise is not recorded in country A and therefore the entrepreneur is not considered to be an entrepreneur in country A.

4.42   Reasons why the SBR population is larger:

- In LFS, the classification of employed or self-employed is based on the main source of income - employment or entrepreneurship. Thus, an entrepreneur may be classified as self-employed. In the SBR-based approach all persons who are owner-operators are considered entrepreneurs regardless of whether they have income from employment.

- In the LFS, high and low age cut-offs are applied in the definition of persons who are regarded as belonging to the labour force population. In the SBR there are no such limitations.

- The labour force population in a country includes only those persons who are resident in that country.

4.4.4   Linking administrative sources to produce data on family businesses

4.43   Family businesses make up a significant share of enterprises in many countries and there may be an interest to stimulate entrepreneurship by family businesses by supportive legislation. Statistical research is needed to develop statistical indicators to provide the necessary information for policy makers. In particular, there is presently no internationally accepted definition of a family business. However, as described in Box 4.7 a definition has been adopted by the EU, and work has been done in bringing together data for family businesses, as outlined in the following paragraphs.

Box 4.7 **European family businesses**

Many countries have a national association to promote the interests of family businesses. Some of those are members of European Family Businesses (EFB)[14] which is the EU federation of national associations representing long-term family owned enterprises, including small, medium-sized and larger companies.

In 1997, the "Groupement Europeén des Enterprise Familiales (GEEF)", defined the concept of a family business that was officially adopted by the European Commission. This definition includes family businesses that have not yet gone through the first generational transfer. It also covers sole proprietors and the self-employed, providing there is a legal entity that can be transferred.

The European Commission has adopted a common European definition of family businesses. According to this definition, a firm, of any size, is a family business, if:

- The majority of decision-making rights are in the possession of the natural person(s) who established the firm, or in the possession of the natural person(s) who has/have acquired the share capital of the firm, or in the possession of their spouses, parents, child or children's direct heirs.
- The majority of decision-making rights are indirect or direct.
- At least one representative of the family or kin is formally involved in the governance of the firm.

A listed company meets the definition of a family business if the person who established or acquired the company (share capital) or his/her family or descendants possess(es) 25 per cent of the decision-making rights mandated by their share capital.

This definition includes family businesses that have not yet gone through the first generational transfer. It also covers sole proprietors and the self-employed, providing there is a legal entity to which the business can be transferred[15].

4.44   For a natural person, who is a regular employee in an enterprise, the business, the family and the ownership of the business are independent systems. However for a family business, these systems are strongly connected. In one of the first overviews of a family business as a specific organizational form, published by the Harvard Business School[16], a family business is visualized in terms of overlapping subsystems, resulting in seven possible sectors, as shown in Figure 4.3. A family business operates in Sector 7.

4.45   The multidisciplinary and complex roles of the family members may in some cases hamper a family business in being successful. Possible hampering factors are conflicts between family members or family interests being mixed with business interests. Also, lack of a strategy or a succession plan could be a reason that a family business does not survive. In some family businesses, problems are caused by the absence of skills, education or experience within the family. On the other hand, within a family business, parents spend part of their lives teaching their children how to run a business.

Figure 4.3 **Model of family business systems**
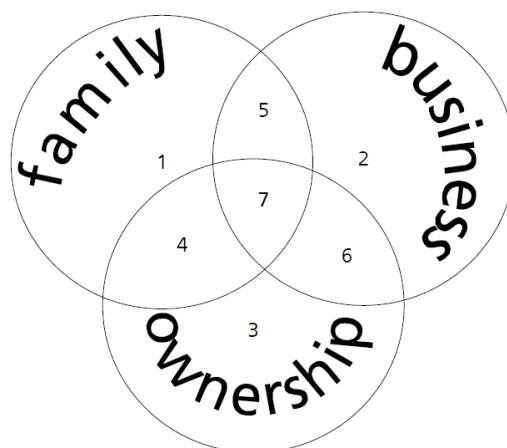
4.46　　Annex 3: Family businesses in the Netherlands describes how Statistics Netherlands developed a satellite register to identify family businesses in the SBR.[17]

4.5　　Data from surveys – limitations and solutions for MDL

4.47　　When a data source is linked to the SBR, data is not necessarily available for all SBR units. This is especially likely to be the case for smaller enterprises in the case of sample surveys for which the focus is to obtain a complete picture of the largest most dominant enterprises, smaller enterprises being included with low probability. An overview of available methods to deal with this problem is given by (De Waal, 2015). Two of the methods are described below.

4.48　　Cross tabulation of characteristics derived from two different sources relies on the overlap of units in the sources. Missing data cannot usually be ignored because they are non-random and because low overlap - especially likely in the lower size-classes - leads to large uncertainties.

Repeated weighting

4.49　　Repeated weighting is a technique that was developed at Statistics Netherlands to solve the problem of inconsistencies among tables of estimates based on multiple data sources. A separate set of weights is assigned to individual sample units for each cross table for which population totals have to be estimated. The tables that need to be estimated are estimated sequentially using in each block as many sample units as possible to keep the sample variance as low as possible. The combined data from administrative data sources and surveys are divided into rectangular blocks. The availability of data determines how a table is estimated.

　　a.　Data from an administrative data source which covers the entire population can simply be counted.

　　b.　Data from surveys are weighted by means of regression weighting. In that case weights must be assigned to all units in the block to be weighted. For a survey, one usually starts with the inverse inclusion probabilities of the sample units. These weights are then further adjusted by calibrating them to previously estimated totals. For a data block containing the overlap of two surveys, the usual approach is to begin with the product of the standard survey weights from each of the surveys as starting weight for each observed unit, and then to correct these starting

---

[17] The method to characterize an FB in the SBR is based on a best practice of Statistics Netherlands to link information of different administrative sources to the statistical units in the SBR.

weights by calibrating to totals known from administrative data sources and/or previously estimated totals.

4.50   When estimating a new table, all cell values and margins of this table that are known or have already been estimated for previous tables are kept fixed, i.e. the regression weighting is calibrated on these known or previously estimated values. This ensures that the cell values and margins of the new table are numerically consistent with previous estimates.

4.51   Repeated weighting was developed with the estimation of a set of related frequency census tables in mind. In principle, the same estimation strategy can also be used for tables containing quantitative variables, such as "income". In Houbiers (2004) more details on repeated weighting are provided. There are advantages and disadvantages of repeated weighting.

4.52   *Advantages and disadvantages of repeated weighting:*

> \+ micro-data consistency, e.g. relations between variables of the same unit, is maintained.

> \- generally unsuitable for estimation of small sub-populations.

Box 4.8 **Repeated weighting example**

> A brief example from social statistics illustrates the repeated weighting procedure.
>
> Assume there are two partly overlapping survey data sets: one containing information on the "educational level" of persons, and one containing information on the "working hours" of persons. In the repeated weighting approach, lower-dimensional tables are in principle weighted before higher-dimensional ones. Hence, the repeated weighting approach would, for instance, start with estimating the table for "educational level". To estimate this table, a standard weighting procedure can be applied. The table for "working hours" can also be estimated by means of a standard weighting procedure. Next, the table "working hours x educational level" can be estimated by calibrating on the estimated numbers for "educational level" and "working hours". In this way, previously estimated values are preserved when making estimates for a new table.

Mass imputation

4.53   The mass imputation approach imputes values for all variables for all units in the population where no value was observed (see, e.g. Whitridge, Bureau & Kovar (1990), Whitridge & Kovar (1990), Shlomo, De Waal & Pannekoek (2009)). This leads to a rectangular data set in which the entire population is represented and for all variables a value is available, either observed or imputed. The imputations are generated by means of an (explicit or implicit) imputation model. Mass imputation makes more use of auxiliary variables for developing imputation models compared to survey weighting (including repeated weighting) and can be made robust to outliers and other potential errors in the covariates. After mass imputation, estimates for population totals can be obtained by simply counting or summing the values of the corresponding variables. In this case the population is the entire SBR population.

4.54   Mass imputation is often based on advanced imputation methods that satisfy edit restrictions on the data, such as the methods developed by Pannekoek, Shlomo & De Waal (2013); Kim, Reiter, Wang, Cox & Karr (2014); and De Waal, Coutinho & Shlomo (2017).

4.55   The mass imputation approach relies on the ability to capture all relevant variables and relevant relations between them in the imputation model, to estimate the model parameters sufficiently accurately. Given that all relevant variables and relations among them can be captured accurately by the imputation model, the approach is very straightforward. When the relationships between variables also hold for sub-groups of the population it is also straightforward to compute totals for these sub-groups summing up the values for the specific variables of these sub-groups.

4.56   *Advantages and disadvantages of mass imputation:*

+ multivariate imputation methods make good use of all the information in the linked data.

+ very flexible method, as different imputation methods may be used for different variables.

+ suitable for estimation for small sub-populations, especially when appropriate imputation models are used.

- a lot of work can go into the specification and checking of imputation models.

- mass-imputed datasets can easily be abused by analysing relationships that were not sufficiently accounted for in the imputation models.

- it is often impossible to capture all relevant variables and relations in the imputation model, simply because there are not enough observations to estimate all model parameters accurately, which implies that relations in the imputed data will be spurious and do not reflect the relations in the population.

Chapter 5: Other approaches to produce business demography statistics

5.1   Introduction

5.1   The previous chapters describe the infrastructure and procedures required to produce business demography and entrepreneurship statistics using the SBR. However, some countries have not developed an SBR to the extent that facilitates the production of business demography and entrepreneurship statistics. To offer alternatives to these countries, this chapter provides examples from three countries, Mexico, South Korea and the United States, that have produced business demography or entrepreneurship statistics based on data sources other than the SBR, namely censuses, surveys, administrative registers, and combinations of these. All three examples describe briefly the statistics that are produced, and the data sources and methodology used.

5.2   In Section 5.2 the Mexican example demonstrates the production of business demography statistics by using data collected in economic business censuses held every 5 years. In this example, the statistical unit is the establishment. Section 5.3 explains the production of business demography statistics in South Korea by linking various sources of administrative data. The example discusses demographic events such as births, deaths, and survivals of businesses and high growth and gazelle enterprises. In this example, the enterprise is the statistical unit. In the example of the United States in Section 5.4 data are collected in an annual survey on entrepreneurship combined with administrative records to produce enterprise and establishment level statistics that describe the characteristics of businesses and business owners.

5.2   Mexico: Study on business demography statistics based on economic censuses

5.3   Mexico has held an economic census every five year since 1930. These data have helped to construct the statistical infrastructure from which business demography statistics are generated. The observation unit is the establishment. Data are obtained on all sectors of the country's economy, except the agricultural sector. During the economic census data collection, all the establishments, both formal and non-formal, are identified on field through a door to door visit in all urban areas in the country, for which an initial directory, obtained from the SBR, is used. During such tours, the data of the surviving establishments are updated, detecting births and deaths. The observation unit for the economic censuses and the SBR is the establishment, which carries out its economic activities in a permanent place.

5.4   This section gives a general overview of a study conducted in Mexico to generate business demography statistics, describing general characteristics of the study, the methodology employed, the results obtained, a comparison with other studies, and conclusions[18, 19].

---

[18] The complete results, as well as more details related to the methodology from this study can be consulted in Spanish on:
http://www.inegi.org.mx/inegi/contenidos/investigacion/Experimentales/Esperanza/default.aspx

[19] A summarized version in English of the methodology from this study can be found at:
https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.42/2015/Session_I_M%C3%A9xico_-_Business_Demography.pdf

5.2.1 Purpose of the study

5.5   The purpose of the study was to obtain survival, mortality and life expectancy indicators that express the demographic dynamics of the population of establishments in Mexico, using applied demographic techniques on establishments identified by Mexico's economic censuses.

5.2.2 General characteristics of the study

5.6   There is a SBR in Mexico. Its objective is to be the backbone of all the economic statistics production processes, as well as the basis for longitudinal and demographic studies. Nevertheless, since it is still in an initial stage of development, the business demography study presented here was based on the results of the economic censuses. The five-year periodicity of the economic censuses has allowed for the construction of a large statistical infrastructure, since there are economic data for each establishment that has existed or exists in the country. For this study, results from six economic censuses, carried out from 1989 to 2014, were used.

5.2.3 Methodology

Year of birth declared by the establishment's respondent as the basis of the study

5.7   In this study, the year in which each establishment started to operate (year of birth) is obtained from the establishment respondents and is used as the base variable. As year of birth is embedded in the unique identifier given to the establishment by the tax administration office, the year reported is considered reliable. The year of birth variable is collected every five years for each establishment in the economic census' target population. Continuity rules between censuses (used to identify surviving establishments, or to determine if an establishment is considered as a birth given a change in its legal name, address and/or economic activity) were not applied, because there was no consistent inter-census unique identifier for each establishment. This may have resulted in increases in the numbers of births and deaths relative to what is required for business demography statistics. Implementing continuity rules under these circumstances was practically impossible given the resulting complexity and associated high costs. Since the creation of the Mexican SBR in 2009, a unique identifier is given to every establishment, so continuity will be easier to assess in the future.

5.8   Although continuity rules were not applied an exercise to estimate the accuracy of the results produced by the method developed in this study was carried out. In the exercise, the results obtained from the study were compared against the results obtained by tracking individual establishments (i.e., by applying continuity rules) during the period 2009-2014. It was found that the differences between the two sets of results were acceptable. A description of this comparison exercise is included in Subsection 5.2.5.

5.9   The year of birth values from the economic censuses allow calculation of how many establishments from a given generation survive and how many die. For example, if in the 1994 economic census x establishments from the 1983 generation (meaning births in 1983) are counted, and if in the 1999 economic census y establishments from the same 1983 generation are counted, this implies that y establishments from the 1983 generation survived the 1994-1999 period (arriving at 16 years of age), while at the same time, x-y establishments from the 1983 generation died aged 11 to 15 years.

5.10   A follow-up of several generations of establishments based on the year of birth enabled the construction of survival and mortality tables for the establishments in Mexico at different levels of disaggregation. With these tables it is possible to answer questions such as the following: How many establishments survive at a specific age? How many establishments die before reaching a certain age? How many years is an establishment expected to live since its birth? How many more years is an establishment expected to live once it has reached a certain age?

Formation of generations of establishments

5.11   Based on the six economic censuses from 1983 to 2013, 31 generations of establishments were formed, one for each year. Each generation includes all the establishments born in the same year. The six censuses included about 20 million records of establishments.

5.12   For each generation, counts of survivors and non-survivors (deaths) were obtained based on the information collected every five years in the economic censuses. For example, for the generation of establishments born in 1983, the number of survivors was obtained every five years, in the period 1989-2014, based on the reported year of birth. In this way, a generational count of survivors table was constructed for the 31 generations considered.

Calculation of survival and death probabilities and life expectancy of establishments

5.13   Using the generational count of survivors table, a table of quinquennial proportions of survivors was established for each of the 26 generations from 1983 to 2008. The quinquennial proportions of survivors approximate the probability of an establishment of age x reaching age x+5. It was not possible to calculate the quinquennial proportions for the 2009-2013 generations because of lack of data.

5.14   Using these generational quinquennial survival proportions, annual survival probabilities were estimated for each of the 26 generations by interpolation. By using the annual survival probabilities, for each generation the number of surviving establishments for each age x was calculated, from an initial population of 100,000 establishments. Next, an average number of survivors was calculated for all 26 generations, obtaining a count of survivors by age, which summarizes the behaviour of all generations taken together. These averaged counts formed the basis for the construction of a survival function, which estimates the number of survivors at any age. The survival function S(x) is a simple, predictive linear regression model in which the number of surviving establishments is estimated as a function of the natural logarithm of the age of the establishments.

5.2.4   Results obtained

5.15   Using the survival function mentioned above, survival and mortality tables were calculated with the following structure:

Age x	S(x)	p(x)	q(x)	d(x)	E(x)

x: Age of establishments
S(x):   Number of surviving establishments at age x
p(x):   Probability of an establishment surviving through age x
q(x):   Probability of an establishment dying before reaching age x
d(x):   Number of establishments dying before reaching age x
E(x):   Life expectancy at age x.

5.16   Several survival and mortality tables were calculated for different categories in three dimensions: a) geographical (33 categories: one at national level and 32 at state level); b) economic (4 categories: one at global level and three for each economic sector: manufacturing, trade and private and non-financial services); and c) establishment size in terms of number of employees (13 categories: one at global level and one for each stratum of number of employees). In this way, 1,716 survival and mortality tables were calculated. The survival and mortality tables calculated at the national level, using an initial population of 100,000 establishments (covering all economic activity sectors and all strata of number of employees) are shown in Table 5.1.

Table 5.1 **Survival, mortality and life expectancy table for businesses at the national level**

x	S(x)	p(x)	q(x)	d(x)	E(x)
0	100000	0.6672	0.3328	33282	7.8
1	66718	0.4794	0.5206	18774	8.2
2	47944	0.4209	0.5791	5855	8.6
3	42089	0.3793	0.6207	4154	9
4	37935	0.3471	0.6529	3222	9.4
5	34712	0.3208	0.6792	2633	9.9
6	32080	0.2985	0.7015	2226	10.4
7	29854	0.2793	0.7207	1928	10.9
8	27925	0.2622	0.7378	1701	11.4
9	26224	0.247	0.753	1521	11.9
10	24703	0.2333	0.7667	1376	12.5
11	23327	0.2207	0.7793	1256	13.1
12	22070	0.2091	0.7909	1156	13.7
13	20914	0.1984	0.8016	1070	14.3
14	19844	0.1885	0.8115	996	15
15	18848	0.1792	0.8208	932	15.7
16	17916	0.1704	0.8296	875	16.5
17	17041	0.1622	0.8378	825	17.2
18	16215	0.1543	0.8457	781	18.1
19	15434	0.1469	0.8531	741	18.9
20	14694	0.1399	0.8601	705	19.8
21	13989	0.1332	0.8668	672	20.7
22	13317	0.1268	0.8732	642	21.7
23	12676	0.1206	0.8794	615	22.7
24	12061	0.1147	0.8853	589	23.8
25	11472	0.1091	0.8909	566	24.9

Source: INEGI Economic Censuses 1989, 1994, 1999, 2004, 2009, 2014.

5.17 The life expectancy of an establishment E(x) is defined as the expected number of years that an establishment will live after having reached an age of x years. It can be calculated based on the survival function. Table 5.1 shows that life expectancy grows with the age of the establishment.

5.18 A study similar to the one described here was carried out five years earlier by considering 5 censuses, right after the 2009 census, in which 21 generations were processed. It was extended to include data from the 2014 census, obtaining very consistent survival and death probabilities, and life expectancies of establishments, which gives more validity to the study's methodology.

5.2.5 Comparison study incorporating continuity rules

5.19 An additional exercise aimed at assessing the validity of the survival and mortality tables generated in the study carried out. It involved a follow-up study using the Eurostat-OECD continuity

rules for each establishment in the 2014 economic census, identifying the surviving establishments in the field, collecting their updated data, and detecting deaths and births of establishments.

5.20   Specifically, for the 2014 Economic Census, the initial directory was extracted from the Mexican SBR, and during the 2014 in-field operation, each establishment was individually monitored through an SBR-based unique identifier assigned to each establishment in 2009, thus enabling the recording of demographic events. The following events were defined based on the SBR's recorded updates for period 2009-2014:

- **Survivors**: establishments detected in 2009 that were collected again during the 2014 Economic Censuses, applying the continuity rules recommended by Eurostat and OECD.

- **Births**: establishments that initiated activities after 2009 or were declared as new establishments, according to the continuity rules.

- **Deaths**: establishments that were active in 2009 but subsequently ceased operations; according to the continuity rules, it can also be determined that an establishment collected in 2009 is no longer the same entity, meaning that it has died and a new establishment has been born.

5.21   It was considered that a new establishment is born from an existing one (which thus has died) when at least two out of its three following characteristics have changed: company name; address; or economic activity. In this manner, continuity rules-based numbers of survivors and deaths for the 2009-2014 period were obtained. The figures were compared against those generated from the survival, mortality and life expectancy tables for the same period. The differences observed were acceptable.[20]

5.2.6   Conclusions

5.22   Some findings from the study outlined in this section are:

- In all survival and mortality tables it is observed that as age increases, the death rate decreases, while life expectancy increases.

- Establishments of small size (according to the number of employees) tend to die earlier than large ones.

- Survival behaviour differs among economic sectors. For example, the manufacturing sector has higher life expectancy; trade is the sector with the lowest life expectancy.

- As of certain age x, establishments are increasingly likely to survive and stabilize over time; this observed stability is reached at earlier ages as businesses grow larger.

- The survival and mortality tables generated by using the methodology outlined in this section, produce demography statistics of reasonable quality, and are a good alternative when application of continuity rules is not possible because of infrastructure and cost limitations.

---

[20] For further details, see http://www.stat.go.jp/english/info/meetings/wiesbaden2016/pdf/6-05_mex_pap.pdf

5.3 South Korea: Business demography statistics based on administrative data

5.23   Statistics Korea (KOSTAT) has compiled and published business demography statistics since 2012. It has generally followed the production methods suggested by the Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007) and the SBR Guidelines (UNECE, 2015), with some adjustments to reflect the special circumstances in South Korea.

5.24   This section presents an overview of the production of South Korean business demography statistics in terms of the collection of basic data, linking data, determining population, computing provisional births and deaths of enterprises, confirmation of demographic events and some 2016 results of the business demography statistics for South Korea.

5.3.1   Collection of basic data

5.25   KOSTAT produces business demography statistics by linking different administrative data. Also, since 2016, South Korea has been developing an SBR, based on which business demography statistics will be compiled in the future.

5.26   The administrative data being used for business demography statistics are provided to KOSTAT by relevant government agencies including the National Tax Service (NTS) and the South Korean Court of Justice, see Table 5.2. The NTS provides monthly records of business registration in South Korea as well as yearly records of value added tax, closures of businesses, registrations of corporations, mergers, split offs, etc.

