

**UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE**

ECE/CES/BUR/2015/FEB/17/Add.1
20 January 2015

CONFERENCE OF EUROPEAN STATISTICIANS

For decision

Meeting of the 2014/2015 Bureau
Geneva (Switzerland), 17-18 February 2015

Item 5 (e) of the Provisional
Agenda

International Guidelines on Statistical Business Registers

Consolidated Draft Version

Prepared by the Task Force

*Attached are the International Guidelines on Statistical Business Registers (document ECE/CES/BUR/2015/FEB/17/Add.1), prepared by the CES Task Force. **The Bureau is invited to comment on the final draft of the Guidelines and to decide whether it can be circulated to all CES members for electronic consultation with a view of presenting it to the CES 2015 plenary session for endorsement.***

Preface

The statistical business register (SBR) plays a central role in the production of economic statistics, both in terms of the way the statistics are produced and in terms of their content and quality. Traditionally the function of the SBR has been to provide a population of statistical units from which frames and samples for economic surveys can be drawn, also to provide the tools for monitoring survey samples and response burden. Today, however, the SBR fulfils two other important roles. First, it is crucial in the integration and use of data from administrative and other sources. Second, a well-developed SBR with comprehensive list of enterprises and other statistical units, and information about their characteristics, can be used as a source of economic statistics in its own right. Business demographics can be derived directly. New statistics can be produced by combining information from the SBR with information from other statistical registers and administrative sources. Thus, the SBR is the backbone in the production of economic statistics, in ensuring coherence across the various statistical outputs, and in developing efficient statistical production processes.

The growing demand for better and more detailed economic statistics together with the need to make production more efficient has put the spotlight on the SBR and its role in the production process and in improving existing statistics and developing new statistics. On-going economic globalization also poses challenges for SBR production and increases the demand for internationally comparable SBRs.

In the light of the growing demand for higher quality and more internationally comparable economic statistics, the Bureau of the Conference of European Statisticians recognised a great diversity in the content and use of SBRs across countries and regions. Therefore, in November 2011, the Bureau established a Task Force to produce a set of international guidelines and recommendations to assist countries in developing and maintaining their SBRs.

The Guidelines, which are the result of the Task Force's work, provide descriptions and recommendations of good practices covering the key areas of the SBR and its various roles in the production of economic statistics. It is the hope that countries will find the Guidelines useful for their work on SBRs.

Background and Acknowledgements

Over the last decade the growing demand for better and more detailed business statistics has put focus on statistical business registers (SBRs) and their role in the production of business statistics. While traditionally the primary role of SBRs has been to provide sample frames for business statistics, SBRs of today often include more information and can be used in their own right as a source from which to derive business statistics. SBRs have also proved to be central for statistical offices' effort to reduce response burden and utilise administrative data sources and for the combination of survey data and administrative data.

There is also growing interest in the role of SBRs as the backbone in the production of business statistics and the potential benefits of integrating the SBR with other statistics. These benefits include both improvement of the quality of existing statistics and development of new statistics by e.g. combining SBRs with information from other administrative or statistical registers. Integration of SBRs also means an opportunity to streamline or "industrialise" the statistical production process. In its vision for the production of official statistics in future, the High-Level Group for Modernising Statistical Production and Services, established by the Bureau of the Conference of European Statisticians, has raised a number challenges for the development of the statistical products and the production process. SBRs will be important in this development.

The need for international guidance and recommendations of good practices on SBRs has been raised on different occasions by countries participating in the Wiesbaden Group and in the joint UNECE/Eurostat/OECD Expert Group on Business Registers. The proposal for international guidelines was raised at the Wiesbaden Group meeting in 2010 in Tallinn. The proposal was subsequently supported by the Steering Group of the Wiesbaden Group and by the joint meeting of the Group of Experts on Business Registers in September 2011. There was agreement that the guidelines should cover the various roles of SBRs, including the provision of services for business surveys, the use of administrative sources and the role of business registers in the production of business statistics.

At its meeting in November 2011 the Bureau of the Conference of European Statisticians recognized a great diversity in the content and usability of SBRs among countries and regions, while the demand for high quality business and internationally comparable business statistics is growing. Given this background the Bureau concluded that a set of international guidelines would be useful, and agreed to establish a Task Force on SBRs with the objective of producing a set of international guidelines and recommendations of good practices to help countries in developing and maintaining their SBRs. The guidelines should be targeted at both developed and less developed statistical systems and provide practical guidance and recommendations on the establishment and maintenance of an SBR.

According to its terms of reference, the Task Force should:

- 1) Provide practical guidance on core issues of establishing and maintaining the SBR.
- 2) Clarify typology, concepts and definitions, including for statistical units.
- 3) Provide guidance on the use of administrative and other sources for the establishment and updating of the SBR.
- 4) Provide guidance on how to use the SBR in its own right for production of statistics and how information from the SBR can be combined with information from other statistical registers, administrative sources or surveys to produce new statistics.
- 5) Provide guidance on the role of SBRs in the modernisation of statistical production and services.

The Guidelines should include country experiences and case studies when found useful to illustrate good practices. Links to other statistical subject areas, such as national accounts and trade statistics should be made when relevant. The Guidelines should take into account other international work in the area of statistical business registers, in particular the *Business Registers Recommendations Manual* of Eurostat (2010) and the *Guidelines for Building Statistical Business Registers in Africa* (African Development Bank, 2012) and consistency with these should be ensured.

During its period of work from 2012 to 2014 the Task Force had three face-to-face meetings in conjunction with international meetings on business registers, and a number of audio conferences throughout the period of work. Drafts of chapters and other relevant materials were shared on a common designated website created for the purpose.

The Task Force had its first audio conference in January 2012. At this conference the Task Force agreed on a time and work plan and on a first outline of the Guidelines. During the first part of 2012 the Task Force developed a detailed outline of the Guidelines, which was presented to the meeting of the Wiesbaden Group on Business Registers in September 2012 in Washington. The Task Force received useful comments and proposals from the participants concerning the content and structure of the Guidelines.

Based on the feedback received the first version of the chapters of the Guidelines were drafted and subsequently presented and discussed at the joint UNECE/Eurostat/OECD Expert Group meeting on Business Registers in September 2013 in Geneva. The participants provided a number of useful comments and suggestions. Countries with less developed statistical systems participated in both the Wiesbaden Group meeting in 2012 and the Expert Group meeting in 2013, which helped to ensure that issues of particular importance to these countries were also raised. Furthermore, UNSD in 2012/2013 conducted a global survey on SBRs the results of which were presented at the Expert Group meeting in September 2013. The survey provided useful information on particular challenges concerning data sources, coverage of business registers and definition of units and classifications, which the Task Force agreed to incorporate in the further drafting.

The Task Force reviewed all draft chapters, the comments received from countries and the outcome of the UNSD global survey on business registers at a meeting on 5 September 2013. At this meeting it was concluded that an editor would be needed to ensure coherence and consistency across the chapters. As a response to this need Statistics Canada agreed to fund an editor, Michael Colledge, who carried out the first round of editing in February – April 2014.

The Task Force discussed the second draft version of the chapters and a number of cross-cutting issues at a meeting on 28 April 2014 in Luxembourg, back-to-back with the Eurostat Working Group on business registers. Following this meeting the authors continued to review and further elaborate the draft chapters and submitted updated drafts in August 2014. After a second round of editing, the third draft version of the chapters were presented to the Wiesbaden Group meeting on 15-18 September 2014 in Vienna to solicit concrete comments and proposals on the draft chapters. The Task Force conducted a survey to all participants of the Wiesbaden Group meeting requesting comments and proposals for all draft chapters, which resulted in written comments from eight countries.

In the course of preparing the Guidelines the Task Force identified a number of questions related to the definitions and delineations of enterprises and institutional units in 2008 SNA. To ensure that the report of the Task Force would be in line with these standards the Task Force submitted a number of questions for clarification to the meeting of the Advisory Expert

Group on National Accounts in September 2014. The Task Force received response from the ISWGNA which was taken into account in the relevant chapters.

From September to November 2014 authors worked to incorporate comments received from the Wiesbaden Group and suggestions for improvements and clarifications by the editor and the chair of the Task Force. Additional country examples of good practices (e.g. from Costa Rica, Georgia and Malaysia) were received to be included in annexes. A final round of reviewing and editing of all chapters and annexes was carried out from November 2014 to January 2015 to ensure coherence and similarity in style across chapters.

[After agreement of the CES Bureau the draft Guidelines were circulated for consultation to all members of the CES in March-April 2015 and a final version incorporating comments and proposal received was submitted to the CES plenary session in June 2015 for endorsement.]