Table 5.2 **Administrative data for the business demography statistics in Korea**

Source	Type of data	Data items	Usage
NTS	Business registration	Business registration no., corporation registration no.	Linkage key
		Name of company, name of representative, address, etc.	Testing the continuity rules
		PIN of representative	Disaggregation by gender and age
		Code of type of main business activity, administrative district code	Disaggregation by industry type and region
		Date of business registration	Verifying business entries
	Value added tax	Revenue (tax base, tax exempted business)	Verifying revenues
	Statement on wage and salary income payment	Number of employees	Verifying number of employees
	Closure of business	Date of closure	Verifying business exits
Court of Justice	Merger/Split off	Merging entity and merged entity, survived entity and newly established entity by split off	Verifying the birth and death of businesses (change of governance structure)
	Registration of corporation	Business type	Verifying the birth and death of businesses (conversion into an incorporation)

5.3.2   Linking data

5.27   KOSTAT currently uses the 10-digit Business Registration Number (BRN) and the 13-digit Corporation Registration Number (CRN) for the identification of statistical units (enterprises) and data linkages. The original administrative use of the BRN is for taxation, and that of the CRN is for corporation registration.

5.3.3 Determining the target population

5.28   The target population is the entire population of active, for profit enterprises in South Korea, including both employers and non-employers, with the exception of enterprises in agriculture. An active, for profit enterprise is defined to be a for profit enterprise that has business turnover (of any amount) and/or more than one employee. Non-profit enterprises can be identified by a non-profit enterprise code on their BRN. Enterprises whose industry classification is indexed as A, O, T, or U are excluded from the population.

5.29   The statistical unit for producing business demography statistics is the enterprise. The definition of an enterprise depends on the legal form. In the case of a corporation, the enterprise includes the headquarters and the branch offices/factories. For a sole proprietorship (or privately-owned enterprise), the individuals who are registered at the NTS for the purpose of business registration are defined as enterprise units.

5.30   The business registration data from the NTS provide basic datasets for the administrative business demography statistics of South Korea, which included 8 million records of registered businesses and 5.55 million records of active for-profit enterprises as of 2015. The active for-profit enterprise population comprised approximately 550,000 incorporated enterprises, approximating 10% of the total population, and 5 million sole proprietorships (or privately-owned enterprises), approximating 90% of the total population.

5.31   KOSTAT is considering the compilation of business demography statistics based on the population of employer enterprises and that of economic enterprises, according to the Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007) which uses an employee threshold to define the scope of the population of active enterprises.

5.3.4 Screening provisional births/deaths of enterprises

5.32   KOSTAT compares the BRNs within the target population of active for-profit enterprises by year to identify provisional births and deaths of enterprises.

5.33   The provisional births of enterprises can be identified by comparing the BRNs in the population of active for-profit enterprises in year t with those in year $t-1$. The provisional births of enterprises are identified as enterprises whose BRNs are only present in year t. The provisional deaths of enterprises are identified as enterprises whose BRNs are only present in year t and not in year $t+1$.

5.34   KOSTAT also uses the continuity rules to identify provisional births and deaths of enterprises. An enterprise is considered as a provisional birth if at least 2 of the 3 continuity factors have changed in year t compared with those in year $t-1$. Those with 2 or more changes in year $t+1$ compared with t are considered as provisional deaths.

5.35   The business sector in South Korea is characterised by frequent entry and exit of businesses due to the large share of privately owned businesses. Considering this characteristic, a "$t\pm1$ approach" is used in handling enterprises that become dormant or restart, as described in the next paragraph.

5.3.5 Confirmation of demographic events

5.36   Based on the screening of provisional births and deaths of enterprises, further steps are taken to identify the actual births and deaths of enterprises according to the Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007). The Manual suggest eliminating merger, split-off, the change of ownership, break up, take-over, etc. from the populations of the provisional enterprise births and deaths to obtain the births and deaths of enterprises for business demography purposes. However, under the $t\pm1$ approach used in KOSTAT, an enterprise dormant for more than one year is a death and if it becomes active again it is a birth. This is a departure from the Eurostat-OECD methodology.

5.3.6 Business demography statistics for South Korea, 2016 results

Active Enterprises

5.37   In 2016, the number of active enterprises was 5.776 million, up by 222,000 (4.0 %) from 2015.

- *Number of enterprises by legal form*. When broken down by legal form, the 5.776 million active enterprises in South Korea were divided into 585,000 corporations (10.1% of the population) and 5.191 million sole proprietorships (and privately-owned businesses) (89.9% of the population).

- The statistical units of corporations and of sole proprietorships (and privately-owned business) are defined differently. The statistical units of corporations are those of enterprises, thus including headquarters and branches, but sole proprietorships use the legal unit (BRN).

Enterprise births

5.38   In 2016, the number of enterprise births was 876,000, up by 63,000 from 2015.

- *Birth rate*. In 2016, the birth rate of enterprises stood at 15.2%, up 0.6% from 2015. According to the legal form, the birth rate of the corporate sector was 11.7%, lower than for sole proprietorships.

- Birth rate (%) = (Number of enterprise births / Number of active enterprises) x 100

Enterprise deaths

5.39   In 2015, the number of enterprise deaths was 640,000, down by 137,000 from 2014.

- *Death rate*. In 2015, the death rate of enterprises stood at 11.5%, down 2.5% from 2014. According to the legal form, the death rate of the corporate sector is 7.2%, lower than for sole proprietorships.

- Death rate (%) = (Number of death enterprises / Number of active enterprises) x 100

Table 5.3 **Number of active enterprises, enterprise births and deaths by legal form (to nearest thousand)**

	Active			Births			Deaths	
	2015	**2016**		**2015**	**2016**		**2015**	
		count	proportion (%)		count	birth rate (%)	count	death rate (%)
Total	5,554	5,776	100.0	813	876	15.2	640	11.5
Sole proprietorship	5,005	5,191	89.9	747	808	15.6	601	12.0
Corporation	548	585	10.1	66	69	11.7	39	7.2

Note: Some sums do not correspond with the totals due to the rounding process.

Number of enterprises by industry

5.40   Within industry, in 2016, Accommodation and Food Service Activities (Table 5.4, Section I) had the highest birth rate at 20.6% and death rate at 17.7%, and a high proportion of sole proprietorship.

Table 5.4 **Number of active enterprises, enterprise births and deaths by industry (thousand)**

ISIC 4	Active			Births			Deaths	
	2015	2016		2015	2016		2015	
		count	proportion (%)		count	birth rate (%)	count	death rate (%)
Total	5,554	5,776	100.0	813	876	15.2	640	11.5
Manufacturing (Section C)	495	503	8.7	50	49	9.7	40	8.2
Construction (Section F)	366	385	6.7	40	43	11.3	27	7.3
Wholesale and Retail Trade (Section G)	1,347	1,377	23.8	202	209	15.2	174	13.0
Transportation (Section H)	520	535	9.3	44	49	9.2	33	6.3
Accommodation and Food Service Activities (Section I)	788	811	14.1	161	167	20.6	139	17.7
Real Estate Activities and Renting and Leasing (Section L)	1,108	1,175	20.3	13	14	15.7	114	10.3
Professional, Scientific and Technical Activities (Section M)	141	152	2.6	21	25	16.3	14	9.7
Membership Organizations, Repair and Other Personal Services (Section S)	264	278	4.8	38	43	15.6	29	11.1

Survival rate

5.41   The 1-year survival rate for enterprises born in 2014 and still active in 2015 was 62.7%, up by 0.3% from 2014. The 5-year survival rate for enterprises born in 2010 and still active in 2015 was 27.5%, up by 0.2% from 2014.

Table 5.5 **Survival rate of born enterprises by reference period (%)**

	1-year survival rate		2-year survival rate		3-year survival rate		4-year survival rate		5-year survival rate	
in 2015	(2014)	62.7	(2013)	49.5	(2012)	39.1	(2011)	32.8	(2010)	27.5
in 2014	(2013)	62.4	(2012)	47.5	(2011)	38.8	(2010)	31.9	(2009)	27.3
in 2013	(2012)	60.1	(2011)	47.3	(2010)	38.2	(2009)	32.2	(2008)	29.0
in 2012	(2011)	59.8	(2010)	46.3	(2009)	38.0	(2008)	33.4	(2007)	30.9
in 2011	(2010)	60.0	(2009)	46.9	(2008)	39.7	(2007)	35.8	-	-

1. ( ) indicates the year of birth
2. - denotes not available

5.4 United States: Annual Survey of Entrepreneurs

5.42   The United States began conducting the Annual Survey of Entrepreneurs (ASE) in 2014. The ASE includes all nonfarm businesses filing Internal Revenue Service tax forms as individual proprietorships, partnerships, or any type of corporation, with annual receipts of $1,000 or more. The ASE covers businesses with paid employees. The ASE is conducted on an enterprise basis rather than an establishment basis. An enterprise is a business consisting of one or more domestic establishments under its ownership or control. Businesses include sole-proprietorships, partnerships, and corporations. For businesses with at least one person owning 10% or more of the business, all business owners are considered entrepreneurs. These owners typically perform one or more of the following functions: manage day-to-day operations, provide services and/or produce goods, or have financial control with the authority to sign loans, leases, and contracts. Businesses without at least one person owning 10% or more of the business are not asked about the individual owners. The data are compiled by combining survey responses with available administrative records.[21]

5.4.1 Purpose of the survey

5.43   The ASE collects race, gender, ethnicity, and veteran status of business owners in the United States.

5.4.2 Scope and limitations

5.44   The ASE uses the Census Bureau's SBR to determine businesses eligible for the survey. They are classified by the North American Industry Classification System (NAICS). The ASE includes U.S.-based, nonfarm businesses except those classified as:

- Crop and Animal Production (NAICS 111, 112)
- Rail Transportation (NAICS 482)
- Postal Service (NAICS 491)
- Monetary Authorities – Central Bank (NAICS 521)
- Funds, Trusts, and Other Financial Vehicles (NAICS 525)
- Religious, Grant making, Civic, Professional, and Similar Organizations (NAICS 813)
- Private Households (NAICS 814)
- Public Administration (NAICS 92)

5.4.3 Inputs and requirements

5.45   The ASE uses the following inputs to determine the sampling frame:

- The Census Bureau's SBR, referred to by the Census Bureau simply as its *Business Register*
- Administrative data from the Social Security Administration
- Lists of minority-owned and women-owned businesses published in syndicated magazines, located on the Internet, or disseminated by trade or special interest groups
- Word strings in the business name indicating possible minority ownership
- Racial distributions for various state-industry classes and racial distributions for various ZIP codes
- Gender, ethnicity, race, and veteran status responses of a single owner business to a previous Survey of Business Owners, ASE, or the 2010 Decennial Census

5.46   After the survey data collection, the ASE uses the County Business Patterns and the SBR to obtain receipts, payroll, employment, industry classification, and geography at an establishment level.

---

[21] More information about the AES is available on https://www.census.gov/programs-surveys/ase.html

5.4.4 Methodology

5.47 Prior to data collection, each business is assigned a likelihood for each race, Hispanic ethnicity, and gender. The likelihoods are used to assign each business to one of nine frames:

- American Indian and Alaska Native (AIAN)
- Asian
- Black or African American
- Hispanic
- Native Hawaiian and Other Pacific Islander (NHOPI)
- Non-Hispanic, white men
- Other
- Public or otherwise unclassifiable by race, gender, or ethnicity
- Women[22]

5.48 Each business is assigned to a single frame regardless of the potential for response in multiple categories. If a business has equal likelihood of belonging to multiple frames, then the frame that generates the highest probability of selection is used.

5.49 The frame, metropolitan statistical area (MSA), and age of business form the strata are used to select the ASE sample. Certainty cut-offs vary by sampling stratum, and sampling occurs at different rates by stratum depending on the number of sampling units available and target variances.

5.50 The ASE began with the collection of 2014 data and continued in survey years 2015 and 2016. Therefore, the MSAs used were the top 50 MSAs based on 2014 population estimates. Additionally, the remainders of states were used to capture companies located outside of the top 50 MSAs for 49 states (the Providence, RI MSA included all of the state of Rhode Island). A category of "multi-state" was used to identify companies that operated in more than one state or MSA.

5.51 The ASE uses hot deck imputation to account for non-response. Estimates are created with the Horvitz-Thompson estimator, and variance estimates are made using the delete-a-group jack-knife variance estimation.

5.52 Businesses are tabulated according to their responses for the main demographic categories and according to each establishment's NAICS and geography that are available from the County Business Patterns (CBP) and the SBR. Note that this becomes a combination of business-level data (survey responses) and establishment-level data (CBP and SBR). For example, a company operating an establishment in New York City and another in Los Angeles would be counted once in each of those MSAs, but only once at the U.S. level. Similarly, a company operating an establishment engaged in manufacturing and another in retail trade would be counted once in each of those sectors, but only once at the U.S. level.

5.53 The company's race/gender/ethnicity/veteran statuses are determined by the percentage of the owners. The frame assigned during sampling no longer applies to tabulation. Tabulation is based on response data. For example, a business with a single owner owning 100%, reporting both white and black for race would be tabulated as both white-owned and black-owned. Each category requires more than 50% ownership to assign the business's ownership. For example, a business with three owners may report owner 1 (34%) as white, owner 2 (33%) as black, and owner 3 (33%) as American Indian. This

---

[22] For sampling, each company is assigned to a single frame. Companies with the potential to be both Hispanic- and women-owned would be assigned to the Hispanic frame to provide a higher probability of selection. The response dictates where the company is tabulated regardless of the sampling frame.

company would not be tabulated in any of the race categories, but as minority-owned since 66% of the business would be classified as minority-owned.

5.4.5   Description of results obtained

5.54   Estimates for employer businesses include the number of firms, sales and receipts, annual payroll, and employment by gender, race, ethnicity, and veteran status. Estimates are available for the United States, each state, and the top 50 MSAs. Estimates are also available by 2-digit NAICS (also referred to as sector), years in business, receipts size of firm, and employment size of firm. Data are also collected and presented for a variety of characteristics of businesses and characteristics of business owners. The characteristics of businesses include sources of financing, for example. The characteristics of business owners include age of the business owner, for example.

5.55   The following table provides some high-level data from the 2014 ASE:

Table 5.6 **Firms and their receipts by ownership in 2014**

	Number of firms	**Receipts ($million)**
All firms	5,437,782	33,036,935
Female-owned	1,057,254	1,321,154
Hispanic-owned	298,563	335,161
Minority-owned	949,318	1,089,710
Veteran-owned	405,235	924,068

5.56   The following table shows how firms are distributed by years in business:

Table 5.7 **Firms and their receipts by years in business in 2014**

Years in business	**Number of firms**	**Receipts ($million)**
less than 2 years	481,981	438, 363
2 to 3 years	723,679	653,333
4 to 5 years	519,712	567,216
6 to 10 years	1,146,177	1,645,418
11 to 15 years	2,398,315	6,609,302
16 or more years	167,917	23,123,303

Chapter 6: Examples of business demography and related entrepreneurship indicators

6.1   Introduction

6.1   This chapter provides illustrative examples of key indicators of entrepreneurial performance produced by NSOs (see Chapter 1, paragraphs 1.8-1.12). The aim of the chapter is to complement the methodological descriptions in earlier chapters with an overview of actual business demography statistics compiled and disseminated at national or international level. Building on the concepts described in Chapter 2 and the methods in Chapter 3 and Chapter 4, this chapter presents a comprehensive set of indicators, accounting for different aspects of entrepreneurship, covering the creation of enterprises, the characteristics of new enterprises including the gender of the business owner, and the export propensities of newly created businesses.

6.2   Core business demography statistics

6.2   As previously noted, business demography statistics support the analysis of a broad range of questions that are of high interest for research and policy. They assist: the monitoring of entrepreneurial activities; the assessment of the contribution of entrepreneurship to achieving social and economic objectives, including the reduction of income inequalities (which is a central concern for policy makers) and the evaluation of the effectiveness of entrepreneurship policies.

6.3   From the economic literature on entrepreneurship, and country reports on the state of entrepreneurship, a set of business demography indicators emerge as the most relevant for the analysis of business dynamics.[23] They include the following core indicators:

- Business birth rate
- Business death rate
- Churn rate of businesses
- Business survival rates
- Share of employment creation by business births
- Share of employment destruction by business deaths
- Share of young businesses
- Employment share of young businesses

6.4   Business demography statistics are usually produced on an annual basis, at the national level, with the enterprise as statistical unit, for the entire population of enterprises and/or for the population of employer enterprises only. However, business demography statistics can also be sub-annual, for different populations, and for establishments or other small statistical units.

6.5   In the past decade, business demography statistics have become part of the regular statistical production of NSOs in developed economies. Typically, breakdowns by activity sector and other stratification characteristics (e.g. size class, legal form) are available, and in many countries also by geographical region.

---

[23] See for instance: https://www.bls.gov/bdm/entrepreneurship/entrepreneurship.htm and (Fischer & Reuber, 2010)

6.6   Examples of core business demography statistics production described above are presented in Boxes 6.1 to 6.5. The first illustration refers to the internationally harmonised data collection conducted in coordination by Eurostat and the OECD (Box 6.1), and it is completed by examples of dissemination at the national level of the harmonised core statistics (Box 6.2 - Box 6.4).

Box 6.1 **Businesses demography databases of Eurostat and the OECD**

In European Union member states the annual collection of core business demography statistics follows a common framework[1] based on the recommendations of the Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007), which is also the reference instrument for the production of the above set of business demography statistics in OECD member and partner countries in the context of the OECD-Eurostat Entrepreneurship Indicators Programme (EIP).

The outcome of these initiatives is that two databases of internationally harmonised statistics on business demography have been created by Eurostat and the OECD, accessible via their respective data portals, i.e. for Eurostat Business Demography: http://ec.europa.eu/eurostat/web/structural-business-statistics/entrepreneurship/business-demography, and for the OECD Structural and Demographic Business Statistics (SDBS): http://dx.doi.org/10.1787/sdbs-data-en.

These databases are widely consulted by the research and policy communities and are a valuable reference for countries not yet compiling business demography statistics.

See also: http://ec.europa.eu/eurostat/web/structural-business-statistics/entrepreneurship/indicators

[1] Regulation (EC) No 295/2008 of the European Parliament and of the Council of 11 March 2008 concerning structural business statistics and Commission Regulation (EU) No 446/2014 of 2 May 2014 amending Regulation (EC) No 295/2008.

Box 6.2 **Birth, death and survival of enterprises in the United Kingdom**

The Office for National Statistics of the United Kingdom produces an annual business demography statistics release. Data are sourced from the Inter-Departmental Business Register (IDBR), and the business demography statistics are compiled following the guidelines of the Eurostat-OECD Manual on Business Demography (Eurostat & OECD, 2007).

The release presents counts and rates of enterprise births and deaths as well as survival rates, with breakdowns by broad industry group and UK region. See: https://www.ons.gov.uk/businessindustryandtrade/business/activitysizeandlocation/bulletins/businessdemography/previousReleases

Box 6.3 **Employment creation and destructions by enterprise births and deaths in Ireland**

In its business demography release, the Central Statistics Office of Ireland analyses the birth and death rates of enterprises, as well as survival rates, in the context of the evolution of the total population of enterprises in the business economy. The business demography statistics are compiled following the guidelines of the Eurostat-OECD Manual on Business Demography (Eurostat & OECD, 2007).

The release also presents breakdowns of births and deaths by broad industry group, and highlights employment creation and destructions by enterprise births and deaths, including at the sectoral level.

See: https://www.cso.ie/en/statistics/multisectoral/businessdemography/

Box 6.4 **Statistical releases on business demography in EU member states**

> All member states of the European Union have regular business demography statistics releases.
>
> The releases are typically in the respective national languages of member states, but English versions of the releases are often available; see for instance Statistics Austria: https://www.statistik.at/web_en/statistics/Economy/enterprises/business_demography_since_2015/index.html,
>
> Statistics Denmark: https://www.dst.dk/en/Statistik/emner/erhvervslivet-paa-tvaers/virksomhedernes-udvikling/erhvervsdemografi, and
>
> Slovenia: http://www.stat.si/StatWeb/en/News/Index/6696.

6.7   In contrast to the EU examples above, Box 6.5 and Box 6.6 illustrate cases where different concepts, definitions and/or methodology for compilation of business demography statistics are used.

Box 6.5 **Businesses counts, entries and exits: example from Australia**

> The Australian Bureau of Statistics produces, on an annual basis, a release of counts of Australian businesses, including entries and exits, where data are sourced from the Australian Bureau of Statistics Business Register. The release contains counts of actively trading businesses, rates of entry to and exit from the market sector of the Australian economy, and rates of business survival. Data are disaggregated by industry, main state of operation, type of legal organisation, institutional sector, employment size range and annual turnover size range. For more details, see Chapter 4.
>
> See: http://www.abs.gov.au/ausstats/abs@.nsf/mf/8165.0

Box 6.6 **Young firms and employment creation: example from the United States**

> The BLS Business Employment Dynamics (BED) programme of the U.S. Bureau of Labor Statistics (BLS) produces data series on new businesses and job creation. This set of statistics is generated from the Quarterly Census of Employment and Wages programme. The BED data series on age of establishment tracks cohorts of new business establishments born in the same year and reports on the associated employment.
>
> See: https://www.bls.gov/bdm/

6.3   Statistics on high-growth enterprises

6.8   In the context of the analysis of business dynamics and entrepreneurial performance, the role of high growth enterprises as drivers of job and wealth creation has attracted the interest of analysts and policy makers. Official statistics on high-growth enterprises are regularly produced by NSOs in the member states of the EU and in countries contributing to the OECD-Eurostat Entrepreneurship Indicators Programme (EIP). The relevant indicators include:

- share of high-growth enterprises
- employment in high-growth enterprises

6.9   Indicators that focus on high-growth enterprises that are younger than a defined age (*gazelles*), are also of interest for economic analysis and are compiled in several EU and OECD countries:

- share of gazelles
- employment in gazelles

6.10   The usual breakdowns and variables (as described in Section 6.2) are available.

6.11   Enterprise growth is measured either based on employment or turnover, although high-growth measures based on employment growth are more common.

6.12   Box 6.7 presents the example of dissemination by Eurostat and the OECD of harmonised statistics on high-growth enterprises produced by the member states of the European Union and member and partner countries of the OECD. The same statistics are also published in national releases, as illustrated by the example from Austria in Box 6.8.

Box 6.7 **High-growth enterprises: example from Eurostat and the OECD**

Eurostat and the OECD regularly publish harmonised statistics on high-growth enterprises compiled following the recommendations of the Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007), in the framework of the EIP. Also, in the case of the EU member states, statistics on high-growth enterprises and their employment are produced in compliance with the Commission Regulation (EU) No 446/2014 of 2 May 2014.