Acknowledgements

The Guidelines are the results of the collective efforts of the members of the Task Force. The Task Force was chaired by Norbert Rainer (Austria) and included the following members: Jenny Foster (Australia), René Beaudoin (Canada), Miguel Angel Cardenas Contreras, Margarita Ramírez Ardila, Sandra Liliana Moreno Mayorga and Martha Poveda Gomez (Colombia), Vibeke Skov Møller and Peter Ottosen (Denmark), Iman Ahmad Ahmad Elhitta (Egypt), Sylvie Mabile, Pierrette Schuhl and Pierre Teillet (France), Monica Consalvi and Caterina Viviano (Italy), Susana Perez (Mexico), Rico Konen and Anita Vaasen-Otten (Netherlands), Marietha Gouws (South Africa), Fabio Tomasini and Claude Macchi (Switzerland), Anton Tovchenko (Ukraine), Andrew Allen (United Kingdom), Amerigo Liotti and Susanne Maus (Eurostat), Mariarosa Lunati (OECD), Besa Muwele (The African Development Bank), Ronald Jansen and Luis Gonzalez Morales (UNSD) and Evan Brand and Carsten Boldsen (UNECE). UNECE provided the secretariat of the Task Force.

The chapters of the Guidelines were drafted by lead authors and with contributing authors providing comments and proposals. The lead authors of the Guidelines were: Norbert Rainer (Austria), Jenny Foster (Australia), René Beaudoin (Canada), Sylvie Mabile, Pierrette Schuhl and Pierre Teillet (France), Monica Consalvi and Caterina Viviano (Italy), Rico Konen (Netherlands), Fabio Tomasini and Paul-André Salamin (Switzerland), Andrew Allen (United Kingdom) and Amerigo Liotti and Susanne Maus (Eurostat).

The annexes of the Guidelines include a number of useful examples and country case studies. These were provided by René Beaudoin (Canada), Martha Poveda Gomez and Sandra Rodriguez Lopez (Colombia), Odilia Bravo Cambronero (Costa Rica), Vibeke Skov Møller (Denmark), Pierrette Schuhl (France), Gogita Todraze (Georgia), Monica Consalvi and Caterina Viviano (Italy), Hartini Yaacob (Malaysia), Rico Konen (Netherlands) and Fabio Tomasini and Paul-André Salamin (Switzerland).

Special thanks go to Michael Colledge for his commitment and substantial work on reviewing and editing the Guidelines to ensure coherence and clarity of language and presentation style across chapters, and to Statistics Canada who generously financed the work of the editor.

Table of Contents

1	Introduction	13
1.1	Context	13
1.2	What is an SBR?.....	14
1.3	Use of an SBR	16
1.4	Aims of SBR Guidelines	17
1.5	Overview of the Guidelines	18
1.6	Terminology.....	21
2	Roles of the SBR.....	24
2.1	Introduction.....	24
2.2	SBR Live Register	26
2.3	SBR Register Snapshot and Frozen Frame.....	27
2.4	SBR Survey Frame	30
2.5	SBR Survey Support	32
2.6	SBR Based Statistics.....	34
2.7	SBR Information Source	37
2.8	SBR International Data Exchange	39
2.9	SBR in Modernisation of Statistical Production and Services.....	40
2.10	Satellite Approach to Extend Functionality.....	41
3	Coverage of SBR.....	43
3.1	Introduction.....	43
3.2	Key Concepts	44
3.2.1	Production Boundaries.....	44
3.2.2	Market and Non-Market Producers	44
3.2.3	Non-Observed and Observed Economies.....	45
3.2.4	Informal Sector	45
3.2.5	Illegal Activities	47
3.2.6	Underground Activities.....	48
3.3	SBR Coverage	48
3.3.1	Type of Institutional Units.....	48
3.3.2	Government Units	48
3.3.3	Corporations	49
3.3.4	Non-Profit Institutions	49
3.3.5	Households	49
3.3.6	Self-employed Professionals	51
3.3.7	Own-use Production	51
3.4	Special Cases.....	51
3.4.1	Agricultural Holdings	51
3.4.2	Illegal and Underground Production	52
3.4.3	Extraterritorial Organisations and Bodies	52
3.5	Summary of Coverage Recommendations	53
4	Units in the SBR.....	56
4.1	Introduction.....	56
4.1.1	Content and References	56
4.1.2	Brief Presentation of Units and Problems Covered.....	56
4.2	Types of Statistical Units.....	58
4.3	Enterprise Group (and Truncated Enterprise Group).....	61
4.3.1	Introductory Remarks	61
4.3.2	Definition of Enterprise Group (and Truncated Enterprise Group)	62

4.3.3	Link between National and Global Views: Truncated Enterprise Group	63
4.3.4	Identifying Enterprise Groups in Practice	64
4.4	Enterprise (and Truncated Enterprise)	65
4.4.1	Introductory Remarks	65
4.4.2	SNA2008 and ISIC Rev 4 Definitions of Enterprise	65
4.4.3	Current European Definition of Enterprise	66
4.4.4	Towards a New Operational Definition of Enterprise	67
4.4.5	National and Global Views: the Truncated Enterprise	68
4.4.6	Simplified Correspondence between Legal units and Enterprises	68
4.5	Establishment	69
4.5.1	Introductory Remarks	69
4.5.2	ISIC Rev. 4 Definition of Establishment	69
4.5.3	Implementation Notes	70
4.6	Other Statistical Units	71
4.6.1	Introductory Remarks	71
4.6.2	Kind-of-Activity Unit	72
4.6.3	Local Unit.....	73
4.7	Legal Unit - Pivotal between Administrative Units and Statistics	74
4.7.1	Introductory Remarks	74
4.7.2	Definition of Legal Unit.....	74
4.7.3	Active Legal Units.....	75
4.7.4	Legal Form	75
4.7.5	Market and Non-Market Legal Units.....	76
4.7.6	Quasi-corporations:	76
4.8	Administrative Units	77
4.8.1	Introductory Remarks	77
4.8.2	Linkage to Statistical Units	77
4.9	Local Unit of Legal Unit	77
4.9.1	Introductory Remarks	77
4.9.2	Definition of Local Unit of Legal Unit	78
4.9.3	Local Unit of Legal Unit in Practice.....	78
4.10	Overall Model of Main Statistical and Administrative Units	78
4.11	Branches, Ancillary Activities, Special Purpose Entities, Franchises and Related Concepts	
	79	
4.11.1	Introductory Remarks	79
4.11.2	Branch.....	79
4.11.3	Ancillary Activities	80
4.11.4	Special Purpose Entities.....	82
4.11.5	Distinction between Head Offices and Holding Companies	85
4.11.6	Franchises.....	86
5	Characteristics of Units.....	87
5.1	Introduction.....	87
5.2	Identification and Contact Characteristics.....	88
5.3	Demographic Characteristics	89
5.4	Economic/Stratification Characteristics	90
5.5	Relationships and Links between Units in the SBR	92
5.6	Relationships and Links with Other Registers	93
5.7	Characteristics by Unit Type.....	94
6	Data Sources for the SBR.....	96
6.1	Introduction.....	96

6.2	General Methods, Procedures and Issues	97
6.2.1	Cooperation with Data Providers	97
6.2.2	Common Identification Systems	98
6.2.3	Identifying and Using External Data Sources.....	98
6.2.4	Combined Administrative and Statistical Register.....	99
6.2.5	Linking and Matching.....	99
6.2.6	Transforming Administrative Units to Statistical Units.....	99
6.3	Administrative Data Sources	100
6.3.1	Typical administrative sources.....	100
6.3.2	Advantages of Using Administrative Data.....	101
6.3.3	Disadvantages of using administrative data	102
6.3.4	Monitoring the Quality of Administrative Data.....	103
6.3.5	Legal Issues and Access to Administrative Sources.....	104
6.4	Using Administrative Sources in Practice	105
6.4.1	Introductory Remark.....	105
6.4.2	Keeping Administrative and Statistical Data Separate	105
6.4.3	Establishing Unique Identifiers.....	105
6.4.4	Using Sector Specific and Additional Sources	106
6.4.5	Thresholds	107
6.4.6	Handling Changes in Administrative Sources	107
6.5	Identifying Statistical Units	107
6.5.1	Identifying Enterprises.....	107
6.5.2	Identifying Local Units and/or Establishments.....	108
6.5.3	Converting from Legal Units to Statistical Units.....	108
6.5.4	Transformation Rules	109
6.5.5	On Line Delivery.....	109
6.6	Statistical Sources.....	109
6.6.1	Economic Census	109
6.6.2	Feedback from Enterprise and Establishment Surveys.....	110
6.6.3	SBR Improvement Surveys	111
6.6.4	Profiling	111
6.7	Combining Administrative and Statistical Sources	112
6.8	Record Linkage in Creation and Maintenance of the SBR.....	113
6.8.1	Introductory Remarks	113
6.8.2	Basic Approach.....	113
6.8.3	Types of Matching.....	114
6.8.4	Standardizing	114
6.8.5	Matching of Text Strings.....	115
6.8.6	Blocking	116
6.8.7	Computing Overall Similarity Measure	117
6.9	Other Data Sources	117
7	Maintenance of the SBR	119
7.1	Introduction.....	119
7.2	SBR Maintenance Strategy	119
7.2.1	Introductory Remarks	119
7.2.2	Use of Administrative Sources	120
7.2.3	Use of Feedback from Surveys	120
7.2.4	Use of SBR Improvement Surveys.....	121
7.2.5	Profiling	121
7.2.6	Dealing with Conflicting Information	121