See: for Eurostat, the database and https://ec.europa.eu/eurostat/documents/2995521/7706167/4-26102016-AP-EN.pdf/20f0c515-ed43-45c3-ad6a-ca0b26b36de5 ; and

for the OECD database http://www.oecd-ilibrary.org/employment/entrepreneurship-at-a-glance-2017/high-growth-enterprises-rate_entrepreneur_aag-2017-17-en;jsessionid=59iifasrkhdpk.x-oecd-live-02.

Box 6.8 **High-growth enterprises in Austria**

Statistics Austria presents detailed statistics on high-growth enterprises on its website. The statistics are disaggregated in several dimensions, including activity sector, legal form and sub-national level. To provide policy context for these statistics, the website also usefully explains that statistics on high-growth enterprises are as well used as a basis for the "high growth innovative enterprises" indicator within the Agenda „Europe 2020" that is aimed at tracking changes in the progress of innovation processes. See:

https://www.statistik.at/web_en/statistics/Economy/enterprises/high_growth_enterprises_since_2015/index.html

6.4   Indicators from micro data linking

6.13   As explained in Chapter 4 of these Guidelines, linking at the micro level data from the SBR to administrative data and other datasets or survey data on businesses, as well as on individuals, provides additional insights into the relationships between the performance of an enterprise and characteristics of the enterprise (e.g., trader/not) and/or of its entrepreneur(s) (e.g., age and educational background). While these developments are relatively new, there are already many successful examples of new statistics generated from the micro linking of data.

6.4.1   Business demography statistics linked with data on individuals

6.14   Profiles of entrepreneurs are produced by linking business demography data with data on individuals (as described in Chapter 4). In contrast to statistics on self-employed, which can only describe the characteristics of the entrepreneur but not of the business he/she owns, linked enterprise-entrepreneur data permit, for instance, the analysis of business dynamics by gender of the business owner. Typical indicators compiled include the numbers of business owners; breakdowns by:

    a.   gender, age, country of origin, education

b. sector of activity

c. own-account workers and employers

6.15   Box 6.9 provides an example of official statistics about entrepreneurs.

Box 6.9 **The profile of the entrepreneur: example from Italy**

The Italian National Institute of Statistics (ISTAT) compiles statistics that support analysis of entrepreneurs in Italy, following the recommendations of the EIP for the development of entrepreneurship statistics.

The methodology developed to profile entrepreneurs exploits administrative sources and official statistics. The basic informative structure is represented by an integrated system of administrative files on occupations (of Linked Employer Employees Database (LEED) type) that enables each individual to be connected to the enterprise where he/she performs an activity. The identification of an entrepreneur within an enterprise is done through the application of deterministic rules that vary according to the juridical form of the enterprises. This informative structure is enriched with demographic information on the enterprise.

See: Cella & De Gregorio (2015)

Also, a statistical release prepared by ISTAT presents the methodology and the new indicators; see: https://www.istat.it/it/archivio/194762 (in Italian).

6.4.2   Business demography statistics linked with other businesses statistics

6.16   As explained in Chapter 4, in the past decade Eurostat has coordinated several projects of micro data linking conducted with the collaboration of EU member states. These projects involved the development of new indicators relevant for the analysis of entrepreneurship and business dynamics, notably by allowing analysis of enterprise characteristics such as business ownership (in particular, domestic versus foreign) and trade status (e.g., traders versus non-traders). New entrepreneurship indicators that can be conceived, based on the linked data, include for example:

- enterprises by age and trade status; and

- performance of small and medium-sized enterprises by ownership status, i.e. independent versus dependent.

Chapter 7: Topics for further work and research

7.1   Introduction

7.1   The Guidelines provide methodological and practical guidance to countries for establishing business demography statistics and business-based entrepreneurship statistics, with the focus on the use of an SBR as the basic database. However, many of the topics discussed are quite new – business demography statistics have only recently been developed, as have methods for linking databases to achieve new kinds of statistics without any additional burden on enterprises. Given the status of these approaches, the Guidelines do not cover all relevant issues. Furthermore, new methodologies and approaches are constantly emerging.

7.2   As developing new methodologies and good practices is resource demanding, sharing experiences between countries and conducting common development projects should be encouraged. International cooperation can help to reduce the gaps between developed and less developed statistical systems. Capacity building and training activities are useful in supporting countries with less experience. Sharing of methods and practices also helps to improve the international comparability of the resulting statistics.

7.3   During the work of drafting the Guidelines, the Task Force noted topics where further work and research is recommended. These topics are listed and briefly described below.

7.2   Recommended work and research topics by domain

SBR functionalities required to support business demography statistics

7.4   The Guidelines focus on the requirements of SBR functionalities to support business demography statistics, especially concerning the creation of a longitudinal database that provides a view of the business units in the SBR for business demography purposes. Such a database is needed as the focus of SBR maintenance is an up-to-date population of the businesses for the provision of survey frames rather than a longitudinal view. The Guidelines provide guidance on the creation of such a database and discuss ways of its integration into the SBR system. The SBR system needs to be enlarged to support the longitudinal aspects. For the creation of the longitudinal data, record linkage methods are required to identify longitudinal (continuing) units in the SBR. These record linkage methods are also required for the maintenance of the SBR, for instance, to link with administrative sources.

7.5   Work on these issues will be needed for the national implementations. However, exchange of experiences and common projects will provide additional insights into these issues and will help other countries as well. Topics include:

- Methods to create a longitudinal database of the units in the SBR for business demography purposes.

- Conceptual enlargement of the SBR system to support the integration of the longitudinal aspects.

- Practical approaches to identifying the characteristics and the demographic changes of statistical units.

- Experiences with record linkage methods.

7.6   The longitudinal database that is emerging from the enlargement of the SBR database is not only important for the compilation of business demography statistics, but also for analysis of the business development and for policy information and advice.

Framework for business demography statistics

7.7   Business demography concerns data on the dynamics of the business population over time and on the core variables that describe the development of the business population. The focus of business demography is on the birth of new businesses, the survival of newly created businesses and the death of businesses. Based on such data, various demography indicators can be derived, such as birth, survival and death rates. There are also different approaches for business demography, depending on the population that is included: including all businesses, or only employer businesses. Furthermore, different statistical units can be used. The focus may be on national or regional data, or on annual or sub-annual demographic changes. Another crucial issue is the choice of definition of entrepreneur. The definition can be based on a legal/administrative basis (e.g. self-employed persons), or on an economic basis (persons that are leading, managing and owning the businesses, irrespective of the legal status).

7.8   The actual implementation in countries may result in statistics based on one or other business demography populations. The population options have not been equally developed and tested. Topics for further work could be:

- Analysis and improved specification of the basic business demography populations and definitions.

- Analysis and testing of regional business demography.

- Analysis and testing of sub-annual (quarterly) business demography.

- Comparative analysis of business demography using different statistical units.

- Analysis and testing of enterprise group demography.

- Specification of high-growth enterprises and gazelles.

- Analysis of the different concepts and definitions of entrepreneurs and possibilities of implementation.

- Analysis and approaches to increase the timeliness of business demography statistics and timely indicators for business creations.

- Analysis and approaches to revise business demography data due to new sources and change of concepts.

- Drawing samples form the business demography population for specific data collection that cannot be achieved by linking exercises, for example "factors of business success", "access to finance".

Use of administrative sources

7.9   Administrative sources are the main data inputs for the maintenance and update of SBRs. Their structure and quality determine the possibilities for SBR maintenance, and the quality of the survey frames thus provided. Cooperation with administrative data providers is therefore crucial for the SBR.

7.10   Administrative sources are also central for the creation of the longitudinal databases that serve as the basis for business demography. Thus, for a country establishing business demography statistics, a topic for further work is analysis of the administrative data available in the country for business demography and their use as input to the SBR.

Survey approaches to compile business demography data

7.11   For various reasons, some countries are compiling business demography statistics based on survey approaches rather than on the SBR. Such approaches are usually quite expensive and may not be repeated on a regular basis. Also, due to the restrictions imposed by the survey, it may not be possible to apply fully international concepts and definitions. However, survey approaches may be a step towards regular data compilation using the SBR. Work on these issues could cover:

- Analysis of the deviations of survey-based approaches from the international business demography concepts and definitions.

- Ways of integrating survey-based data into the SBR system.

- Strategies for switching from a survey-based approach to an SBR based approach.

Linking of databases to achieve additional characteristics of the businesses and the entrepreneurs

7.12   An SBR covers only a limited set of variables for each business unit. Identifying demography events per se does not increase the variables available. Thus, linking the business units, especially those that belong to the business demography population, with data from other business or social statistics is required to enlarge the analytical possibilities and to gain new insights into the dynamics of business development.

7.13   Linking of databases at the micro level has become an important approach as enlarged databases are established without additional burden on the businesses. Work on the following issues should be encouraged:

- Linking methodology in the absence of a common identifier.

- Linking methodology in cases where one of the linked databases covers only a sample of the population.

- Compilation and analysis of databases achieved by linking business demography data with other business data, such as foreign trade, including trade in services, structural information, FATS, R&D, innovation, as well as data on international sourcing/global value chains.

- Compilation and analysis of databases achieved by linking business demography data with data on entrepreneurs (such as gender, age, education status, entrepreneurial experience) and on persons employed.

- Approaches to compile data on different business arrangements, such as family businesses, home based businesses, indigenous businesses.

International comparability

7.14   For improvement of quality and international comparability, Countries are encouraged to implement international methods and concepts as far as possible. The Guidelines support this goal by providing guidance on concepts, definitions and methods, and by providing some country examples. Work in this area would improve international comparability, in particular analysis of how national implementations diverge from international standards as concerns concepts and definitions or using thresholds in their statistical business registers and developing of strategies in order to increase compliance.

Glossary

Where no source is indicated the source is these Guidelines.

Active unit

A unit is active when it has any economic activity or when it has no economic activity but is legally or administratively registered and part of another unit that has economic activity at any time during a respective reference period.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Activity, economic activity

Activity

An activity is a process, i.e. the combination of actions that result in a certain set of products. Activities are defined as the use of inputs (e.g., capital, labour, energy and materials) to produce outputs. The outputs that result from undertaking activities can be transferred or sold to other units (in market or non-market transactions), placed in inventory or used by the producing units for own final use. In practice the majority of units carry on activities of a mixed character. One can distinguish between three types of economic activity:

Principal activity: The principal activity is the activity which contributes most to the total value added of the unit under consideration.

Secondary activity: A secondary activity is any other activity of the unit that produces goods or services.

Ancillary activity: Any ancillary activities are those that exist solely to support the main productive activities of a unit by providing non-durable goods or services for the use of that entity.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Active unit, ancillary activity, classification of activities, economic activity, principal activity, secondary activity

Activity [business demography]

Within business demography context, activity is defined as any turnover and/or employment in the period from 1st January to 31st December in a given year. This definition complements the concept of activity. There are cases where the reporting year of the businesses is deviating from the calendar year. Where the calendar year cannot be met, this should be noted for transparency.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Activity

Administrative business register

An administrative business register is a regularly updated structured list of specific business units in a territorial area, which is maintained by administrative authorities for administrative, legal or taxation purposes (e.g. recording and maintaining certain details of businesses or taxation).

Source: Guidelines on Statistical Business Registers (UNECE, 2015)

Link: http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

Related terms: Administrative register, statistical business register

Administrative data

Data originally collected for non-statistical purpose. Control of the methods by which the administrative data are collected and processed rests with the administrative agency. In most cases the administrative authority will be a government unit.

Source: Using Administrative and Secondary Sources for Official Statistics - A Handbook of Principles and Practices (UNECE, 2011)

Link: http://www.unece.org/fileadmin/DAM/stats/publications/Using_Administrative_Sources_Final_for_web.pdf

Related terms: Administrative register, administrative source

Administrative register

A register is a written and complete record containing regular entries of items and details on a set of objects. Typically, a register is a structured list of units, containing a number of attributes for each of those units, and having some sort of regular updating mechanism. Registers maintained by administrative authorities for administrative purposes can be considered to be administrative registers. 'Administrative register' is an umbrella term and covers for example records collected for business registers or personal registers for administrative purposes.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Administrative business register

Administrative source

Administrative sources are files of data collected by government bodies for the purposes of administering taxes and benefits or monitoring populations. More generally, administrative sources contain information that is not primarily collected for statistical purposes.

Source: Using Administrative and Secondary Sources for Official Statistics - A Handbook of Principles and Practices (UNECE, 2011)

Link: http://www.unece.org/fileadmin/DAM/stats/publications/Using_Administrative_Sources_Final_for_web.pdf

Related terms: Administrative data, administrative register

Administrative unit

An administrative unit is designed for the purposes of conforming with an administrative regulation, for example for registration purposes or for accounting purposes of VAT and other taxes.

Source: Guidelines on Statistical Business Registers (UNECE, 2015)

Link: http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

Related terms: Statistical unit

Ancillary activity

Ancillary activities are undertaken to support principal and secondary productive activities of a unit by providing goods or services entirely or primarily for the use of that entity, such as bookkeeping,

transportation, storage, purchasing, sales promotion, cleaning, repair and maintenance, security etc. The output is always intended for intermediate consumption within the same unit and is therefore usually not recorded separately. Although most ancillary activities produce services, some goods-producing activities may, by exception, be regarded as ancillary. The goods thus produced, however, may not become a physical part of the output of the main productive activities. Ancillary activities are usually fairly small-scale compared with the principal activity they support.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Activity, economic activity, principal activity, secondary activity

Attribute

See: Characteristic

Autonomy of decision

See: Enterprise

Birth (of enterprise)

A birth is characterized by the creation of a combination of production factors with the restriction that no other enterprises are involved in the event. Births do not include creation of entries into the population due to mergers, break-ups, split-offs or restructuring of a set of enterprises. It does not include entries into a sub-population resulting only from a change of activity. A birth means the enterprise starts from scratch and actually starts activity. An enterprise creation can be considered as an enterprise birth if new production factors, in particular new jobs, are created. If a dormant unit is reactivated within two years, this event is not considered a birth.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Break-up (of enterprise), creation (of business), date of creation (of enterprise), new enterprise, split-off (of enterprise)

Birth (of enterprise group)

The birth of an enterprise group is the establishing of a link of control, direct or indirect, between two or more independent legal units, where no link of control existed before and no other enterprise group is involved.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Break-up (of enterprise group), creation (of business), date of creation (of enterprise groups)

Birth date (of business)

The date on which a unit was born. For enterprises, the date could be decided by referring to the definition of the enterprise: the birth takes place at the (first) moment the conditions of the definition are met, so the moment there is an organisational unit producing goods or services. The convention is that the date on which the first financial commitments for investments are made should be taken. This may seem somewhat early, since actual production will take place afterwards, but this allows for the statistical coverage of all important variables, such as investments, from the beginning. This date is not

considered too early since serious commitments have been made, however, from the point of view of cost-efficiency and response burden it may not always be desirable to actually collect the date of birth from the enterprise; in that case the registration date at the administrative source has to be taken.

Source: Business registers – Recommendations manual (Eurostat, 2010), paragraphs 13.13, 13.30

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Date of creation

Birth rate

The birth rate of a given reference period (usually one calendar year) is the number of births as a percentage of the population of active enterprises. This birth rate may vary depending on the birth concept that is used. The use of thresholds affects both the enumerator (births) and the denominator (population of active enterprises), but may do so to a different degree.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Break-up (of enterprise)

A break-up results in one enterprise before and more than one enterprise after the event. In a break up, the enterprise is divided in such a way that neither (none) of the new enterprises keeps the identity of the original enterprise. There is no continuity or survival, but the closure of the previous enterprise is not considered to be a real death. Similarly the new enterprises are not considered to be real births. A break up is similar to a split-off and can be seen the opposite of a merger.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Birth (of enterprise), creation (of business), new enterprise, split-off (of enterprise)

Break-up (of enterprise group)

A break-up results in one enterprise group before and more than one enterprise group after the event. In a break-up, the enterprise group is divided in such a way that neither (none) of the new enterprise groups keep the identity of the original enterprise group.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Birth (of enterprise group), creation (of business)

Business

The term is used as a type of enterprise, namely a "commercial enterprise" or legal unit with commercial economic activity.

Related terms: Enterprise

Business closure

See: Cessation of business

Business demography

Business demography covers events, like births and other creations of units, deaths and other cessations of units, and their ratio to the business population. It covers follow-up of units in the time dimension, thus gaining information on their survival or discontinuity. It also covers development in the time

dimension according to certain characteristics like size, thus gaining information on the growth of units, or a cohort of units, by type of activity. Demographic information can in principle be produced for any statistical unit; however, a clear political interest in the European Union is on enterprise demography. In other regions business demography data are often calculated based on establishments. The demography of enterprises can be assessed by studying enterprise births and enterprise deaths and by examining the change in the number of enterprises by type of activity, i.e. by examining the flows and stocks to get a complete picture of the enterprise dynamism.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Continuity, survival

Business register

See: Statistical business register

Business start-up

See: Creation of business

Cessation (of business)

The cessation of activities of a unit can occur either due to a (real) death of the unit, or due to other cessation by a merger, take-over, break-up or discontinuity point according to the continuity rules.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Business closure, death (of enterprise), death (of enterprise group)

Characteristic

A characteristic is one of a set of information that is stored in a business register to describe a statistical unit. Characteristics are provided for identification of a unit like name, address, and identification numbers, for economic description of a unit, like activity code, turnover or employment of a unit or for the structure of a unit, like the relationship to other statistical units.

Related terms: Variable

Churn rate (of employer enterprises)

The churn rate is defined as the sum of the employer enterprise birth rate and the employer enterprise death rate.

Source: Entrepreneurship at a Glance 2017 (OECD, Entrepreneurship at a Glance 2017, 2017)

Link: http://www.oecd.org/sdd/business-stats/entrepreneurship-at-a-glance-22266941.htm

Related terms: Employer enterprise birth rate, employer enterprise death rate, enterprise churn rate

Classification of activities

The main purpose of a classification of activities is to classify economic entities and statistical units, such as establishments or enterprises, according to the economic activity in which they mainly engage. The main aim is to provide a set of activity categories that can be utilised when dissecting statistics according to such activities. Different classifications are needed to cater for the different functions which statistics are required to perform, so at international and national levels classifications have been developed for different purposes. Economic classifications can be broadly divided into two categories:

Classifications of economic activities form the basis for compiling statistics on output, production factors (labour, raw materials and supplies, energy etc.), capital formation and financial transactions. The international classification for activity is the International Standard Industrial Classification of All Economic Activities (ISIC), maintained by the United Nations, used at world level.

Classifications of products: The outputs of the economic entities are termed products and are generally divided into goods and services and grouped according to a product classification. The international classification for products is the Central Product Classification (CPC), maintained by the United Nations and used worldwide.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008); and Central Product Classification (CPC) Version 2.1 (United Nations, 2015)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

https://unstats.un.org/unsd/classifications/unsdclassifications/cpcv21.pdf

Related terms: Activity

Company

See: Corporation

Concentration of enterprises

Concentration of enterprises refers to demographic events involving more than one enterprise before and one enterprise after events like merging or taking-over. The term may also be used to denote that the population of enterprises gets fewer owners or is spread over a reduced number of enterprise groups.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Merger, take-over

Continuation

See: Survival

Continuity (of enterprise)

In theory, the continuity rules would be derived from the definition of the enterprise (or other units) and its statistical uses. In principle, the continuity of an enterprise depends on the continuity of its production factors: employment, machines and equipment, land, buildings, management, and intangible assets. The continuity of these factors can be measured and weighted to decide upon the continuity of the enterprise. In practice, the continuity rules consider three main criteria: continuity of control, economic activity and location.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Business demography, survival

Corporation

In the legal sense, corporations may be described by different names: corporations, incorporated enterprises, public limited companies, public corporations, private companies, joint-stock companies, limited liability companies, limited liability partnerships, and so on. In the SNA, the term corporation covers legally constituted corporations and also cooperatives, limited liability partnerships, notional resident units and quasi-corporations. The term corporation is used more broadly than in just the legal

sense. In general, all entities that are: capable of generating a profit or other financial gain for their owners, recognized at law as separate legal entities from their owners who enjoy limited liability, set up for purposes of engaging in market production.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/SNA2008.pdf

Related terms: Enterprise, enterprise group, establishment, legal unit

Creation (of business)

The emergence of a new business unit. This can be either due to a (real) birth of the unit, or due to other creation by a merger, break-up, split-off or discontinuity point according to the continuity rules.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Birth (of enterprise), birth (of enterprise group), break-up (of business), business start-up, new enterprise, take-over

Date of birth (of enterprise)

See: Date of creation (of enterprise)

Date of birth (of enterprise group)

See: Date of creation (of enterprise group)

Date of birth (of establishment)

See: Date of creation (of establishment)

Date of birth (of legal unit)

See: Date of creation (of legal unit)

Date of birth (of local unit)

See: Date of creation (of local unit)

Date of cessation (of enterprise)

Date of final cessation of activities. It refers to the death or other deletion date of the enterprise (when it becomes historical) and is interpreted similarly to the corresponding variable for local units. A death amounts to the dissolution of a combination of production factors with the restriction that no other enterprises are involved in the event. Deaths do not include exits from the population due to mergers, take-overs, break-ups or restructuring of a set of enterprises. They do not include exits from a sub-population resulting only from a change of activity. An enterprise is included in the count of deaths only if it is not reactivated within two years. Equally, a reactivation within two years is not counted as a birth.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Cessation (of business), death (of enterprise)

Date of cessation (of enterprise group)

Date of cessation of the all-resident/truncated enterprise group. Cessation of a group means either death of the group (dissolution of the links of control between the units belonging to the group), or (more commonly) other cessation date by merger with or take-over by another group, or break-up, split-off, or restructure into two or more groups. The death of an enterprise group is the cessation of all control links, direct or indirect, between the legal units of which the enterprise group consists. The legal units become independent again or cease to exist. No other enterprise group is involved.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Cessation (of business), death (of enterprise group)

Date of cessation (of establishment)

Date of cessation of activities. It refers to the death or the deletion date of the establishment. Since the establishment is a part of an enterprise, being active in a certain activity at a certain location, the death of an establishment amounts to the dissolution of a (partial) combination of production factors in a certain activity and at a geographically identified place.