7.2.7	Maintenance Groups	121
7.2.8	Updating Schedule.....	123
7.2.9	Historical Register	123
7.3	Handling Births, Deaths and Linkage Changes of Statistical Units	124
7.3.1	Demographic Events.....	124
7.3.2	Demographic Events, Continuity and Information Sources for Enterprises	126
7.3.3	Demographic Events, Continuity and Information Sources for Local Units.....	127
7.3.4	Demographic Events, Continuity, Information Sources - Enterprise Groups	128
7.4	Handling Changes in Characteristics of Statistical Units	135
7.4.1	Changes in Identification and Contact Characteristics	135
7.4.2	Changes in Economic/Stratification Characteristics	135
7.5	SBR Maintenance Procedures	136
7.6	Treatment of Errors.....	137
7.6.1	Introduction	137
7.6.2	Taking Account of Different User Needs.....	138
7.6.3	Data about Errors and Corrections	138
7.6.4	The Time Dimension	139
7.6.5	Error Correction Policy	139
8	Survey Frame Methodology	141
8.1	Introduction.....	141
8.2	Establish Output Objectives.....	142
8.3	Identify Concepts.....	144
8.4	Design Outputs	144
8.5	Design Frame and Sample	145
8.6	Design Collection	150
9	Dissemination	152
9.1	Introduction.....	152
9.2	Dissemination of Economic Statistics.....	152
9.2.1	Dissemination Conditions	152
9.2.2	Confidentiality Considerations	153
9.3	Dissemination of Business Demographics	153
9.3.1	Background.....	153
9.3.2	Advantages of Using the SBR	155
9.4	Micro-Data Dissemination.....	156
10	Quality.....	158
10.1	Introduction.....	158
10.2	SBR Quality Dimensions.....	159
10.3	Difference between Survey Quality and SBR Quality.....	160
10.3.1	Use of Administrative Data	160
10.3.2	Heterogeneity of Inputs.....	160
10.3.3	Variability of Inputs over Time.....	160
10.3.4	Technological Requirements	161
10.3.5	Primary Output is Micro-Data	161
10.3.6	Heterogeneity of Users.....	161
10.3.7	Continuous Updating of SBR Data	162
10.4	Assessing the Quality of Administrative Data for the SBR.....	162
10.5	Frame Errors and Their Implications for Surveys	163
10.5.1	Types of Frame Errors	163
10.5.2	Impact of Errors on Survey Frames.....	164

10.6	Classification of Metadata	164
10.7	Quality Assessment Methods	165
10.7.1	Introductory Remark	165
10.7.2	User Survey	165
10.7.3	SBR Audit	166
10.7.4	SBR Improvement Surveys and Quality Measurement Surveys	166
10.7.5	Auditing Clerical Work	166
10.7.6	Macro-Editing	167
10.7.7	Defining and Monitoring Quality Indicators	167
10.8	SBR Quality Indicators	167
10.8.1	Introductory Remarks	167
10.8.2	Conceptual Framework for Quality Indicators	167
10.8.3	Quality of SBR Inputs	168
10.8.4	Quality of SBR Processes	169
10.9	Quality Policy and Improvement	171
10.9.1	Quality Policy	171
10.9.2	General Approach to Improving Quality	172
10.9.3	Improving Timeliness	172
10.9.4	Improving Completeness	173
10.9.5	Improving Coverage	173
10.9.6	Improving Quality Reports	173
10.9.7	Providing Survey Support	173
10.10	Examples of Quality Indicators by Quality Dimension	174
11	Key Considerations in Establishing an SBR	176
11.1	Introduction	176
11.2	Planning Considerations	176
11.2.1	Initial versus Longer Term Scoping: the Importance of a Modular Approach	176
11.2.2	Key Considerations for Establishing a Survey Frame	177
11.3	Governance and Organizational Considerations	178
11.3.1	Introduction	178
11.3.2	Legislative Framework	180
11.3.3	Funding and Development Phases	181
11.3.4	Human Resources	182
11.4	IT Considerations	185
11.4.1	Introductory Remarks	185
11.4.2	Project Management Methodology	186
11.4.3	Software Development Methodology	186
11.4.4	Solution Architecture	186
11.4.5	Database	187
11.4.6	Database Management System (DBMS) Options	188
11.4.7	Programming Requirements	190
11.4.8	IT Environments	191
11.4.9	Programming Language	191
11.4.10	Relationships with Other Systems and Registers	191
11.4.11	Establishing a Unique Identifier for Statistical Units	192
11.4.12	Tools/Software for Record Linkage	193
11.4.13	Job Scheduling Software	193
11.4.14	Documentation	193
11.5	Data Retention	194

Annex A: References	197
UN and Other International Agencies.....	197
European Commission and EUROSTAT.....	198
National Statistical Institutes and Statisticians	199
Other Sources.....	200
Annex B: Glossary	201
Annex C: Characteristics of Units by Unit Type	249
Annex D: Classification of institutional sectors.....	286
Annex E: Examples of Statistical Business Registers	289
Annex E1: SBR Development at Statistics Denmark.....	290
E1.1 Summary.....	290
E1.2 First Establishment of Danish Business Register	291
E1.3 Breaking Point One – Precondition for Second SBR.....	291
E1.4 Content of Second SBR.....	292
E1.5 Statistical Versus Administrative Business Register	292
E1.6 Work to Establish LKAUs.....	293
E1.7 Work to Establish a Coherent Business Register	293
E1.8 Common Identification	294
E1.9 Third Establishment of Danish Business Register.....	294
E1.10 Fourth Establishment of Danish Business Register	294
E1.11 Further Development.....	295
E1.12 Lessons Learned	295
Annex E2: Costa Rica’s SBR programme	297
E2.1 Overview of Costa Rica’s SBR Programme	297
E2.2 Background.....	298
E2.3 Updating Process	298
E2.4 Unique Identifier	299
E2.5 Administrative Records.....	299
E2.6 Cooperation Agreements with Sources of Administrative Records	301
Annex E3: Statistics Canada’s Statistical Business Register.....	302
E3.1 Introduction.....	302
E3.2 Browser Module	303
E3.3 Update Module.....	303
E3.4 Structure Manager Module	303
E3.5 Collection Entity Module	303
E3.6 Workload Module.....	303
E3.7 Survey Interface Module	304
E3.9 Response Burden Module.....	304
E3.10 Reporting and Analysis Tools	304
Annex E4: Business Register at the National Statistics Office of Georgia.....	305
Annex E5. Redeveloping the Malaysian Statistical Business Register.....	309
E5.1 Brief History of Business Register in Malaysia	309
E5.2 Streamlining Business Processes.....	309
E5.3 Integration of the BR with the Collection System.....	310
E5.4 Future Challenges	311
Annex F: Examples Relating to SBR Quality Assurance	313
Annex F1: Istat’s SBR Quality Indicators.....	314
F1.1 Introduction.....	314

F1.2 Quality of Inputs	314
F1.3 Quality of Processes	316
F1.4 Quality of Outputs.....	316
Annex F2: Colombian Experience in Implementation of Quality Processes in the Statistical Directory of Enterprises	319
F2.1 Introduction	319
F2.2 Description of the Business Register of Colombia	319
F2.3 Quality improvement of the administrative information	320
F2.4 Results of the Information Preparation and Processing Processes	321
F2.5 Proposed Quality Indicators	323
Annex F3: Statistics Netherlands' Administrative Data Source Evaluation Checklist.....	327
Annex G: Additional Concepts and Procedures.....	333
Annex G1: Register Based Census.....	334
G1.1 General Description.....	334
G1.2 The Italian 'Virtual' Economic Census (VEC).....	336
Annex G2: Illustration of Time Lag in SBR Maintenance Based on Annual Administrative Source.....	338
Annex G3: Calculating a Check Digit for an Identification Number	340
Annex H: Topics for Further Work and Research.....	342
H1. Introduction	342
H2. Statistical Units and Profiling.....	342
H3. SBR as the Backbone of Business Statistics	343
H4. Use of Administrative Data Sources	343
H5. Use of New Data Sources	343
H6. Economic Globalisation.....	343
H7. Development of New Statistics.....	344
H8. Methodological Developments	344
H9. International Comparability	344