Date of cessation (of legal unit)

Date of cessation is not easy to collect but the registration of the event is far more important than the precise day and month when it took place. Basically, the legal unit ceases to be part of an enterprise when: the legal unit ceases to exist or the legal unit ceases to be economically active and it is not part of the control chain within the enterprise group. Between activity and real death, there is therefore often a period of inactivity during which the unit may be regarded as 'dormant'. A sign of such a situation would be the lack of employees, the cessation of tax compliance or the inability to contact the unit after repeated efforts.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Cessation (of business)

Date of cessation (of local unit)

Date of cessation of activities. It refers to the death or other deletion date of the local unit. As for legal units, this date may not be available with any precision, only the fact that the local unit has ceased to exist during the reference year may be known. Since the local unit is a part of an enterprise, situated in a geographically identified place, and the enterprise is a combination of production factors, the death of a local unit amounts to the dissolution of a (partial) combination of production factors at a geographically identified place.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Cessation (of business)

Date of creation (of enterprise)

Date of commencement of activities. The date refers to the date of birth, i.e. in principle the date on which the first financial commitments are made, although in practice it may refer to the registration date in the administrative source, if the unit starts its economic activities immediately after that. However, the legal unit may change and be reregistered for instance after a change of legal form, while the enterprise remains the same, because the continuity rules for enterprises should be applied.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Birth (of enterprise), creation (of business)

Date of creation (of enterprise group)

Date of commencement of the all-resident/truncated enterprise group. The date refers either to a date when a new all-resident group is born, or other creation date of a new group (by merger, break-up, split-off, or restructure). The birth of a new group may be difficult to define in practice, if the smallest groups of no statistical importance are not monitored. The date from which the group is being monitored shall then be used as a proxy. However, the approximate dates are important in order to know from which year a certain multinational group is monitored in different countries.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Birth (of enterprise group), creation (of business)

Date of creation (of establishment)

Date of commencement of the activities. This should refer to the birth or other creation date of the establishment according to the continuity rules.

Date of creation (of legal unit)

Date of incorporation for legal persons or date of official recognition as an economic operator for natural persons. The date of official recognition should be the date on which an identification number is given, or the date on which the legal existence was approved, be it a company/trade register number, a VAT number or other.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Creation (of business)

Date of creation (of local unit)

Date of commencement of the activities. This date should refer to the birth or other creation date of the local unit according to the continuity rules.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Creation (of business)

Date of death (of enterprise)

See: Date of cessation (of enterprise)

Date of death (of enterprise group)

See: Date of cessation (of enterprise group)

Date of death (of establishment)

See: Date of cessation (of establishment)

Date of death (of legal unit)

See: Date of cessation (of legal unit)

Date of death (of local unit)

See: Date of cessation (of local unit)

Death (of business)

See: Date of cessation (of business)

Death (of enterprise)

The death of an enterprise refers to the dissolution of a combination of production factors with the restriction that no other enterprises are involved in the event. Deaths do not include exits from the population due to mergers, take-overs, break-ups or restructuring of a set of enterprises. It does not include exits from a sub-population resulting only from a change of activity.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Business closure, cessation (of business), date of cessation (of enterprise)

Death (of enterprise group)

The death of an enterprise group is the cessation of all control links, direct or indirect, between the legal units of which the enterprise group consists. The legal units become independent again or cease to exist. No other enterprise group is involved.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Business closure, cessation (of business), date of cessation (of enterprise group)

Death (of establishment)

See: Date of cessation (of establishment)

Death (of legal unit)

See: Date of cessation (of legal unit)

Death (of local unit)

See: Date of cessation (of local unit)

Death rate

The death rate of a given reference period (usually one calendar year) is the number of deaths as a percentage of the population of active enterprises. This death rate may vary depending on the birth concept that is used. The use of thresholds affects both the enumerator (deaths) and the denominator (population of active enterprises), but may do so to a different degree.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Decline

See: Growth

De-concentration

De-concentration refers to demographic events involving one enterprise before and more than one enterprise after the events of break-ups and split-offs. The term may also be used to denote that the population of enterprises gets more owners or is spread over a larger number of enterprise groups.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Break-up, split-off

Demography of enterprises

See: Business demography

Dormant unit

A unit is said to be dormant if it is legally alive and has legal personality, but does not carry on any activity and has neither employment nor turnover.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Cessation of business, reactivation

Economic activity

Any activity comprising the offer of goods and services on a given market is an economic activity. Additionally, non-market services contributing to the GDP as well as direct and indirect holdings of active legal units are economic activities for the purpose of SBRs.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Active unit, activity

Economic business demography

See: Economic enterprise demography

Economic census

A survey conducted on the full set of observation objects belonging to a given business population.

Source: Terminology on Statistical Metadata (UNECE, 2000)

Link: https://www.unece.org/fileadmin/DAM/stats/publications/53metadaterminology.pdf

Related terms: Economic survey

Economic enterprise

A subset of the population of employer enterprises comprising enterprises that have at least two employees.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Employer enterprise, enterprise

Economic enterprise birth

Birth of an enterprise with at least two employees. This population consists of enterprise births that have at least two employees in the birth year and of enterprises that existed before the year in consideration, but were below the threshold of two employees.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Economic enterprise death, employer enterprise birth, entry by growth

Economic enterprise death

An Economic Enterprise death occurs either as an enterprise death with at least two employees in the year of death or as an exit by decline, moving below the threshold of two employees.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Economic enterprise birth, employer enterprise death, exit by decline

Economic enterprise demography

Business demography covering only the population of enterprises with two or more employees.

Economic operator

See: Economic unit

Economic organisation

See: Economic unit

Economic producer

See: Economic unit

Economic production

Economic production may be defined as an activity carried out under the control and responsibility of an institutional unit that uses inputs of labour, capital, and goods and services to produce outputs of goods or services.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Activity, economic unit

Economic statistics

Economic statistics describe the activities of economic transactors and the transactions that take place between them. In the real world, economic entities engaged in the production of goods and services vary in their legal, accounting, organizational and operating structures.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Business statistics

Economic survey

An investigation about the characteristics of a given business population by means of collecting data from a sample of that population and estimating their characteristics through the systematic use of statistical methodology

Source: Terminology on Statistical Metadata (UNECE, 2000)

Link: https://www.unece.org/fileadmin/DAM/stats/publications/53metadaterminology.pdf

Related terms: Economic census

Economic unit

An economic unit is a legal unit, or part of a legal unit, with economic production as defined in the current version of the SNA.

Source: Guidelines on Statistical Business Registers (UNECE, 2015)

Link: http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

Related terms: Economic production, legal unit, statistical unit

Employees

Employees are persons who, by agreement, work for a resident enterprise and receive a compensation for their labour. The relationship of employer to employee exists when there is an agreement, which may be formal or informal, between the employer and a person, normally entered voluntarily by both parties, whereby the person works for the employer in return for remuneration in cash or in kind. The measurement is realized by the actual number of persons employed, and number of employees, both as head counts and, in the latter case, also in full-time equivalents (FTEs) defined as total hours worked divided by average annual hours worked in full-time jobs. The main uses of these characteristics are in stratification for sampling, analysis and dissemination purposes.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms Employment, number of employees, number of persons employed

Employer business demography

See: Employer enterprise demography

Employer enterprise

A subset of the population of enterprises comprising enterprises that have at least one employee.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Economic enterprise, enterprise

Employer enterprise birth

Birth of an enterprise with at least one employee. This population consists of enterprise births that have at least one employee in the birth year and of enterprises that existed before the year in consideration but were below the threshold of one employee.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Economic enterprise birth, employer enterprise death, entry by growth

Employer enterprise birth rate

The employer enterprise birth rate is the percentage of newly born employer enterprises in year t in relation to all active employer enterprises in year t.

Related term: Employer enterprise birth

Employer enterprise death

An employer enterprise death occurs either as an enterprise death with at least one employee in the year of death or as an exit by decline, moving below the threshold of one employee.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Economic enterprise death, exit by decline

Employer enterprise death rate

The employer enterprise death rate is the percentage of all employer enterprises closed in year t in relation to all active employer enterprises in year t.

Related term: Employer enterprise death

Employer enterprise demography

Business demography covering only the population of employer enterprises, i.e. enterprise with at least one employee.

Employer enterprise survival rate

The survival rate of newly born employer enterprises in a given reference period is the number of employer enterprises that were born in year t-n and survived to year t as a percentage of all employer enterprises born in year t-n.

Related term: Employer enterprise survival

Employment

Employment includes all persons, both employees and self-employed persons, engaged in some productive activity that is undertaken by a resident enterprise.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Employees, self-employed

Enterprise

An enterprise is a legal unit (or the smallest set of legal units) producing economic goods and services with autonomy in respect of financial and investment decision-making, as well as authority and responsibility for allocating resources for the production of goods and services. It may be engaged in one or more productive activities. An enterprise may be a corporation (or quasi-corporation), a non-profit institution or an unincorporated enterprise. Corporate enterprises and non-profit institutions are complete institutional units. On the other hand, the term "unincorporated enterprise" refers to a household or government unit in its capacity as a producer of goods and services. The enterprise is the level of statistical unit at which information relating to its transactions, including financial and balance-sheet accounts, are maintained, and from which international transactions, an international investment position (when applicable), consolidated financial position and net worth can be derived.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Multinational enterprise, standard statistical unit

Enterprise birth

See: Birth (of enterprise)

Enterprise birth rate

The enterprise birth rate is the percentage of newly born enterprises in year t in relation to all active enterprises in year t.

Related term: Enterprise birth

Enterprise churn rate

The churn rate is defined as the sum of the enterprise birth rate and the enterprise death rate.

Source: Entrepreneurship at a Glance 2017 (OECD, Entrepreneurship at a Glance 2017, 2017)

Link: http://www.oecd.org/sdd/business-stats/entrepreneurship-at-a-glance-22266941.htm

Related terms: Enterprise birth rate, enterprise death rate

Enterprise closure

Enterprises that are not active in a given period but were active in the previous period. The number of enterprise deaths is derived from the population of enterprise closures by removing reactivations within two years and closures that do not meet the definition of enterprise deaths.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Activity, new enterprise, number of deaths of enterprises

Enterprise death

See: Death (of enterprise)

Enterprise death rate

The enterprise death rate is the percentage of all enterprises closed in year t in relation to all active enterprises in year t.

Related term: Enterprise death

Enterprise survival rate

The survival rate of newly born enterprises in a given reference period is the number of enterprises that were born in year t-n and survived to year t as a percentage of all enterprises born in year t-n.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related term: Enterprise survival

Enterprise group

An enterprise group is an association of enterprises bound together by legal and/or financial links. A group of enterprises can have more than one decision-making centre, especially for policy on production, sales and profit. It may centralise certain aspects of financial management and taxation. It constitutes an economic unit which is empowered to make choices, particularly concerning the units which it comprises. An enterprise group is a set of enterprises controlled by the group head.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: All-resident enterprise group, multinational enterprise group, truncated enterprise group

Entrepreneur

An entrepreneur is a person who owns an enterprise and seeks to generate value through the creation or expansion of economic activity, by identifying and exploiting new products, processes or markets.

Source: Defining Entrepreneurial Activity: Definitions Supporting Frameworks for Data Collection (OECD, 2008)

Link:http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=std/doc(2008)1

Entrepreneurial activity

Entrepreneurial activity is enterprising human action in pursuit of the generation of value through the creation or expansion of economic activity, by identifying and exploiting new products, processes or markets.

Source: Defining Entrepreneurial Activity: Definitions Supporting Frameworks for Data Collection (OECD, 2008)

Link:http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=std/doc(2008)1

Entrepreneurship

Entrepreneurship is the phenomenon associated with entrepreneurial activity.

Source: Defining Entrepreneurial Activity: Definitions Supporting Frameworks for Data Collection (OECD, 2008)

Link:http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=std/doc(2008)1

Entrepreneurship statistics

Entrepreneurship statistics is statistics on the entrepreneurship phenomenon based on the indicators framework of OECD and Eurostat. It comprises data on determinants of entrepreneurial activity, entrepreneurial performance and impact of entrepreneurial activity.

Source: Definition based on Defining Entrepreneurial Activity: Definitions Supporting Frameworks for Data Collection (OECD, 2008)

Link:http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=std/doc(2008)1

Entry by growth

An entry by growth occurs if an enterprise was already active, but its employment was below the employee threshold for at least two years before the year when it reaches the employee threshold. This is an event that occurs only in the context of business demographic data based on a threshold of one or two employees.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Economic enterprise birth, employer enterprise birth, exit by decline

Establishment

The establishment is defined as an enterprise or part of an enterprise that is situated in a single location and in which only a single (non-ancillary) productive activity is carried out or in which the principal productive activity accounts for most of the value added.

Establishments are also referred to as "local kind of activity units" (local KAU).

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Standard statistical unit

European System of Accounts (ESA 2010)

The European System of National and Regional Accounts (ESA) is an internationally compatible accounting framework for a systematic and detailed description of a total economy (that is a region, country or group of countries), its components and its relations with other total economies. The ESA is fully consistent with the world-wide guidelines on national accounting, namely the System of National Accounts (SNA).

Source: European System of National Accounts 2010 (Eurostat, 2013)

Link: http://ec.europa.eu/eurostat/documents/3859598/5925693/KS-02-13-269-EN.PDF/

Related terms: System of National Accounts

Event [business demography]

The table below contains the main demographic events for enterprises, the number of enterprises involved in the events and their consequences for business registers in terms of number of register creations and deletions. For enterprise groups a similar table could be constructed as most of the events can be defined for groups as well. For local units only births and deaths are generally observed.

Event	Real, observable world		Business register	
	Number of enterprises before the event	*Number of enterprises after the event*	*Number of creations*	*Number of deletions*
Birth	-	1	1	-
Death	1	-	-	1
Change of ownership	1	1	-	-
Merger	n	1	1	n
Take-over	n	1	-	n-1
Break-up	1	n	n	1
Split-off	1	n	n-1	-
Creation of a joint venture	n	n+1	1	-
Cessation of a joint venture	n	n-1	-	1
Restructuring within an enterprise	1	1	-	-
Restructuring within an enterprise group	n	N	0 or more	0 or more
Change of group	1	1	-	-
Complex restructuring	n	n	0 or more	0 or more

Note: n = 2 or more

Source: Business registers – Recommendations manual (Eurostat, 2010), paragraph 13.45

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Exit by decline

An exit by decline occurs if an enterprise continues to be active but moves below the employee threshold for at least two years. This is an event that occurs only in the context of demographic business data based on a threshold of one or two employees.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Economic enterprise death, employer enterprise death, entry by growth

Factors of production

A factor of production is any good or service used to produce an output. In economics, factors of production are normally grouped into the categories land, labour and capital. Capital includes intermediate inputs.

Source: Business registers – Recommendations manual (Eurostat, 2010), paragraph 14.9

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

False match

A pair wrongly designated as a match in the matching process

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Matching

False non-match

A pair which is a match in reality, but is designated as a non-match in the matching process.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Matching

Family business

There is no internationally agreed definition of family businesses so far. The European Commission has adopted a common definition of family businesses for the European Union. According to this definition, a firm, of any size, is a family business, if:

1. The majority of decision-making rights are in the possession of the natural person(s) who established the firm, or in the possession of the natural person(s) who has/have acquired the share capital of the firm, or in the possession of their spouses, parents, child, or children's direct heirs.

2. The majority of decision-making rights are indirect or direct.

3. At least one representative of the family or kin is formally involved in the governance of the firm.

4. Listed companies meet the definition of family enterprise if the person who established or acquired the firm (share capital) or their families or descendants possess 25 per cent of the decision-making rights mandated by their share capital.

Source: European Commission, Directorate Growth

Link: http://ec.europa.eu/growth/smes/promoting-entrepreneurship/we-work-for/family-business/

Financial corporation

See: Institutional sector

Firm

The term firm usually refers to a market business unit. It is often used as synonym for company or unincorporated business.

Related terms: Business, Company

Foreign affiliate

An enterprise resident in the compiling country over which an institutional unit not resident in the compiling country has control, or an enterprise not resident in the compiling country over which an institutional unit resident in the compiling country has control.

Source: Recommendations Manual on the Production of Foreign AffiliaTes Statistics (FATS) (Eurostat, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/FATS_2007_EN.pdf

Related terms: Foreign Affiliates Trade in Services (FATS)

Foreign Affiliates Trade in Services (FATS)

FATS mean statistics describing the overall activity of foreign affiliates. As there are two points of view with respect to the location of the affiliate, FATS is divided in two statistics, inward and outward FATS: Inward statistics on foreign affiliates describe the activity of foreign affiliates resident in the compiling economy. Outward statistics on foreign affiliates describe the activity of foreign affiliates abroad controlled by the compiling economy.

Source: Foreign AffiliaTes Statistics (FATS) Recommendations Manual (Eurostat, 2012)

Link: http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-12-016

Related terms: Foreign affiliate

Frame

The frame for a given survey is the subset of the frozen frame, comprising the set of units that match the specification of the survey target population. Thus, for example, a survey of employment will include units in all (or at least most) industries that are employers, i.e., will exclude units that are non-employers. A survey of manufacturing will include all units that have an ISIC code in the manufacturing group, whether they have employees or not. A survey of capital expenditure may include all units above a certain size. Thus the survey frames are typically different from one another but may be extracted from the same common set of units, i.e. a frozen frame. A frame may be referred to as a survey frame also as a sampling frame or a survey sampling frame.

Source: Guidelines on Statistical Business Registers (UNECE, 2015)

Link: http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

Related terms: Frozen frame

Frozen frame

The frozen frame is a subset of the snapshot that comprises all statistical units that are active, or potentially active, or active within the previous reference year. It also includes administrative units that are linked to these statistical units. The aim is to include all units and all characteristics that are used by subsequent processes. In other words it is a trimmed down version of the snapshot that is easier to manipulate because the possible large number of inactive units are not there. It may be further restricted by containing only units for which there are values for the characteristics that are to be used for frame extraction and sample selection for at least one survey.

Source: Guidelines on Statistical Business Registers (UNECE, 2015)

Link: http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

Related terms: Frame, live register

Frozen register

See: Register snapshot

Full-time equivalent units

Full-time equivalent units are used in annual business statistics to improve the comparability of measures of employment. Figures for the number of persons working less than the standard working time of a full-year full-time worker, should be converted into full time equivalents, with regard to the working time of a full-time full-year employee in the unit. Included in this category are people working less than a standard working day, less than the standard number of working days in the week, or less than the standard number of weeks/months in the year. The conversion should be carried out on the basis of the number of hours, days, weeks or months worked.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Head count, number of employees

Gazelle

A gazelle is a high-growth enterprise that is up to 5 years old.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: High-growth enterprise

General government

See: Institutional sector

Growth

The term growth is used in business demography to study how groups of enterprises develop. Growth is measured in terms of a change in size (in this case employment) over time. It is expected that growth for births will generally be positive (for those enterprises that have survived) as the vast majority are very small at the time of start-up. There will be occasional cases for births, and more frequent cases for the population of active enterprises, where the change in size is negative. The term used for this case is decline.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007), Section 6.2.

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Decline

Head count

The number of physical persons (full time and part time) employed by a unit

Source: Business registers – Recommendations manual (Eurostat, 2010), paragraph 5.5

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Full-time equivalent units, number of persons employed

High-growth enterprise

A high-growth enterprise is an enterprise with average annualised growth greater than 10% per annum, over a three year period. Growth can be measured by the number of employees or by turnover.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Gazelle

Historical data

When a unit has ceased and is not dormant (temporary inactive), for purpose of reconstructing demographic events the record needs to be marked as historical and not to be deleted physically.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Cessation of business, historical register

Household

See: Institutional sector

Identifier

The purpose of an identifier is to identify a unit and to link it with other units in the register and with administrative and statistical sources. The identity number of a legal unit can be either specific to the statistical business register or an external one, common or shared with other institutions in the country, a so called unique identifier.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Characteristic

Industry

An industry consists of a group of establishments engaged in the same, or similar, kinds of activity. At the most detailed level of classification, an industry consists of all the establishments falling within a single class of the ISIC. At higher levels of aggregation corresponding to the groups, divisions and, ultimately, sections of the ISIC, industries consist of a number of establishments engaged in similar types of activities.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Establishment, kind-of-activity

Institutional sector

In the 2008 SNA all resident institutional units are grouped together to form five institutional sectors, on the basis of their principal functions, behaviour and objectives:

S.11. Non-financial corporations are institutional units which are independent legal entities and market producers that are principally engaged in the production of goods and non-financial services.

S.12. Financial corporations are institutional units which are independent legal entities and market producers that are principally engaged in financial services including financial intermediation.

S.13. General Government consists of institutional units that, in addition to fulfilling their political responsibilities and their role of economic regulation, produce services (and possibly goods) for individual or collective consumption mainly on a non-market basis and redistribute income and wealth.

S.14. Households are institutional units consisting of individuals or groups of individuals as consumers and as entrepreneurs producing market goods and non-financial and financial services provided that the production of goods and services is not by separate entities treated as quasi-corporations. It also includes individuals or groups of individuals as producers of goods and non-financial services for exclusively own final use.

S.15. Non-profit institutions serving households (NPISHs) are separate legal entities which are non-market producers that are principally engaged in the production of services for households or the community at large and whose main resources are voluntary contributions.

ESA 2010 further divides the institutional sectors S.11 – S.14 into sub-sectors according to type of production (market or non-market) and control (government or non-government) of the institutional units.

No single method of sub-sectoring may be optimal for all purposes or all countries, so that alternative sub-sectoring may be useful. Dividing the total economy into sectors enhances the usefulness of the accounts for purposes of economic analysis by grouping together institutional units with similar objectives and types of behaviour. Sectors and subsectors are also needed to be able to target or monitor particular groups of institutional units for policy purposes. Only the following sectors and subsectors of ESA 2010 are relevant for business demography statistics:

- S.11 Non-financial corporations
- S.12 Financial corporations
- S 14.1 Household sector: Employers
- S 14.2 Household sector: Own account workers

Source: System of National Accounts, 2008 (United Nations, 2009); European System of Accounts 2010 (Eurostat, 2013)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf and
http://ec.europa.eu/eurostat/documents/3859598/5925693/KS-02-13-269-EN.PDF/

Related terms: Institutional unit

Institutional unit

An institutional unit is an economic unit that is capable, in its own right, of owning assets, incurring liabilities and engaging in economic activities and in transactions with other entities. Thus, an institutional unit is entitled to own goods or assets in its own right; to exchange ownership of goods or assets in transactions with other institutional units, is able to take economic decisions and engage in economic activities for which it is itself held to be directly responsible and accountable at law, is able to incur liabilities on its own behalf, to take on other obligations or future commitments and to enter into contracts, has a complete set of accounts or it would be possible to compile a complete set of accounts if they were required.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Institutional sector, standard statistical unit

ISIC Rev. 4

International Standard Industrial Classification of All Economic Activities (ISIC) Rev.4 is a standard classification of economic activities arranged so that entities can be classified according to the activity they carry out. The categories of ISIC at the most detailed level (classes) are delineated according to what is, in most countries, the customary combination of activities described in statistical units and considers the relative importance of the activities included in these classes. While ISIC Rev.4 continues to use criteria such as input, output and use of the products produced, more emphasis has been given to the character of the production process in defining and delineating ISIC classes. The groups and divisions, the successively broader levels of classification, combine the activities of producing units according to: similarities in the character of the goods and services produced, the uses to which the goods and services are put, and the inputs, process and technology of production.

ISIC is widely used nationally and internationally in classifying data according to kind of economic activity in the fields of production, employment and other areas. ISIC is a basic tool for studying economic phenomena, fostering international comparability, providing guidance for the development of national classifications and for promoting the development of sound national statistical systems. In providing more up-to-date detail, this revision of the classification provides a closer representation of current economic reality. In addition, the Revision 4 of ISIC has improved comparability with other regional activity classifications in use around the world

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: NACE Rev. 2

Jobs

Individuals may have more than one source of income from employment because they work for more than one employer or, in addition to working for one or more employer, they work on their own account as self-employed. The agreement between an employee and the employer defines a job. The number of jobs in the economy thus exceeds the number of person employed due to the extent that employees have more than one job.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Employment

Joint venture

A joint venture is created when two or more independent enterprises agree to commit some of their resources to work together on a common project or towards a common goal. An important feature of a joint venture enterprise is that none of the original enterprises exercise outright control over the entity created, therefore it is considered an enterprise. For business demography purposes, joint ventures may be considered real births if they involve the creation of new factors of production. The cessation of a joint venture mirrors the above. It can be considered a real death if less than half of the employment is transferred to the participating enterprises. The proportion of the new factors of production necessary for a joint venture to be considered a real birth should be at least half, i.e. if less than half of the total employment of the joint venture enterprise is transferred from the participating enterprises, it is considered a real birth.

Source: Business registers – Recommendations manual (Eurostat, 2010), paragraph 13.23; and Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007), Section 4.1 and 5.1.

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf and

http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Kind-of-activity unit (KAU)

A kind-of-activity unit is an enterprise or part of an enterprise that engages in only one kind of productive activity or in which the principal productive activity accounts for most of the value added. Compared with the establishment, the KAU is not restricted on the geographic area in which the activity is carried out but it is characterized by homogeneity of activity.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Establishment

Legal form

The legal form (also known as legal status) is defined according to national legislation. It is useful for eliminating ambiguity in identification searches and as the possible criterion for selection or stratification for surveys. It is also used for defining the institutional sector. Statistics according to legal form are produced e.g. in business demography. The character of legal or natural person is decisive in fiscal terms, because the tax regime applicable to the unit depends on this. It means that any statistical register fed with fiscal records will have that information. Experience has shown that legal form will often be useful to make adjustments to information collection processes and questionnaires on the legal unit operating an enterprise. A code representing the legal form should therefore be recorded in accordance with the classification of legal forms or categories. The following legal forms can be found in most countries:

- Sole proprietorship: Enterprise owned exclusively by one natural person.
- Partnership: Association of persons who conduct a business under a collective name. It can take the form of a limited partnership.
- Limited liability companies: Enterprises comprising joint-stock companies, limited partnerships with share capital and private limited company.
- Co-operative societies: These are bodies set down by law in each country. They observe a number of general principles, for example they may only be entitled to provide their services to members, profits are often distributed in proportion to members' dealings with the society, etc.
- Non-profit institutions.
- Enterprises with other forms of legal constitution: This group includes non taxative nationalised industries in form of publicly-owned enterprises and state or local authority monopolies.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Legal person, legal unit

Legal person

The term "legal person" corresponds to all forms of legal construction organised by the constitution and laws of countries and endowed with rights and obligations characteristic of legal personality.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Legal form, legal unit, natural person

Legal unit

Legal units include:

- Legal persons whose existence is recognised by law independently of the individuals or institutions which may own them or are members of them.
- Natural persons who are engaged in an economic activity in their own right.

The legal unit is usually recorded in one or more administrative sources. The sources used for statistical business registers do not necessarily provide identical views of legal units. These units can vary both between different sources within a country and between countries. Thus, the legal unit is not suitable as a statistical unit, particularly for international comparisons. The characteristics of a legal unit are: it owns goods or assets, it incurs liabilities and it enters into contracts. The legal unit always forms, either by itself or sometimes in combination with other legal units, the basis for the statistical unit known as the "enterprise".

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Legal form, legal person, natural person

Live register

An important role of the SBR is to maintain and keep track of changes in statistical units and their characteristics. Maintenance is a continuous process in which constant modifications of the set of statistical units occur over time. The extent of the modifications depends on the update strategy of the SBR. In this respect the SBR is considered a live register in which the composition and characteristics of units constantly change over time.

The live register is a vehicle for bringing together data from the various sources that provide the basis for derivation of statistical units. It is the starting point for communication with data source owners and for discussions about units. Legal units are usually the building blocks for creating statistical units. In some countries the SBR is the only environment in which legal units of all forms are brought together. Statistical units are created in the live register.

Source: Guidelines on Statistical Business Registers (UNECE, 2015)

Link: http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

Related terms: Administrative business register

Local kind-of-activity unit (local KAU)

See: Establishment

Local unit

A local unit is an enterprise or part of an enterprise (for example, a workshop, factory, warehouse, office, mine or depot) that is engaged in productive activity at or from one location. The definition has only one dimension in that it does not refer to the kind of activity that is carried out.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Enterprise, statistical unit

Longitudinal data

Longitudinal data is data collected from the same individuals (or household, businesses or other entities) repeatedly over time. The data can show how actions and events can affect outcomes later in life.

Source: National Center for Longitudinal Data, Australia Government, Department of Social Services.

Link: https://www.dss.gov.au/about-the-department/national-centre-for-longitudinal-data

Longitudinal identifier

The longitudinal identifier serves to identify a longitudinal unit (continuing statistical or legal unit). A longitudinal or continuing unit can be associated to a single or, following business demographic event(s) (see definition of Event), to multiple statistical or legal units (see definition of Continuity and Survival). Longitudinal identifiers are required to link the multiple statistical or legal units to their corresponding longitudinal unit.

Related terms: Identifier, Longitudinal unit

Longitudinal unit

The longitudinal unit is a unit that is deemed to continue (see definition of Continuity) over two or more consecutive reference periods.

Related terms: Continuity, Standard Statistical Unit

Main activity

See: Activity, principal activity

Main location

The main location of an enterprise is the location of the local unit with the largest number of persons actually employed.

Source: Business registers – Recommendations manual (Eurostat, 2010), paragraph 14.11

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Market producer

A market producer is a corporation that is created for the purpose of producing goods or services for sale on the market at prices that are economically significant.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Non-market output, production unit

Master frame

See: Frozen frame

Match

A pair that represents the same entity in reality.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Matching

Matching

Matching is the process of linking data from different sources. There are various forms of matching, including:

- exact matching (record Linkage) - linking corresponding records from two or more sources using a common identifier; and
- probabilistic matching - determining a probable link between records from two or more sources using an algorithm based on common fields, e.g. name, address, economic activity code.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: False match, false non-match, match, non-match,

Merger

Enterprises may integrate to the extent that the number of existing enterprises is reduced. If two enterprises integrate entirely, the enterprises involved may lose their identity because they are dissolved beyond recognition in the new organisation. If both enterprises lose their identity, the event is called a merger. There is no continuity or survival, but the closures of the previous enterprises are not considered to be real deaths. Similarly, the new enterprise is not considered to be a real birth. This event can be seen as the opposite of a break-up.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Enterprise, take-over (of enterprise)

NACE Rev. 2

NACE ("Nomenclature générale des Activités économiques dans les Communautés Européennes" - Statistical classification of economic activities in the European Communities) is the EU standard classification of economic activities. NACE presents the universe of economic activities partitioned in such a way that a NACE code can be associated with a statistical unit carrying them out. NACE provides the framework for collecting and presenting a large range of statistical data according to economic activity in the fields of economic statistics (e.g. production, employment, national accounts) and in other statistical domains. NACE is derived from ISIC, in the sense that it is more detailed than ISIC. ISIC and NACE have exactly the same items at the highest levels, where NACE is more detailed at lower levels. The coding of the NACE nomenclature comprises:

- a first level consisting of headings identified by an alphabetical code (sections);
- an intermediate level consisting of headings identified by a two-character alphabetical code (subsections);
- a second level consisting of headings identified by a two-digit numerical code (divisions);
- a third level consisting of headings identified by a three-digit numerical code (groups); and
- a fourth level consisting of headings identified by a four-digit numerical code (classes).

Source: NACE Rev. 2 - Statistical classification of economic activities in the European Community (Eurostat, 2008)

Link: http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

Related terms: Activity, ISIC Rev. 4

Natural person

The term natural person is used by the law and by many administrative authorities to denote a human being endowed with all the rights constituting legal personality.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Legal person

New enterprise

See: Birth (of enterprise), creation (of business)

Non-financial corporation

See: Institutional sector

Non-match

A pair that represents two different entities in reality

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Matching

Non-profit institution (NPI)

Most NPIs are separately identified institutional units. That is, they are capable in their own right of owning assets, incurring liabilities and engaging in economic activities and in transactions with other entities. It follows that a complete set of accounts for the unit, including a balance sheet of assets and liabilities, exists or could be constructed if required. NPIs are categorized as follows:

a. those providing services to corporations whose output is sold to the corporations concerned and treated as intermediate consumption;

b. those that are controlled by government and provide individual or collective services on a non-market basis; and

c. those providing goods and services to households, divided between: those that provide goods and services to individual households at economically significant prices; those providing services to individual households free or at prices that are not economically significant and those that provide collective services free or at prices that are not economically significant.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Non-market output, NPISH

Non-profit institution serving households (NPISHs)

See: Institutional sector

Number of active enterprises

A count of enterprises that had either turnover or employment at any time during a given reference period.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007), Section 3.1

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Number of births of enterprises

A count of the number of births of enterprises registered to the population concerned in the business register corrected for errors. A birth amounts to the creation of a combination of production factors with the restriction that no other enterprises are involved in the event. Births do not include entries into the population due to mergers, break-ups, split-offs or restructuring of a set of enterprises. It does not include entries into a sub-population resulting only from a change of activity.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Enterprise birth, reactivation

Number of deaths of enterprises

A count of the number of deaths of enterprises registered to the population concerned in the business register corrected for errors. A death amounts to the dissolution of a combination of production factors with the restriction that no other enterprises are involved in the event. Deaths do not include exits from the population due to mergers, take-overs, break-ups or restructuring of a set of enterprises. It does not include exits from a sub-population resulting only from a change of activity.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Enterprise death

Number of employees

Number of employees is defined as those persons who work for an employer and who have a contract of employment and receive compensation in the form of wages, salaries, fees, gratuities, piecework pay

or remuneration in kind. A worker from an employment agency is considered to be an employee of that employment agency and not of the units in which they (temporarily) work.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms Employees, employment, number of persons employed

Number of employees in full-time equivalent units

This heading is defined as a count of the number of employees converted into full time equivalents (FTE).

Figures for the number of persons working less than the standard working time of a full-year full-time worker should be converted into full time equivalents, with regard to the working time of a full-time full-year employee in the unit. Included in this category are people working less than a standard working day, less than the standard number of working days in the week, or less than the standard number of weeks/months in the year. The conversion should be carried out on the basis of the number of hours, days, weeks or months worked.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Employees, employment, number of persons employed

Number of persons employed

Number of persons employed is defined as the total number of persons who work in the unit, including wage-earners and self-employed persons (i.e., working proprietors, partners working regularly in the unit and unpaid family workers) as well as persons who work outside the unit but who belong to it and are paid by it (e.g., sales representatives, delivery personnel, repair and maintenance teams). It excludes manpower supplied to the unit by other enterprises, persons carrying out repair and maintenance work in the unit on behalf of other enterprises, as well as those on compulsory military service.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Employees, employment, number of employees

Over-coverage

Units which are included in the sampling frame but do not belong to the target population. These cases are usually observed for contacted units, but not necessarily for non-contacted units or those excluded from a sample. Reasons for over-coverage are the death of units, misclassification and a non-updated frame. The over-coverage in a register generally biases the estimators down from that sampling list.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Coverage, frame, under-coverage

Partnership

See: Legal form

Persons employed

See: Number of persons employed

Population of active enterprises

See: Number of active enterprises

Principal activity

The principal activity is the activity which contributes most to the total value added of the unit under consideration. Ideally, the principal activity of the unit should be determined with reference to the value added to the goods and services produced, by applying the top-down method. The top-down method follows a hierarchical principle: the classification of the unit at the lowest level of the classification must be consistent with the classification of the unit at the higher levels. The principal activity so identified does not necessarily account for 50% or more of the entity's total value added. In case that the value added is not known, substitute criteria can be used as proxies to obtain the best approximation possible compared to the result which would have been obtained based on value added data.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Activity, ancillary activity, economic activity, secondary activity

Producer

See: Production unit

Producing unit

See: Production unit

Production factors

See: Factors of production

Production unit

A production unit carries out an economic activity under the control and responsibility of an institutional unit using inputs of labour, capital and goods and services to produce outputs of goods and services. Enterprises can be very heterogeneous if they have several secondary activities that are quite different from their principal activities. In order to obtain groups of producers whose activities are more homogeneous, enterprises are partitioned into smaller and more homogeneous units of production like local units, kind-of-activity units, and establishments.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Economic unit, institutional units, market producer, product, statistical unit

Property

See: Characteristic

Reactivation

This event involves an enterprise becoming dormant for a period of less than two years, then recommencing activity in a way that complies with the definition of continuity. In terms of business demography this event does not constitute a birth or death.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Dormant unit

Real birth

See: Enterprise birth

Real death

See: Enterprise death

Register [business registers manual]

See: Statistical business register

Register picture

See: Register snapshot

Register snapshot

A copy of all the statistical units in the live register as of a given point in time, including all administrative units or links from statistical units to administrative units. The snapshot is an intermediary step between the live register and the frozen frame. It is used to check for errors that have crept in during processing since the previous snapshot, also as the basis for a historical record. It may contain many inactive enterprises.

Source: Guidelines on Statistical Business Registers (UNECE, 2015)

Link: http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

Related terms: Frozen frame, live register

Restructuring

Restructuring within an enterprise does not affect the continuity of the enterprise, only its structure. An example could be the creation or deletion of a local unit. Restructuring may affect key characteristics such as size or principal activity. It could be argued that this is not really a demographic event at the level of the enterprise and does not impact on the demographic variables relating to the enterprise, but it may affect the way the enterprise is included in demographic statistics. Restructuring within an enterprise group is a change (e.g. creation and/or cessation of one or more enterprises) involving more than one enterprise before and more than one enterprise after the event, where all enterprises involved are under common control. It affects the identity of at least one enterprise, though the total number of enterprises before and after the event may be the same. A typical example is the complete reorganisation of the production capacity of a large enterprise group. Complex restructuring is a similar event, but this is not constrained to one enterprise group. Restructuring within an enterprise group, or complex restructuring, may entail any number of register creations and deletions.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Continuity, survival

Revenue

See: Turnover

Sampling Frame

See: Frame

Satellite

A statistical business register may be quite a complex network of databases and functionalities. An approach that can be used to extend the functionality with minimum complication is to extract parts or the whole of one or more frozen frames from the SBR and afterwards link information from other data sources to its units. The resulting product, which is maintained outside and independent of the SBR, is called an SBR satellite. The responsibility for and control of a satellite are separated from the SBR and usually take place in a different environment.

Source: Guidelines on Statistical Business Registers (UNECE, 2015)

Link: http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

Related terms: Statistical business register

Secondary activity

A secondary activity is each separate activity that produces products eventually for third parties and that is not the principal activity of the unit in question. The outputs of secondary activities are secondary products. Most economic entities produce at least some secondary products.

Source: International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Activity, ancillary activity, economic activity, principal activity

Self-employed

Self-employed persons are persons who are the sole or joint owners of the unincorporated enterprises in which they work. Persons who work in unincorporated enterprises are considered self-employed f they are not in paid employment that constitutes their principal source of income; in that latter case, they are classified as employees.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Employment

Small and medium-sized enterprises (SMEs)

Small and medium-sized enterprises (SMEs) are non-subsidiary, independent firms which employ less than a given number of employees. This number varies across countries. The most frequent upper limit designating an SME is 250 employees, as in the European Union. However, some countries set the limit at 200 employees, while the United States considers SMEs to include firms with fewer than 500 employees. Small firms are generally those with fewer than 50 employees, while micro-enterprises have at most 10, or in some cases 5, workers. Financial assets are also used to define SMEs. In the European Union, a new definition came into force on 1 January 2005 applying to all community acts and funding programmes as well as in the field of state aid where SMEs can be granted higher intensity of national and regional aid than large companies. The definition includes the financial ceilings: the turnover of medium-sized enterprises (50-249 employees) should not exceed EUR 50 million; that of small enterprises (10-49 employees) should not exceed EUR 10 million while that of micro firms (less than 10 employees) should not exceed EUR 2 million. Alternatively, balance sheets for medium, small and micro enterprises should not exceed EUR 43 million, EUR 10 million and EUR 2 million, respectively.

Source: OECD Glossary of statistical terms

Link: http://stats.oecd.org/glossary/

Related terms: Enterprise

Sole proprietorship

See: Legal form

Special purpose entity (SPE)

There is no common definition of an SPE but some of the following characteristics may apply. SPEs often have no employees and no non-financial assets. They may have little physical presence beyond a "brass plate" confirming their place of registration. They are always related to another corporation, often as a subsidiary, and are often resident in another territory than the related corporations. In the absence of any physical dimension to an SPE, its residence is determined according to the economic territory under whose laws the enterprise is incorporated or registered. Entities of this type are commonly managed by employees of another corporation which may or may not be a related one. The unit pays fees for services rendered to it and in turn charges its parent or other related corporation a fee to cover these costs. This is the only production the unit is involved in though it will often incur liabilities on behalf of its owner and will usually receive investment income and holding gains on the assets it holds. Whether a unit has all or none of these characteristics, and whether it is described as an SPE or some similar designation or not, it is treated in the SNA in the same way as other institutional units by being allocated to sector and industry according to its principal activity unless it falls into one of the three following categories: Captive financial institutions, artificial subsidiaries of corporations, or special purpose units of general government.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Institutional unit

Split-off (of enterprise)

Split-off involves one enterprise before and more than one enterprise after the event. In a split-off, the new enterprise(s) is (are) generally much smaller and the identity of the original enterprise is retained by the larger enterprise. There is no death, but one or more new enterprises are created. This event can be seen as the opposite of a takeover.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Break-up (of enterprise), new enterprise

Standard statistical unit

Statistical units are the units for which information is sought and for which statistics are ultimately compiled. Commonly used types of statistical units for economic statistics are the enterprise, the enterprise group, the kind-of-activity unit (KAU), the local unit and the establishment (in the EU called local kind-of-activity unit (LKAU)). In national accounts also the institutional unit is of importance.

Source: International Standard Industrial Classification of all Economic Activities, Rev. 4, (United Nations, 2008)

Link: https://unstats.un.org/unsd/publications/catalogue?selectID=396

Related terms: Legal unit, institutional unit

Statistical business register (SBR)

The statistical business register is a full and comprehensive, regularly updated and structured list of business units engaged in the production of goods and services, which is maintained by national statistical authorities for statistical purposes to assist the compilation of statistical data and particular as a (backbone) tool for the preparation and coordination of surveys, as a source of information for statistical analysis of the business population and its demography, for the use of administrative data, and for the identification and construction of statistical units.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Administrative business register

Status of Employment classification

The International Classification of Status in Employment (ICSE-18) classifies jobs in employment for pay or profit into ten detailed categories based on the concepts of type of authority and type of economic risk. These categories may be aggregated according to two alternative classification hierarchies: the International Classification of Status in Employment according to type of authority (ICSE-18-A) and the International Classification of Status in Employment according to type of economic risk (ICSE-18-R).

Source: (ILO, 2018)

Link: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms_648693.pdf

Statistical unit

See: Standard statistical units

STEC

Statistics on service trade by enterprise characteristics (STEC)

Link: http://ec.europa.eu/eurostat/web/international-trade-in-services/methodology/services-trade-by-enterprise-characteristics-stec-

Structural business statistics

Structural business statistics describe the structure, main characteristics and performance of economic activities. They can provide answers to questions on the creation of wealth (value added), investment and labour input of different economic activities. The data can be used to analyse structural shifts, for example between industry and services, country specialisations in particular activities, sectoral productivity and profitability, as well as a range of other topics. Because they are available broken down by enterprise size class, structural business statistics also permit a detailed analysis of small and medium-sized enterprises (SMEs), which is of particular use to EU policymakers and analysts wishing to focus on entrepreneurship and the role of SMEs. Furthermore, structural business statistics provide useful background information on which to base an interpretation of short-term statistics and the business cycle.

Source: Eurostat: Statistics Explained: Structural business statistics overview (2017).

Link: http://ec.europa.eu/eurostat/statistics-explained/index.php/Structural_business_statistics_overview

Survey Frame

See: Frame

Survival

In general, survival occurs when a unit is active and identifiable both before and after a specific (business) demographic event. The unit may be changed in some way, e.g. in terms of economic activity, size, ownership or location, but there should be continuity of the unit reference number in the statistical business register. In business demography context, survival occurs if an enterprise is active in terms of employment and/or turnover in the year of birth and the following year(s). An enterprise is also considered to have survived if the linked legal unit(s) have ceased to be active, but their activity has been taken over by a new legal unit set up specifically to take over the factors of production of that enterprise (= survival by take-over).

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Business demography, continuity

Survival [business demography]

In the business demography context, survival occurs if an enterprise is active in terms of employment and/or turnover in the year of birth and the following year(s). Two types of survival can be distinguished:

1)   An enterprise born in year t or having survived to year t is considered to have survived in year t+1 if it is active in terms of turnover and/or employment in any part of year t+1 (= survival without changes).

2)   An enterprise is also considered to have survived if the linked legal unit(s) have ceased to be active, but their activity has been taken over by a new legal unit set up specifically to take over the factors of production of that enterprise (= survival by take-over).

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Continuity

Survival rate (of enterprises)

The survival rate of newly born enterprises in a given reference period is the number of enterprises that were born in year t-n and survived to year t as a percentage of all enterprises born in year t-n.

Source: Eurostat – OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Employer enterprise survival rate

System of National Accounts (SNA)

The System of National Accounts (SNA) is the internationally agreed standard set of recommendations on how to compile measures of economic activity in accordance with accounting conventions based on economic principles. The recommendations are expressed in terms of a set of concepts, definitions, classifications and accounting rules that comprise the internationally agreed standard for measuring of economic activities. The accounts present in a condensed way a mass of detailed information, organized according to economic principles and perceptions, about the working of an economy. They provide a comprehensive and detailed record of the complex economic activities taking place within an economy and of the interaction between the different economic agents, and groups of agents, which takes place on markets or elsewhere. The SNA is designed to give a realistic and compact view of the economy that is suitable for policy and analytical use.

Source: System of National Accounts, 2008 (United Nations, 2009)

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: European System of Accounts (ESA)

Take-over (of enterprise)

Enterprises may integrate to the extent that the number of existing enterprises is reduced. If two enterprises integrate entirely, one of the enterprises may remain largely the same. In this case the other enterprise is generally much smaller, it is merely absorbed by the larger enterprise, which remains the same. If one of the enterprises keeps its identity, the event is called a take-over. Enterprises taken over are not considered to be real deaths. In this case, one of the original enterprises does survive in a recognisable form, and therefore there is both continuity and survival. The remaining original enterprises are closed. This event can be seen as the opposite of a split-off.

Source: Business registers – Recommendations manual (Eurostat, 2010)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Cessation (of business), death (of enterprise), merger

Take-over (of enterprise group)

Like enterprises, enterprise groups may have many kinds of intergroup relations and integrate their operations partly or totally. Two (or more) enterprise groups may integrate entirely and become one group. In this process either both groups involved may lose their identity, because they are dissolved beyond recognition in the new organisation, or one group may remain largely the same. In the latter case the other group is generally much smaller; it is merely absorbed by the larger group, which remains largely the same. If one of them keeps its identity, it is called a take-over.

Source: Business registers – Recommendations manual (Eurostat, 2010); and Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf, and

http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: Cessation (of business), death (of enterprise group), merger

Trade by enterprise characteristics (TEC)

Statistics on trade in goods by enterprise characteristics (TEC).

Link: http://ec.europa.eu/eurostat/web/international-trade-in-goods/data/focus-on-enterprise-characteristics-tec

Turnover

Turnover comprises the totals invoiced by the observation unit during the reference period, and this corresponds to market sales of goods or services supplied to third parties. Turnover also includes all other charges (transport, packaging, etc.) passed on to the customer, even if these charges are listed separately in the invoice.

Turnover excludes VAT and other similar deductible taxes directly linked to turnover as well as all duties and taxes on the goods or services invoiced by the unit. Reduction in prices, rebates and discounts as well as the value of returned packing are not taken into account. Income classified as other operating income, financial income and extraordinary income in company accounts is excluded from turnover.

Source: Business registers – Recommendations manual (Eurostat, 2010), p. 340

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Related terms: Revenue

Under-coverage

Failure to include all units belonging to the defined study population in the frame. This mainly includes (new) enterprises not included in the frame, either through real birth, or break-ups and split-offs, or misclassified units. This generally leads to biases in the estimators.

Source: Eurostat-OECD Manual on Business Demography Statistics (Eurostat & OECD, 2007)

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07-010-EN.pdf

Related terms: coverage, frame, over-coverage

Unincorporated enterprise

An unincorporated enterprise represents the production activity of a government unit, NPISH or household that cannot be treated as the production activity of a quasi-corporation. An unincorporated enterprise is a producer unit which is not incorporated as a legal unit separate from the owner (household, government or foreign resident); the fixed and other assets used in unincorporated enterprises do not belong to the enterprises but to their owners, the enterprises as such cannot engage in transactions with other economic units nor can they enter into contractual relationships with other units nor incur liabilities on their own behalf; in addition, their owners are personally liable, without limit, for any debts or obligations incurred in the course of production.

Source: System of National Accounts, 2008 (United Nations, 2009), paragraphs 4.21 and 5.1

Link: http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

Related terms: Enterprise

Value added tax (VAT)

A value added type tax (VAT) is a tax on goods and services collected in stages by enterprises and which is ultimately charged in full to the final purchasers. This heading comprises the value added tax which is collected by the General government and which is applied to national and imported products, as well as other deductible taxes applied under similar rules to those governing VAT. The common feature of VAT is that producers are obliged to pay to the government only the difference between the VAT on their sales and the VAT on their purchases for intermediate consumption and gross fixed capital formation.

Source: European System of Accounts 2010 (Eurostat, 2013), paragraph 4.17

Link: http://ec.europa.eu/eurostat/documents/3859598/5925693/KS-02-13-269-EN.PDF/44cd9d01-bc64-40e5-bd40-d17df0c69334

Variable

The term equals the meaning of the term "characteristic" but is more appropriate in the case of surveys where sampling is involved and thus the notion of variability due the probability mechanism applied for selecting samples.

Related terms: Characteristic

Annex 1: Developing and using longitudinal businesses data in Canada

1   Introduction

1.   An SBR, referred to at Statistics Canada and in this annex as the Business Register (BR), has been an invaluable pillar supporting Statistics Canada's ability to provide high-quality statistical information on the Canadian economy. Its main purpose is to serve as the sampling frame for Statistics Canada's business surveys. It also plays a central role in the modernization of the Agency's approach to producing economic statistics with the use of administrative data, data collected by government agencies and/or private sector companies. In this expanding role, it serves as a hub through which a wide range of external data sources can be linked to each other and to Statistics Canada's surveys. This hub facilitates data integration, which allows the Agency to improve data quality, reduce costs and respondent burden, and meet emerging needs in a timely fashion.

2.   Emerging needs include a demand for more disaggregated statistics - mostly at the geographical level but also at the business owner level. Indeed, some business-related topics are garnering renewed interest on the part of the policy community and the general public: entrepreneurship, start-ups, foreign control, patents and innovation to name a few. These emerging needs require micro-level business data on different units of analysis (ultimate parents, legal entities, locations and business owners) in order to study firm behaviour to changing macro-conditions.

3.   This expanding role has pushed the capabilities of the current BR beyond its original purpose. Not only is it being called upon to help ensure consistency and linkage across sources at a point in time, but also to ensure the consistency and linkage of data across time at a very detailed level. Preferably, this detailed longitudinal consistency would extend down to the locations of a business, and even to the business owners.

4.   In response to the need for more timely and detailed data on businesses, as well as for new types of business information, Statistics Canada decided to embark on developing new longitudinal modules on the BR systems. These modules would include longitudinal identifiers, predecessor/successor tables with detailed information, as well as historical updates of the past data to reflect the most recent information available. The new features in the BR will provide the infrastructure for a more rapid response to the evolving needs for detailed longitudinal business data, and allow for a greater consistency in the statistics produced across the Agency.

5.   That said, the BR as it is today, and as it was decades ago, has supported the production of longitudinal statistics and analysis since the 1980s. Statisticians and economists developed methodologies to exploit the 'cross-sectional' BR to its fullest extent in order to inform precise research questions and to produce official statistics. This annex reviews examples of these data sets and the techniques used to create the longitudinal structure tailored to a precise purpose. This will give insights into methods that could be adopted in the creation of a more general-purpose longitudinal structure residing within a statistical business register. The lessons learned in creating those longitudinal databases as well as in producing official statistics and research with them, will guide Statistics Canada in the development of its new longitudinal modules. The datasets are presented in the chronological order in which they were initially created. Some aspects of the Canadian business administrative data and the Canadian BR are explained in detail in order to facilitate comparisons between this country and others.

2   Longitudinal Annual Survey of Manufactures (ASM)

6.   For over two decades, a longitudinal file from the Annual Survey of Manufacturing and Logging (known previously as the Census of Manufactures, and Annual Survey of Manufactures) has been created to support research and analysis at Statistics Canada.[24] The comparative strength of this file came from the fact that plants were assigned a unique identification number that persisted despite changes in name or ownership (a special feature of the survey, absent from the BR itself). Each plant could be linked to its owning firm through the BR. This permitted the construction of consistent files of plants and firms over long time spans despite changes in census thresholds, industrial classification systems, and redesign of the manufacturing census (e.g. 1961-1990, 1973-1999, 1990-2010 and 2000-2012). The greatest advantage came from being able to divide changes occurring at the firm level because of changes in the underlying plants (Baldwin, 1995).

7.   The opportunity to develop a longitudinal database with detailed information at the plant level demonstrated the advantages of this construct. First, legal entities in the BR did not necessarily have to be consistently defined, as the consistent plant-level information could be used to validate the consistency of the legal entity information in the BR. For example, a change in the identity of a legal entity for a continuing plant could be flagged for further analysis to determine whether the change in legal entity was genuine or not (like a change in name for marketing purposes). Second, the lifecycle and activities of the legal entity could be characterized by the lower level data which provided richer detail on the extensive amount of change that is taking place at that level (i.e. a firm could be both creating new plants itself and acquiring/divesting from other firms).

8.   As (Baldwin, Dupuy, & Penner, 1992) pointed out, all this does not suggest that a business register needs to adopt this approach. It may be inordinately expensive to keep up-to-date longitudinal data of plants or locations. However, these costs may lower in the future with improvements in geocoding and as on-line details about business locations become more readily available. As previously mentioned, there is demand for detailed statistics on small areas. The source of the basic data may be the greatest challenge for progress in this area in Canada since the current administrative sources are at the level of the tax-filling entity, not the operating location (discussed more in depth below). Even with the successful development and adoption of methodologies to provide small area statistics through estimation or allocation methods[25], some data at the more detailed geographic level is needed as inputs in the methods.

3   Longitudinal Employment Analysis Program (LEAP)

9.   The LEAP has been used to produce annual measures of entries, exits and continuing employer enterprises, and associated employment creation and destruction since 1980 (Statistics Canada, 1988). It was developed to show how administrative data could produce statistics at a lower cost than traditional survey methods.[26]

10.   The current LEAP database uses the T4 Statement of Remuneration Paid from the Canada Revenue Agency (CRA). Each year, Canadian businesses must issue T4 slips to their employees for personal income tax purposes. A T4 slip details, among other things, the employment earnings and tax deductions for an employee at a particular business during the year. Each employee needs his/her T4

---

[24] Examples of research produced with the longitudinal ASM include (Baldwin & Gu, 2004), (Baldwin & Gu, 2006) (Baldwin, Dar-Brodeur, & Yan, 2016), and (Gu & Li, 2017)

[25] Allocation methods are used by (Hidiroglou & You, 2016), (Brown & Rispoli, 2014), and (Leung, Rispoli, & Gibson, 2011)

[26] To make the data as widely accessible as possible, estimates were initially made available by print and CD, and then on-line when Statistics Canada's CANSIM database became free to users in 2012.

slip to file his/her personal income taxes, and the penalties that employers face for not issuing T4s make the file comprehensive.[27]

3.1 The Business Number and the Business Register

11.   The business identifier used by the CRA is a Business Number (BN) - a business needs to apply for one before filing documents with the CRA (income tax, sale taxes, employer information, import and export activities). A BN represents a tax-filing entity with the CRA, not an entity of interest for statistical purposes. The CRA supplies Statistics Canada with a list of BNs (and basic information like legal name, address of head office, and industry) on an on-going basis, and this list is used as an input into the BR. A statistical enterprise in the BR is defined as the unit that reports a complete set of financial statements (income statement and balance sheet) – this can represent one or many BNs depending how the business decides to interact with the CRA. A BN may cover more than one of the statistical enterprise's locations, and even only a part of a location. The vast majority of enterprises (about 99%) have one BN and location, but the complex enterprises with multiple BNs and locations account for about 50% of business activity.

12.   Businesses can undergo many transformations: change in ownership, change in name, expansion into new domestic and international markets, change in legal structure, and merger and acquisition activity. These transformations, either substantive or purely administrative, may trigger the creation of new BNs.

13.   When the BR receives the list of new BNs from the CRA, it must decide how to integrate the new information. Given the volume involved, the priority is given to keeping up to date the profiles of BNs that have associated with them a larger level of economic activity. BNs with little or no associated economic activity, or with little apparent connection with an existing firm are less of a priority. They are added in the BR as tax-filing entities, but they do not get classified (industry and size) and are not part of the sampling frame for some time. In a similar fashion, the BR reactivates BNs that had become inactive in the past and are becoming active again only with some delay.

14.   When producing the LEAP database, analysts first link the T4 data to the BR in order to assign statistical enterprise numbers to the BNs. Table A1.1, reproduced from (Rollin, 2013), shows the proportion of BNs for calendar years 1999 to 2009 that can be found in the BR files for January in the following year. The proportion that can be assigned a statistical enterprise number falls in the 70 to 79 percent range. Table A1.2, also reproduced from (Rollin, 2013), shows that in terms of the T4 payroll, the coverage rate improved post-2005, with only less than 5% of payroll not assigned to a statistical enterprise.

15.   The T4 data are available only with a lag. Thus, it is possible to use more than just the following January BR file to assign a statistical enterprise. (Rollin, 2013) shows that many cases can be resolved by looking at other BR files. In the current LEAP, this is reflected by that fact that only 7% of active employer businesses in 2015 cannot be assigned an industry (are in the unclassified category). This percentage is higher for entrants as defined by LEAP, at 21% for 2015. This is expected, as the BR has the most difficulty in providing up-to-date information on the newest units.

3.2 Turning LEAP into a longitudinal file

16.   The linked database is then made longitudinal. Since the principal purpose of the BR has been to serve as a sampling frame for surveys, the BR focuses primarily on creating an accurate snapshot of the business population at a point in time. Additional procedures therefore must be implemented to ensure that the cross-sectional links from the BN to the enterprise number are also longitudinally consistent.

---

[27] The T4 file is one of the main input into the estimate of wages and salaries in Statistics Canada's Macroeconomic Accounts.

17.   The continuity of the enterprise can be established by looking at the continuity of its production factors such as its employees, buildings, machines and equipment, land, and management (Eurostat, 2010, p. 129); (Eurostat & OECD, 2007, p. 26). The continuity of employment was chosen for LEAP to remove false deaths and births (Baldwin, Dupuy, & Penner, 1992). Given the analytical purpose of the LEAP, which is to study entry and exit and the associated job creation and job destruction, this seemed an especially pertinent approach. Since the T4 slips contain both business and individual identifiers, the creators of LEAP decided to track the movement of individuals across businesses and developed a "labour tracking" technique to improve the longitudinal consistency of the database. In particular, they needed to detect when births and deaths were 'false', and link predecessor and successor firms over time when this occurred.

Table A1.1 **Proportion of filing business numbers, by presence in next January BR and by calendar year**

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
In next January BR, and enterprise identifier found	73	73	72	71	71	71	70	68	77	76	75
In next January BR, but no enterprise identifier found	20	17	19	19	21	24	29	30	22	23	23
Not in next January BR	7	10	9	10	7	5	1	1	1	1	1

Table A1.2 **Share of total economic activity for filing business numbers with missing enterprise identifier**

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
T4 Payroll	27	24	20	22	23	15	4	4	3	3	2

18.   The technique adopted identified the possible employer entrants and exits with at least five employees. Employees of the employers that purportedly exited would be tracked to employers in the next year, and employees of employers that purportedly entered would be tracked to employers in the previous year. Data from this tracking would be compiled in a supplementary database during processing. In this supplementary database, data for each exit and entrant would be combined with data for up to five target businesses with the largest number of common workers.

19.   Criteria to be considered a target for LEAP are shown in Table A1.3. As the size of the potential entrant or exit decreases, the shared percentage of employees necessary to be a target increases. A potential exit with five employees would be linked to another business in the supplementary database, if in the following year all of its workers moved to that other business. Using this supplementary database, a name matching algorithm is used to flag identical or similar business names, some automatic longitudinal matches are made, and a file is outputted for manual review.

Table A1.3 **Criteria used in LEAP to identify potential linkages across time using labour tracking**

Size (number of employees) of potential entrant or exit	251+	51 to 250	16 to 50	8 to 15	6 to 7	5
% shared workers in target firm	25%	30%	50%	60%	70%	100%

20.   Labour tracking implicates roughly 6,000 to 7,000 employers per year out of roughly 300,000 entrants and exits. Most are simple one-to-one relationships, but many also involve large employers. Thus, while the impact of labour tracking on entry and exit statistics are minor, the impact on employment creation and destruction can be substantial.

21.   The final LEAP database is created using the supplementary file by establishing linking and merging rules, which are more difficult to establish for relationships that are not one-to-one match between units over time. The approach adopted in LEAP is to accurately represent the structure of the business population in the most recent year. For example, if in between the most recent year and the previous year two businesses merged to become one business, the LEAP microdata base shows only one business in the previous year. This business in the previous year consists of a single record of the two businesses merged together. In this way, the employment growth of that one continuing business represents the "organic" growth in employment, the growth in employment that abstracts from changes due to possible merger and acquisition activity.

22.   A new LEAP vintage file is created every year, and each vintage pushes the current structure of the business population back in time. For example, the 2015 vintage of LEAP pushes the 2015 structure back to 1983, and the 1999 vintage (the earliest vintage available) pushes the 2000 structure back to 1983. A given vintage provides consistent structures for comparisons over time - a sound approach for many research projects studying organic employment creation and destruction. However, this approach has limited usefulness when producing official statistics on business demographics because the firm structures pushed back in time did not actually exist in the past. Using a particular vintage over long periods of time introduces biases. For example, the proportion of small versus large firms in 2000 is quite different in the 2015 LEAP vintage compared to the 'original' 2000 LEAP vintage. This limitation explains why the LEAP statistics on Statistics Canada's website are based on only the last two years of each vintage. The concepts of entry, exit and continuing employer firm are defined over a two-year window.

23.   Microeconomic longitudinal analysis is possible using the LEAP, but attention must be paid to the research purposes, and whether or how to make best use of all available vintages (see for example, (Dixon & Rollin, 2014)). Long-run analyses need to take into account changes in structure - holding structure constant may not be the ideal methodology for all research projects.

3.3   Lessons learned

24.   The experience of creating the LEAP and using it to produce official statistics and research brings to the fore a number of lessons learned.

   a) Compared to the total number of entrants and exits, the number implicated by the labour tracking is small. Thus, with respect to the count of entrants and exits, the "cross-sectional" BR serves as good starting point for any development.

   b) While labour tracking affects a small fraction of potential entrants and exits, the job creation and job destruction estimates are affected substantially by this technique because the cases that involve labour shifting from one firm to another tend to involve larger firms.

   c) The labour tracking technique has proved to be a good criterion to rely on to link firms over time or to define entry and exit. It is not the only criterion, however. There are other characteristics that define newness or change: name, address, industry, ownership change, incorporation status, transition from private to publicly-traded corporation, etc.

   d) A detailed analysis of potential entrants and exits flags not only real changes in firm activity, but also purely administrative changes or errors in the BR. The creation of longitudinal modules within the BR will likely improve the quality of the BR cross-sectional data.

e) A significant proportion (20%) of new tax-filling entities cannot currently be linked to a BR statistical enterprise over a short time horizon. Considering the demand for statistics on start-ups, novel avenues have to be explored to introduce and classify the new entities more quickly.

f) Producing official statistics over short-time horizons reduces the bias that necessarily occurs over long periods because of the changing nature of the business population. Data and definitions measuring year-on-year changes or changes over a few years produce precise estimates of entry and exit, as well as job creation and destruction.

4 National Accounts Longitudinal Microdata File (NALMF)

4.1 Background

25.   In order to address the need for a database to study research questions besides the employment dynamics of businesses, the LEAP was expanded in subsequent years by adding additional variables from other administrative sources. For example, the LEAP was linked to T2 corporate income tax files to create a T2-LEAP file, a longitudinal file of corporations with employees. The core income and balance sheet variables in T2 (net income, assets, liabilities, equity, total revenues, gross profits, and operating expenses) permitted studies on how firm financing affected the growth and survival of firms, and studies on the profitability of firms. The basic longitudinal structure from the LEAP, however, remained unchanged.

26.   With the codification of the many possible items on a firm's financial statements, and the electronic filings of tax returns, the CRA was able to supply Statistics Canada with more of their administrative data. The use of administrative data from the CRA became widespread in many statistical programs post-2000. This allowed the creation of new business microdata bases that could be benchmarked to the statistical aggregates produced at Statistics Canada. These databases could be used to produce estimates at levels of disaggregation that were possible in the past only with costly large scale surveys.

27.   The NALMF was created in 2011 using key administrative files from the CRA: the T4s, T2 corporate income tax returns, PD7 payroll deductions and remittance accounts, and goods and services tax (GST) or harmonized sales tax (HST) filings. The BR was again the central hub through which all the sources were linked. The detailed profile of the operating locations in the BR was used to allocate the values in the administrative data to detailed industry, and firm size/geographical area categories. These allocated values were used as splitters to produce experimental estimates of GDP by firm size and GDP by Census Metropolitan Area consistent with higher level aggregates from the Macroeconomic Accounts (see (Baldwin, Leung, & Rispoli, 2013) and (Brown & Rispoli, 2014)).

28.   With the creation of the NALMF data base, it was decided to launch a multi-year project to update the techniques used in the LEAP file and to create a new longitudinal structure for NALMF. Whereas the LEAP focused on removing false entrants and exits, more information on predecessor/successor relationships is kept in the NALMF so that custom decisions on the creation of the longitudinal structure can be taken. The predecessor/successor relationship tables document all types of relationships: one-to-one; two-to-one or many-to-one (merger or acquisition); one-to-two or one-to-many (spin-off or divesture); and many-to-many (a continuing entity that acquires and divests some of its parts).

4.2 Updating the tools to create a longitudinal structure

29.   Since the labour tracking methodology in LEAP has been proved over the last two decades to provide useful information in the creation of longitudinally consistent database and has even been adopted by other countries (see (Geurts, 2016) for the Belgium example), it formed a key part of the development of the longitudinal structure of NALMF. The first year of the multi-year project revisited the labour tracking methodology. The goal of the project was to improve the methodology in such a

way as to reduce the manual revisions that were required in the past, and to provide a more formal justification for thresholds, such as the ones shown in Table A1.3. It questioned the appropriateness of the size classes, and the employment threshold for each size class. This required acquiring a deep understanding of the behaviour of continuing firms to be able to identify false births and deaths with greater certainty. More precisely, what is the fraction of employees that remain with the same continuing firm year-on-year, the 'pass-through rate'? Statistical tests compared the distribution of pass-through rate for continuing firms, by size and industry groupings. This allowed the groupings, over which the thresholds would be defined, to be chosen based on the similarity of their pass-through rates. The thresholds could then be determined by a point in the distribution of pass through rates for each grouping. For example, for a certain grouping of firms determined by firm size and/or industry, the labour tracking threshold to identify an entry or exit as a false one can be defined as the pass-through rate above which 75% of continuing firms in that grouping are found. An entry or exit displaying such a high pass-through behaves quite like a continuing firm and is probably 'false'.

30.   In the second year, a project was launched to improve the name matching methodology. This included an environmental scan of how business names were standardized at Statistics Canada, the development of a common approach to standardization (that was subsequently applied to the names in the BR), and the development of an approach to link names probabilistically. This approach was tested and fine-tuned by linking names in the BR across time, and with ad-hoc projects that required the linkage of external data sources to Statistics Canada data sources via the BR. These methods can then be applied to adjacent cross-sectional files of the BR in order to identify false births and deaths as well as mergers or acquisitions.

31.   In the third year, a project was launched to geo-code entities on the BR by latitude and longitude so that a distance metric could also be used to match entities. Again, these geographic indicators become useful to detect organizational changes or false births and deaths.

4.3   Future work on the NALMF

32.   Many challenges remain in the development of the longitudinal structure of NALMF, and work has begun to address them. New tools have been developed to compare some key characteristics of businesses across time: the proportion of common employees across time, their names, and the latitude and longitude coordinates of their addresses. Other characteristics like industry or telephone numbers could also be added. All of these characteristics already reside on the BR or could be placed on the BR.

33.   The next major step is to choose the characteristics to be used to create the longitudinal structure, and the rules that will be used to combine the information available ((Oyarzun & Wile, 2016) propose one approach). Labour tracking has already revealed that the relationship between enterprises across time can be complex with large enterprises opening new locations, buying/selling operations from/to other enterprises, and shutting down operations. One approach would be to provide detailed information to researchers on the many changes taking place so that they can build a custom structure that suits their own particular purposes. However, the production of official statistics on firm entry and exit requires the development of a baseline longitudinal firm structure that responds to most needs. The placement of this type of information on the BR will be beneficial to the Agency as it will remove duplication of effort and establish consistency in the longitudinal business statistics produced across Statistics Canada.

4.4   Lessons learned

34.   The main lesson learned from the NALMF is that the creation of good quality longitudinal business data is a long term project that is best accomplished in a step-wise fashion. Such a large scale project requires the implication of people with different background and expertise, as well as the testing of different options.

35.   NALMF proved that, when used in conjunction with administrative data, the current BR allows the creation of consistent cross-sectional datasets and statistical measures over a long time period (back

to 2000).[28] The historical richness of the current BR needs to be recognized, and taken into consideration in the creation of the new longitudinal modules. Ideally, the new modules would be created moving forward, but also backward in time.

36.   The cross-sectional data in NALMF already cover more than fifteen years. Such a long period revealed the variety and complexity of the changes going on in the business population. It also exposed two key challenges in the creation of a longitudinal BR. The first one has to do with concepts and definitions. How should the concepts of entry, exit, and continuity be defined? Once there is consensus on this, how should the longitudinal business identifiers be created and kept track of? When business A buys B and C while getting rid of locations C and D, should it receive a new identifier, or should it keep living as A? The creation and maintenance of business identifiers over a long time horizon is one of the greatest puzzles moving forward.

5   Recent initiatives – more timely statistics and statistics on entrepreneurs

37.   To remain relevant in the digital world where large volumes of data are readily available, statistical agencies need to be able to produce timely, detailed and high quality statistical products that transform the raw data into information and knowledge. In response to the needs of its stakeholders, Statistics Canada has developed a number of products that extend the usefulness of the databases and statistics discussed above. These information products leverage the administrative sources already available at Statistics Canada to produce more timely information and statistics along dimensions not previously available. This section discusses two such initiatives.

5.1   Quarterly estimates of business entry and exit

38.   The LEAP is used to produce annual statistics on business and employment dynamics. The series produced using LEAP are updated in the summer for a reference year that is two years in the past. For example, data for 2015, was released in the summer of 2017. The key constraint preventing the publication of more timely statistics using the LEAP is that the annual T4 slips are available with a considerable lag. To overcome this constraint, the monthly PD7 payroll deduction and remittances files were used in conjunction with the T4 slips in the LEAP. Like the T4 slips, the PD7 is filed by businesses (identified with a BN) with the CRA. The T4 and PD7 were used together to create a new experimental quarterly series on business entry and exit and the associated employment creation and destruction (see (Leung, 2015) for more details); series that can be published four months after the end of the reference quarter. The statistics were created in such a way that they were consistent with the pre-existing LEAP annual data and the quarterly jobs numbers published in the Canadian Productivity Accounts at Statistics Canada.

39.   During the course of this project, several issues were flagged. The number of entrants that are unclassified in the quarterly data are even larger than in the annual LEAP because there is even less time for the BR to fully integrate them. An internal study was conducted to examine the industries to which entrants were eventually classified. It was decided to adopt a raking procedure (distribute the unclassified entrants across industries in proportion to the entrants with an industry code). While the internal study showed that historically, on average, this is a reasonable approach, it is expected that this approach is more prone to error during macroeconomic events that affect disproportionately certain sectors of the economy. The internal study also revealed that many entrants exited before they could be classified to an industry; the short lifespan of many entrants is consistent with the Canadian evidence on the survival of firms (Baldwin, Bian, Dupuy, & Gellatly, 2000)

40.   Hence, looking forward, the production of non-experimental statistics based on a longitudinal structure on the BR will likely require: improvements in the speed in which the BR can integrate new

---

[28] This is true despite the fact that a major BR redesign took place in 2008.

units and classify them (possibly with the use of tools such as the ones described in Section 4.2 of this example); or the development of techniques to better handle the spreading of the unclassified entrants across industries that takes into account observable signals in the economy.

41. Second, the use of the PD7 files in conjunction with the T4 slips sheds light on the filing patterns of these documents over the life of a business. Both data sources come from the CRA, and theoretically employers would be filing both and they would be consistent with one another. However, the T4s are issued annually after the close of a calendar year by businesses to individuals for personal income tax purposes, while the PD7 is filed on a more frequent basis by businesses but does not implicate individuals. The fact that the T4 involves both individuals and businesses explains why the T4 is a more complete file in general, but also at the start of a business: employees will pursue their employer because they need to file their personal income taxes on time. The filing rates for PD7 are higher than for T4 at the end of the life of businesses. For example, firms that exit early in a given year will file a last PD7 in that year but will be gone when it will be time to issue T4s to its old employees in the beginning of the following year. The lesson learned here is that there are strengths and weaknesses associated with different files that seem to cover the universe of businesses. The use of a broader set of data sources is essential to capture businesses on the margins of their lifecycles.

5.2 Canadian Employer-Employee Dynamics Database (CEEDD)

42. The increased use of record linkages expands greatly the statistics that can be produced. Although the BR is a register of businesses and hence well placed with some enhancements to produce statistics on business demographics and/or longitudinal statistics on businesses, the combination of administrative data on businesses with administrative data on persons allows the coming together of the work on entrepreneurship (a term inherently linked to an individual), and statistics on businesses.

43. The CEEDD is an example of one such linkage of longitudinal data on the individuals with businesses. It essentially links the firm performance data in NALMF with personal income tax files in order to create a matched employer-employee dataset. It also allows the identification of the self-employed and the business owner of private corporations (see (Green, Liu, Ostrovsky, & Picot, 2016)) for a complete description of the database). It is now possible to start developing statistics on businesses based on the characteristics of the business owner (female, immigrant, youth, indigenous, family-owned, serial business owner, etc.). Studies can be done, for example, not only on the contribution of entry cohorts of businesses to employment, but also on the contribution of a cohort of business owners. This will improve the links between entrepreneurs (the individuals) and their contribution to the business dynamics in the economy.

44. Thus, the BR could consider including not only flags associated with characteristics of the business, but also indicators associated with the characteristics of the business owner. This would allow the longitudinal study of businesses and entrepreneurs.

6   Concluding remarks

45. The BR has been used to support the development of longitudinal business databases at Statistics Canada for the past two decades. These experiences have shown areas in which improvements could be made, but also demonstrated the opportunities for the BR to support the production of a wide range of statistics related to longitudinal business data across the entire Agency.

46. In response to the need for more timely and detailed data on businesses, Statistics Canada is embarking on developing new central modules on the BR systems over the next few years. Building on the work described in the previous sections, these modules would allow historical updates, include longitudinal identifiers, and document predecessor/successor relationships. The new longitudinal features in the BR will provide the infrastructure for a more rapid response to evolving data needs and allow for a greater consistency in the statistics produced across the Agency.

47.   There are many new challenges for the BR in this environment - each of which needs to be evaluated and prioritized. Preserving the maximum degree of flexibility to group and regroup entities at a point in time and across time would allow the widest range of statistics and analytical studies to be produced. A shift in effort areas will be essential. In the past, the production of more highly aggregated statistics demanded that the largest business units in the economy be carefully profiled and kept up to date. More recently, the production of more timely, disaggregated statistics, and statistics on businesses that focus on the new entities, requires dealing with the much larger number of small units that have relatively less economic activity. The length of life of those businesses can be short and detailed information about them is not currently available on a timely basis.

Annex 2: Profiling entrepreneurs by linking business and social statistics in Denmark

48.   There are strong links between a new enterprise and the owner behind – the entrepreneur; especially as the vast majority of new enterprises start up without employees and often as sole proprietorships. This close linkage between the enterprise and the entrepreneur makes obvious the need to consider profiling the new enterprise not only from the business statistics perspective (e.g. activity, location, turnover, employment), but also to utilise social data and statistics to describe the entrepreneur (e.g. gender, age, education) in order to better understand the nature and performance of the new enterprise.

Figure A2.1 **The linking of enterprise and social statistical registers in the Danish statistical system**
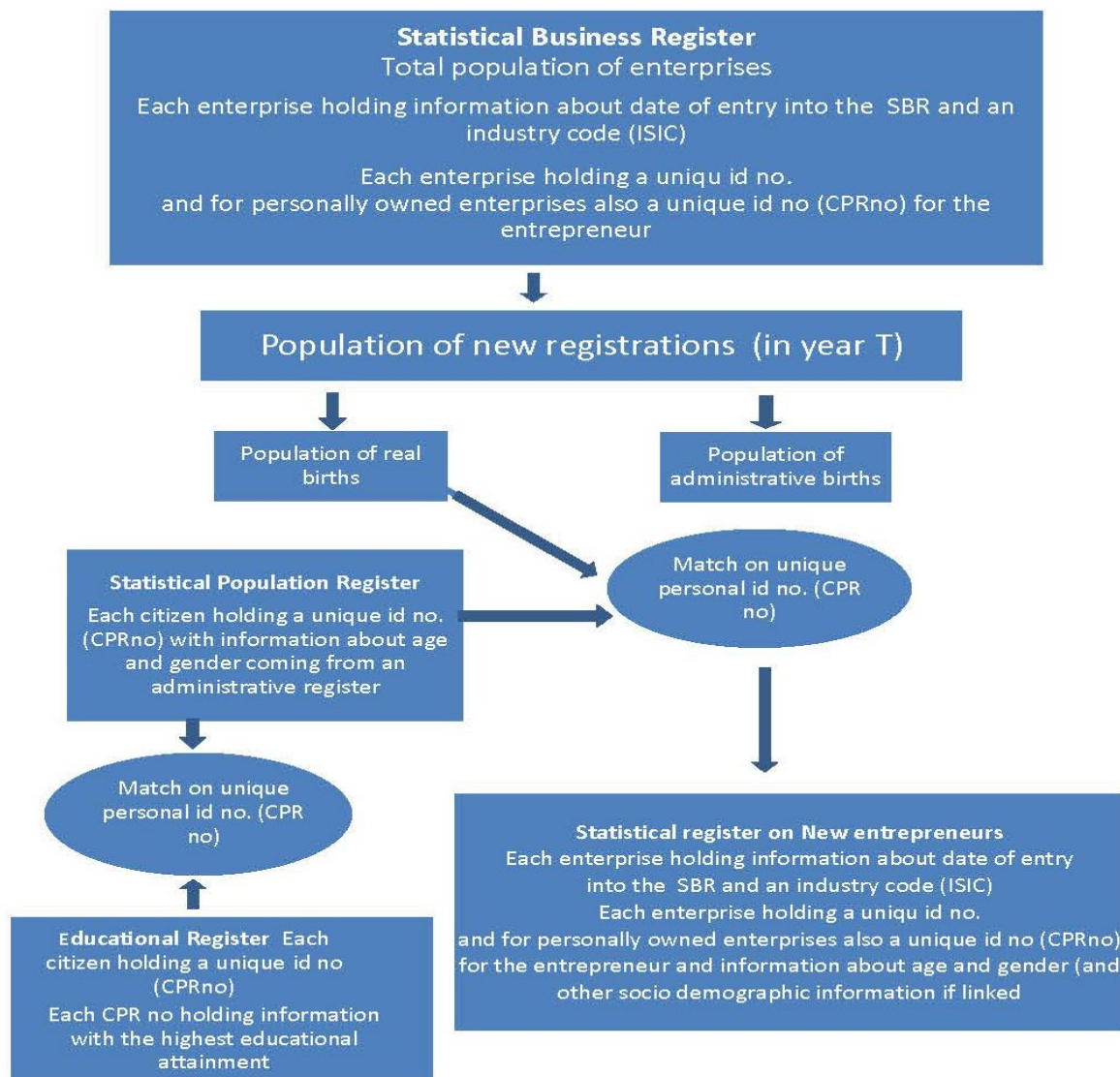
49.   In statistical business registers (SBR) some basic information about the owner of the personally owned enterprises are often stored at the individual enterprise level. This is also the case for the Danish SBR where a unique personal identification number (so-called CPR no) is linked to the enterprises via the administrative business register. This identification number automatically holds information about age and gender of the owner and can be linked to other statistical registers holding information about education, income, etc. at the personal level identified by the unique personal identifier, see Figure A2.1 above.

50.   Statistical information about the persons behind the start-ups is important for policy shaping as, for example, there is a focus on creating more women-owned start-ups, as women entrepreneurs are seen as a source for future growth and innovation in the Danish economy that has yet to be fully tapped. Furthermore, there is also a policy focus on encouraging younger persons to establish their own businesses; to better utilize the digital competences posed especially by these age groups.

51.   Although the Danish entrepreneurship policy has had the focus on increasing the share of female entrepreneurs, the share of female entrepreneurs establishing new enterprises has remained almost unchanged in Denmark since 2001 varying from 25 to 29 per cent of all new entrepreneurs; see Figure A2.2.

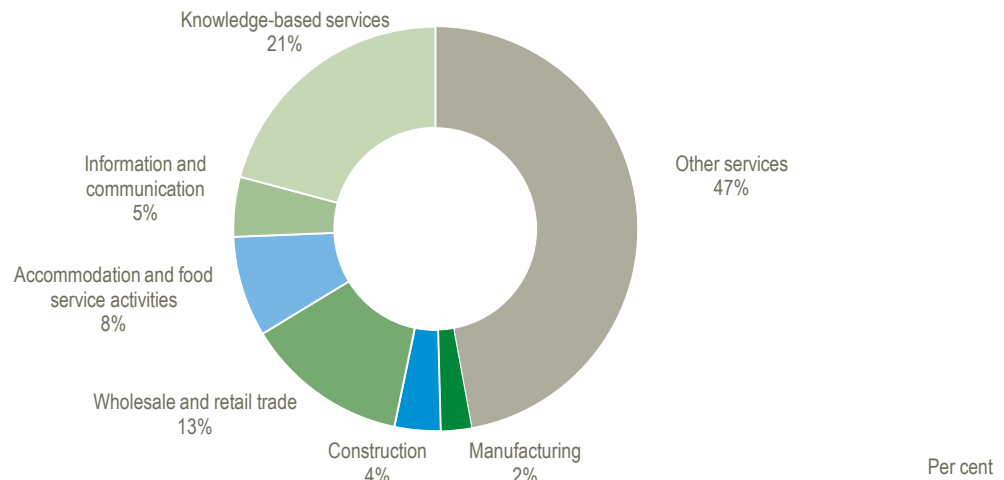Figure A2.2 **Entrepreneurs by gender (2001-2013)**
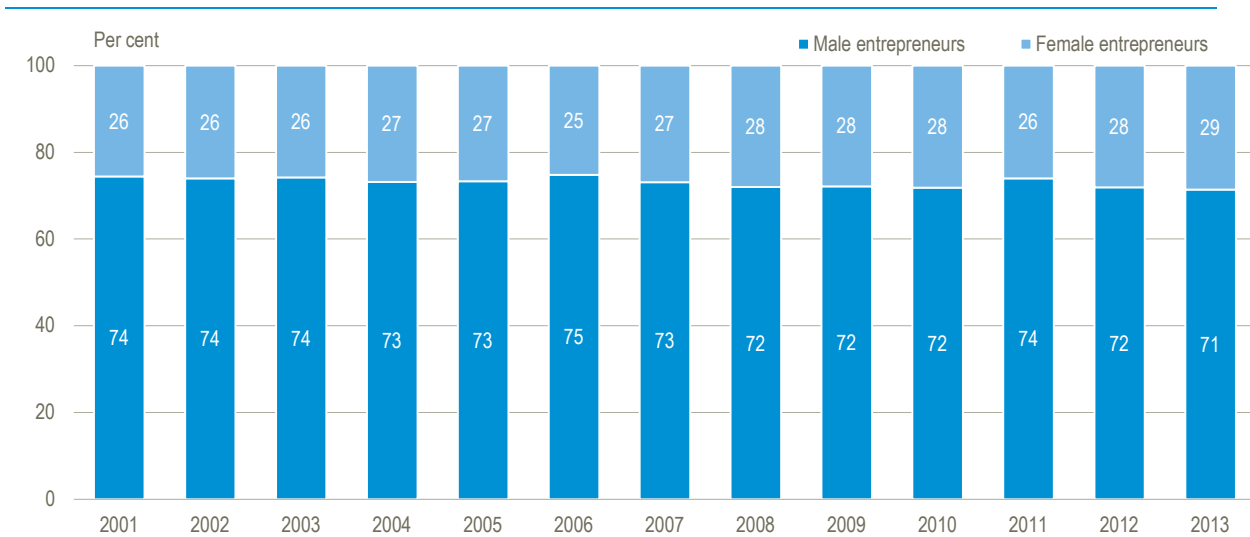
Source: Statistics Denmark

52.   The information about gender of the new entrepreneurs is not only interesting from a policy point of view but also of interest when analysing and understanding the activity breakdown of the new enterprises as females do not start up in the same activities as males, see Figures A2.3 and A2.4.
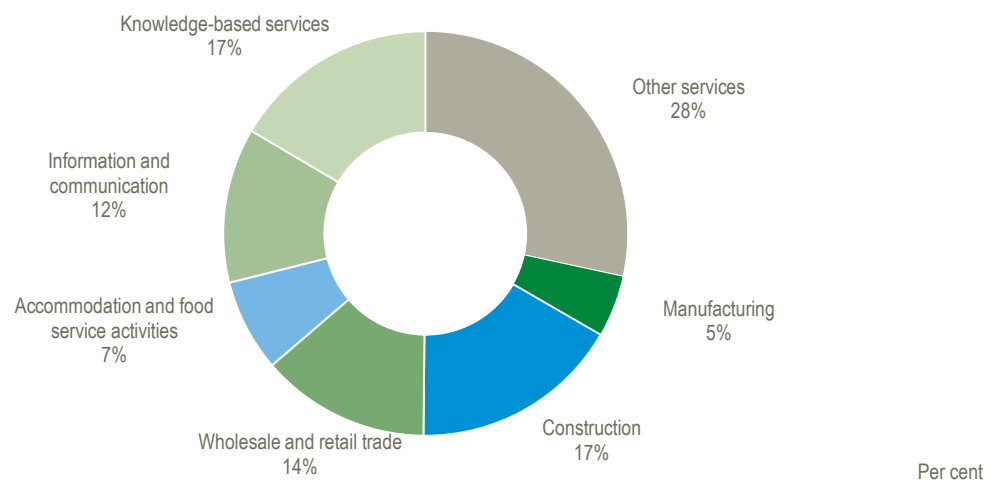
53.   A larger share of female entrepreneurs starts their businesses in services industries, e.g. cleaning, while male entrepreneurs to a much larger extent start up within construction or information and communication industries.

Figure A2.3 **Women entrepreneurs broken down by sector (2013)**

Knowledge-based services
21%

Information and
communication
5%

Accommodation and food
service activities
8%

Wholesale and retail trade
13%

Construction
4%

Manufacturing
2%

Other services
47%

Per cent

Source: Statistics Denmark

Figure A2.4 **Male entrepreneurs broken down by sector (2013)**

Knowledge-based services
17%

Information and
communication
12%

Accommodation and food
service activities
7%

Wholesale and retail trade
14%

Construction
17%

Manufacturing
5%

Other services
28%

Per cent

Source: Statistics Denmark

54. If we compare the age distribution among the new entrepreneurs with the total population, we see an overrepresentation of entrepreneurs in the age groups 25-54 years old, cf. Figure A2.5. This fits well with the educational patterns in Denmark as the age groups 25-39 years are representing the age groups where young people have finalised their education and having had their first work experience.

55. Obviously, not only gender and age determine the pattern for establishing a new enterprise, but also characteristics such as education, work experience, wealth and income are important factors when analysing and designing policy to establish the creation of better framework conditions for new entrepreneurs – including the design of measures to stimulate entrepreneurship. In Denmark, all these socio-demographic characteristics can be compiled using administrative registers as the personal identification number is used across the public sector administration in Denmark. This condition has allowed Statistics Denmark already from the nineties to develop a battery of indicators for measuring and profiling the new entrepreneurs which must be considered relatively unique, except for the other Nordic countries. The core list of indicators can be found in Box A2.1.

Figure A2.5 **New entrepreneurs and the entire population broken down by age (per cent)**
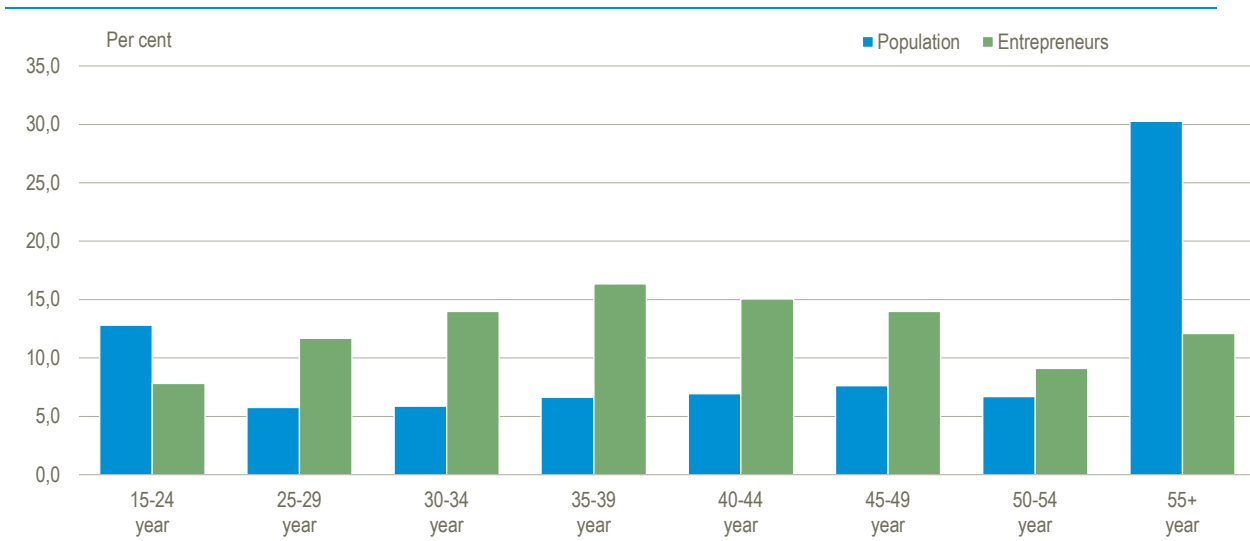
56.   To assess the contribution and importance of new enterprises for economic growth and employment, it is necessary to follow enterprises for a longer period than the year of start-up. Not all new activities are sustainable, and quite a large proportion of the new enterprises will fail within the first 3-5 years. The establishment of a new enterprise is not in itself a criterion of success; the decisive element is the contribution of the new enterprises to innovation and the creation of value added and employment. An analysis of the role played by new enterprises in the development of the economy calls for knowledge about the specific characteristics of the surviving enterprises and the entrepreneurs behind.

Box A2.1 **Core socio-economic indicators**

1) Age of entrepreneur
2) Nationality *(interesting from the aspect of integration of foreign citizens into the domestic economy; in Denmark we find relatively many immigrants finding their way to the labour market as entrepreneurs)*
3) Highest level of education *(interesting from the perspective of motivation (push or pull factor) for setting up an enterprise, e.g. a person having obtained a certain knowledge and know-how through his studies might be pulled to realise his innovative ideas by setting up a new enterprise to obtain a better income or simply by implementing the product idea)*
4) Relation to the labour market the year prior to start-up *(interesting from the perspective of motivation (push or pull factor) for setting up an enterprise, e.g. a person being unemployed might be pushed to set up a new enterprise to secure an income)*
5) Gross income or wealth prior to start-up *(as access to finance can be difficult so the possibility of financing the new enterprise by own capital)*

57.   This is the reason why Statistics Denmark has analysed not only the different characteristics of the surviving new enterprises but also of the entrepreneurs behind to identify the importance of different factors. In general, the activity of the enterprise is the most important factor determining survival of the enterprise but also some of the characteristics of the entrepreneur can via the regression model used by Statistics Denmark be identified as having effects on the possibility of survival. The model

identifies the factors outlined in Box A2.2 as the most important factors related to the survival of the new enterprises.

Box A2.2 **Indicators related to the survival of enterprises**

> *1)* Industry (ISIC code) *(the industry is the most statistically significant factor for survival of the new enterprise)*
> 2) Previous branch experience *(informing about the experience gained from previous job to better understand factors of survival; in the case of Denmark, statistics has shown the importance of previous branch experience for the survival of the new enterprises)*
> 3) Age of the entrepreneur
> *4)* Highest level of education *(informing about the formal competences of the entrepreneur being of importance for survival)*

58.   The longitudinal analysis of the 19 600 new enterprises that started in 2011 show that after three years 10 700 or 55% had survived, see Figure A2.6. There is a small difference in survival rates between women and male entrepreneurs, with 56% of the new enterprises started by males surviving, compared to 52% of those started by women. This pattern can be found in all industry groupings – except perhaps surprisingly manufacturing where only few females started business. The survival rate of male entrepreneurs is especially in Information and communication (9 percentages points) and construction (8 percentages points) higher than for female entrepreneurs. These are also the two industry groupings where we find the relative largest share of male entrepreneurs. The results of the model have shown that previous work experience from an enterprise in the same economic activity as the entrepreneur starts up is the crucial factor for raising the possibility of survival. New entrepreneurs starting up in the same activity in which they were employed the year prior to start-up have a survival rate of 69%, compared with 51% for entrepreneurs with no previous experience from that particular activity.

Figure A2.6 **Survival rates of new enterprises broken down by gender and activity (2011 – 2014)**
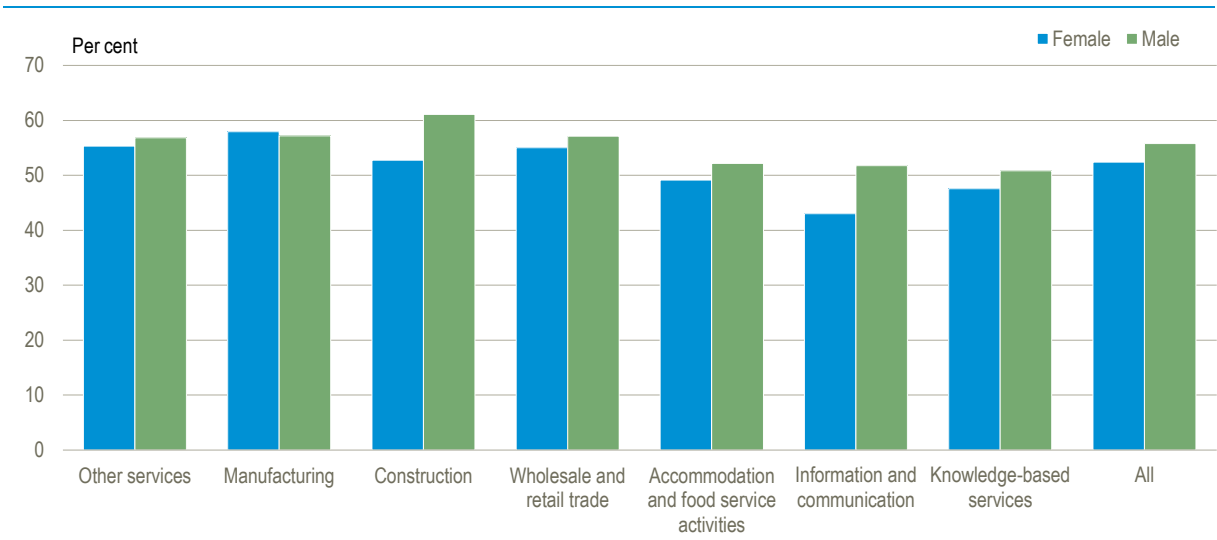
Source: Statistics Denmark

Annex 3: Family businesses in the Netherlands

With the availability of suitable administrative registers and identifiers it may be possible to identify family businesses by linking administrative information to the statistical units in the SBR. Based on best practices experiences in Statistics Netherlands, Table A3.1 describes how a satellite register of family businesses (FBs) in the SBR may be developed.[29]

Table A3.1 **Characterising and recording family businesses - Statistics Netherlands approach**

Nr	Action	Notes
1.	The starting point is to analyze what role a FB has in the enterprise population	FBs are considered as playing an important role in the economy of a country, because of their role in supporting innovation and entrepreneurship and their commitment to the region they are operating in. The second important aspect an FB is control of business knowledge, skills, property, assets and liabilities within the kin of the family and keeping and protecting it for later generations.
2.	Based on the results of the first step, a definition of an FB should be developed and the corresponding "statistical universe".	Since statistical units may consist of more than one legal unit it is feasible to involve the characteristics of the linked legal units to characterize an FB. An FB may be characterized as when • at least one representative of the family or kin is formally involved in the governance of the enterprise (group); • a majority of decision-making rights is in the possession of the natural person(s) who established the enterprise (group), or in the possession of the natural person(s) who has (have) acquired the share capital of the enterprise(group), or in the possession of their spouses, parents, children or children's direct heirs; and • the assets of the enterprise (group) are transferable to a successor.
3.	Parallel to this stage, investigations on the administrative concepts should lead to appropriate laws and useful administrative sources to characterize a FB in the "legal universe". Furthermore, the information available might help defining the "statistical universe".	Information on legal form and the natural persons linked as an owner, a partner or a shareholder are often registered in the trade register, by the tax authorities, or as a financial relationship known by the central bank. Family relationships are often registered in the national registration of inhabitants. It might be necessary that a special file with the appropriate family-relationships is derived from it. The shareholders of a limited liability company are the owners of the business. To distinguish a group which keeps the majority of the shares, information about the number of shares held is desirable. At a minimum, information about the reception (or not) of social security insurance for employees is needed. The underlying idea is that board members without social security insurance for employees are likely to own shares. In this step the administrative concepts on control, family and transferability of wealth to the next generations are subject to

---

[29] More information was presented at the meeting of the Group of experts on Business Registers in September 2017, see: http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.42/2017/Netherlands.pdf

		investigation in administrative sources. To obtain a complete population of 'owners' it is important to involve the information of all natural persons who are an 'owner' of the legal units linked by a control relationship of more than 50% (i.e. enterprise group).
		Control by one family: The concept of control of a legal unit depends on the legal form:
		- Sole proprietor: owner is the family - Limited liability: most of the shares or voting power in one family - Partnership: most of partners in one family
		Family: A family business always involves two generations. Two persons are considered in "family" if they have a common parent or great-parent. The word parent in this case also includes the parents of the registered partner, husband or wife. Parents of former registered partners, husbands and wives, persons who used to live together like husband and wife but never registered their relationship (assuming they started the business when they lived together) and persons that have a common child but were never married or never had their relationship registered, are also included.
		Transferability: A legal unit is considered 'transferable' if it is possible to leave the knowledge and/or the assets to a successor.
4.	The next challenge is to develop a "metadata translator", that translates administrative concepts into statistical concepts.	The metadata translator describes how the information from the different administrative sources are linked to the statistical units in the SBR. In the SBR, each statistical unit is linked to one or more legal units. Each legal unit is uniquely identified by the ID of the National Statistical Agency (NSA-ID) and each NSA-ID is linked to a national ID (NAT-ID) of the administrative source. Aggregation of the linked administrative data to the level of the statistical unit in the SBR, should meet the agreed requirements to be a FB according the definition in 2.
5.	In the last step the "legal universe" and the "statistical universe" are linked by the "Micro Data Translator" to detect a FB in the SBR.	Selections and aggregations in the micro data translator: Make a selection in the appropriate administrative source (see 2.) of all sole proprietors, shareholders or partners, which are registered as the owner of a legal unit. Link the information on 'control' for each legal unit to the enterprise(group) in the SBR. Investigate the family relations in the enterprise(group). Calculate the indicator '% of Family Relations': *#Family relations pairwise / #Relations pairwise.* If the indicator is above 50% the control is linked to the same family and the enterprise (group) meets the definition of a FB. Sole proprietors can be selected directly from the SBR. The knowledge or assets of a sole proprietor is linked directly to the expertise of the owner. Transferability of this knowledge to a successor is possible, if at least one person, beside the owner is working in the 'self-employed' enterprise(group).

References

Ahmad, & Hoffmann. (2008). *A Framework for Addressing and Measuring Entrepreneurship.* Retrieved from http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=std/doc(2008)2

Ahmad, & Seymour. (2008). *Defining Entrepreneurial Activity: Definitions Supporting Frameworks for Data Collection.* Retrieved from http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=std/doc(2008)1

Alajääskö. (2015). *Statistics on trade by enterprise characteristics – brief review of past, current and future projects linking business statistics in the European Statistical System.* Retrieved from https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.42/2015/Session_II_Eurostat_-_Statistics_on_trade_by_enterprise_characteristics.pdf

Baldwin. (1995). *The Dynamics of Industrial Competition.* Retrieved from https://doi.org/10.1017/CBO9780511664700

Baldwin, & Gu. (2004). *Innovation, Survival and Performance of Canadian Manufacturing Plants.* Retrieved from http://publications.gc.ca/Collection/Statcan/11F0027M/11F0027MIE2004022.pdf

Baldwin, & Gu. (2006). *Competition, Firm Turnover and Productivity Growth.* Retrieved from http://www.publications.gc.ca/Collection/Statcan/11F0027M/11F0027MIE2006042.pdf

Baldwin, Bian, Dupuy, & Gellatly. (2000). *Failure Rates for New Canadian Firms: New Perspectives on Entry and Exit.* Retrieved from http://publications.gc.ca/Collection/Statcan/61-526-XIE/61-526-XIE1999001.pdf

Baldwin, Dar-Brodeur, & Yan. (2016). *Innovation and Export-market Participation in Canadian Manufacturing.* Retrieved from https://www150.statcan.gc.ca/n1/pub/11f0019m/11f0019m2016386-eng.htm

Baldwin, Dupuy, & Penner. (1992). *Development of longitudinal panel data from business registers: the Canadian Experience.* Retrieved from https://content.iospress.com/articles/statistical-journal-of-the-united-nations-economic-commission-for-europe/sju9-4-01

Baldwin, Leung, & Rispoli. (2013). *Canadian Labour Productivity - Differences Across Firm Size Classes, 2002 to 2008.* Retrieved from http://www.statcan.gc.ca/pub/15-206-x/15-206-x2013032-eng.pdf

Brown, & Rispoli. (2014). *Metropolitan Gross Domestic Product: Experimental Estimates, 2001 to 2009.* Retrieved from https://www150.statcan.gc.ca/n1/pub/11-626-x/11-626-x2014042-eng.htm

Cella, & De Gregorio. (2015). *Measuring the entrepreneur: a comparative analysis using business and household data.* Retrieved from https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.42/2015/Session_I_Italy__-_Measuring_the_Entrepreneur.pdf

De Waal. (2015). *General Approaches for Consistent Estimation based on Administrative Data and Surveys.* Retrieved from https://www.cbs.nl/nl-nl/achtergrond/2015/37/general-approaches-for-consistent-estimation-based-on-administrative-data-and-surveys

De Waal, Coutinho, & Shlomo. (2017). *Calibrated Hot Deck Imputation for Numerical Data Under Edit Restrictions.* Retrieved from https://doi.org/10.1093/jssam/smw037

Dixon, & Rollin. (2014). *The Distribution of Employment Growth Rates in Canada: The Role of High-growth and Rapidly Shrinking Firms.* Retrieved from https://www150.statcan.gc.ca/n1/pub/11f0027m/11f0027m2014091-eng.htm

Eurostat. (2007). *Recommendations Manual on the Production of Foreign AffiliaTes Statistics (FATS).* Retrieved from http://ec.europa.eu/eurostat/ramon/statmanuals/files/FATS_2007_EN.pdf

Eurostat. (2008). *NACE Rev. 2 - Statistical classification of economic activities in the European Community.* Retrieved from http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

Eurostat. (2010). *Business registers – Recommendations manual.* Retrieved from http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-32-10-216-EN-C-EN.pdf

Eurostat. (2012). *Foreign AffiliaTes Statistics (FATS) Recommendations Manual.* Retrieved from http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-12-016

Eurostat. (2013). *European system of accounts – ESA 2010.* Retrieved from http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-02-13-269

Eurostat. (2016). *Compilers guide on European statistics on international trade in goods by enterprise characteristics (TEC).* Retrieved from http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-16-002-EN-N

Eurostat, & OECD. (2007). *Eurostat-OECD Manual on Business Demography Statistics.* Retrieved from http://www.oecd.org/sdd/business-stats/39974460.pdf

Eurostat, & OECD. (2017). *Compilers Guide for statistics on Services Trade by Enterprise Characteristics (STEC) - 2017 edition.* Retrieved from http://ec.europa.eu/eurostat/documents/3859598/8466641/KS-GQ-16-007-EN-N.pdf

Fischer, & Reuber. (2010). *The State of Entrepreneurship in Canada.* Retrieved from https://www.ic.gc.ca/eic/site/061.nsf/eng/h_rd02468.html

Gersick, Hampton, Lansberg, & Davis. (1997). *Generation to generation.* Retrieved from https://hbr.org/product/generation-to-generation-life-cycles-of-the-family-business/555X-HBK-ENG

Geurts. (2016). *Longitudinal firm-level data: problems and solutions.* Retrieved from https://doi.org/10.1007/s11187-015-9693-6

Green, Liu, Ostrovsky, & Picot. (2016). *Business Ownership and Employment in Immigrant-owned Firms in Canada.* Retrieved from https://www150.statcan.gc.ca/n1/pub/11-626-x/11-626-x2016057-eng.htm

Gu, & Li. (2017). *Multinationals and Reallocation: Productivity Growth in Canadian Manufacturing Sector.* Retrieved from https://www150.statcan.gc.ca/n1/pub/11f0019m/11f0019m2017398-eng.htm

Hidiroglou, & You. (2016). *Comparison of unit level and area level small area estimators.* Retrieved from https://www150.statcan.gc.ca/n1/pub/12-001-x/2016001/article/14540-eng.htm

Houbiers. (2004). *Towards a Social Statistical Database and Unified Estimates at Statistics Netherlands.* Retrieved from https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/towards-a-social-statistical-database-and-unified-estimates-at-statistics-netherlands.pdf

ILO. (2018). *Resolution concerning statistics on work relationships - 20th International Conference of Labour Statisticians.* Retrieved from https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms_648693.pdf

Kim, Reiter, Wang, Cox, & Karr. (2014). *Multiple Imputation of Missing or Faulty Values Under Linear Constraints.* Retrieved from https://doi.org/10.1080/07350015.2014.885435

Leung. (2015). *Quarterly Business and Employment Dynamics: Experimental Estimates, First Quarter 2001 to Third Quarter 2014.* Retrieved from https://www150.statcan.gc.ca/n1/pub/11-626-x/11-626-x2015045-eng.htm

Leung, Rispoli, & Gibson. (2011). *Small, Medium-sized and Large Businesses in the Canadian Economy: Measuring Their Contribution to Gross Domestic Product in 2005.* Retrieved from http://publications.gc.ca/collections/collection_2011/statcan/11F0027M/11f0027m2011069-eng.pdf

OECD. (2008). *Defining Entrepreneurial Activity: Definitions Supporting Frameworks for Data Collection.* Retrieved from http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=std/doc(2008)1

OECD. (2017). *Entrepreneurship at a Glance 2017.* Retrieved from http://dx.doi.org/10.1787/entrepreneur_aag-2017-en

Oyarzun, & Wile. (2016). *An Overview of Business Record Linkage at Statistics Canada: How to link the 'Unlinkable'.* Retrieved from https://www.statcan.gc.ca/eng/conferences/symposium2016/program/14741-eng.pdf

Pannekoek, Shlomo, & De Waal. (2013). *Calibrated impution of numerical data under linear edit restrictions.* Retrieved from https://projecteuclid.org/euclid.aoas/1387823307

Rollin. (2013). *Developing a Longitudinal Structure for the National Accounts Longitudinal Microdata File.*

Schumpeter. (1912). The Theory of Economic Development.

Shlomo, De Waal, & Pannekoek. (2009). *Mass Imputation for Building a Numerical Statistical Database.* Retrieved from https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2009/wp.31.e.pdf

Statistics Canada. (1988). *Developing a Longitudinal Database on Businesses in the Canadian Economy: An Approach to the Study of Employment.* Retrieved from http://publications.gc.ca/site/eng/9.835281/publication.html

Sterman. (2000). *Business Dynamics: Systems thinking and modelling for a complex world.*

UNECE. (2000). *Terminology on Statistical Metadata.* Retrieved from https://www.unece.org/fileadmin/DAM/stats/publications/53metadaterminology.pdf

UNECE. (2011). *Using Administrative and Secondary Sources for Official Statistics - A Handbook of Principles and Practices.* Retrieved from http://www.unece.org/fileadmin/DAM/stats/publications/Using_Administrative_Sources_Final_for_web.pdf

UNECE. (2014). *In-depth review of entrepreneurship statistics.* Retrieved from https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2014/8-In-depth_review_of_entrepreneur_statistics.pdf

UNECE. (2015). *Guidelines on Statistical Business Registers.* Retrieved from https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

United Nations. (2008). *International Standard Industrial Classification of All Economic Activities. Revision 4.* Retrieved from https://unstats.un.org/unsd/publications/catalogue?selectID=396

United Nations. (2009). *System of National Accounts 2008.* Retrieved from
https://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf

United Nations. (2015). *Central Product Classification (CPC) Version. 2.1.* Retrieved from
https://unstats.un.org/unsd/classifications/unsdclassifications/cpcv21.pdf

Vale. (2006). *The international comparability of business start-up rates final report.* Retrieved from
http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=
std/doc(2006)4

Whitridge, & Kovar. (1990). *Use of Mass Imputation to Estimate for Subsample Variables.*

Whitridge, Bureau, & Kovar. (1990). *Mass Imputation at Statistics Canada.*

Guidelines on the use of statistical business registers for business demography and entrepreneurship statistics

This publication provides practical guidance on developing and maintaining statistical business registers to support the production of business demography and entrepreneurship statistics. The guidelines provide definitions and compilation guidance for demographic events such as births and deaths of enterprises, describe the requirements for the statistical business register to produce business demography statistics and propose business demography indicators that can be produced directly from the statistical business. The Guidelines also give guidance on linking the statistical business register with information from administrative registers and other data sources to produce business and entrepreneurship statistics.