1 Introduction

1.1 Context

In considering national and international economic policy actions, statistical data are an important input for the analysts, researchers and politicians. Also the general public are interested in the status and development of the economy as they are part of the economic system being consumers, entrepreneurs, employers or employees. Therefore, the production of high-quality official economic statistics is an important task of national statistical offices. Such statistics usually comprise structural information, business cycle indicators, production indices, price indexes and various other types of statistical domains. Also the production of national accounts data for analysis from a macro-economic point of view is part of the economic statistical system of a country.

The demand for economic statistics is growing in all countries as well as at international level. To a large extent this is the result of the fact that economies have become more complex, with increasing division of labour both at national and at global level. To remain relevant and to meet user needs economic statistics should, therefore, cover new areas and should provide more detailed information about activities and variables. There is also a general need to improve timeliness, coherence and reliability, where this is possible, at both national and international level.

While business statistics should achieve these new and enlarged goals, it should do so in an efficient way and with as low a response burden as possible. This means efforts must be made to industrialize the production processes, to use administrative data to reduce survey burden, to redesign survey systems, to harmonize surveys and variables, to comprehensively link administrative data and survey sources, to cooperate more closely with administrative authorities and statistical institutions, and to perform similar projects, procedures and processes. Better consistency between the various statistical areas at national and international level has also become a much more important goal than it was in the past, when the economic development was less effected by international dependencies, and the globalisation of production and markets was at a lower level.

In all these developments of business statistics the statistical business register (SBR) plays a key role. Therefore, the challenges of current and future business statistics are at the same time also challenges for the SBR. The SBR delivers the basic information for conducting economic surveys, by providing the populations of statistical units and their characteristics. Further, the SBR provides links to administrative units and registers, thus enabling the use of administrative data for statistical purposes. It also provides unique identifiers enabling linkages at the micro-level across statistical domains as needed.

High quality business statistics depend on high quality SBR. A high quality SBR is one that fulfils the user needs in an optimal way, and is based on international concepts, definitions and classifications. Thus, it serves as the primary basis for international harmonisation of economic statistics in terms of coverage, statistical units and frame methodology. The main users of SBR information are survey staff in the national statistical offices/institutes (NSIs). These staff use the SBR as the source of input for survey design and survey operations. As SBR and survey staff are in the same organisations, close cooperation and exchange on user needs is easily possible.

1.2 What is an SBR?

An SBR is a regularly updated, structured database of specific business units in a territorial area, maintained by an NSI, and used for statistical purposes.

In this definition the terms have the following meanings.

- *Structured database.* This is the most complex property. An SBR is a structured (electronic, relational) database, where defined characteristics are stored for each business unit. In circumstances where there is more than one kind of business unit (which is normally the case) the relationships between these units are also included. For example, for each local unit there is a link to the enterprise to which this unit belongs; for each legal unit there is a link to which enterprise with which it is associated. Business units are assigned unique numerical identifiers so that they are easily identified and to make sure that no units are included twice.
- *Business units.* The business units comprise *observation units, statistical units, and legal/administrative units.* The most important types of statistical units are the *enterprise, the local unit* and the *establishment.* Legal/administrative units are the units registered in administrative registers, such as taxation registers, social security register, company register, register of the chamber of commerce, etc. They serve as the basis for delineation of statistical units. In most cases a legal/administrative unit corresponds to a statistical unit. However, in specific cases, a legal unit might not correspond to a statistical unit. Therefore, in an SBR database these two types of units are recorded as two separate but related types of entities. This sort of approach applies to all types of units. Thus enterprises and local units are recorded as separate types of entities. Even in the case where an enterprise has a single local unit (which is, in fact, the most common situation) two units are to be recorded in the SBR: an enterprise unit and a local unit.
- *Regularly updated.* An SBR that was not regularly updated would soon lose its value as ongoing changes in the numbers and structure of the business units were not incorporated in the database. Such changes include: new business units being created and others being closed; business units being merged or changing legal form, location or activity; units having higher or lower turnover over time or engaging more or less employees than in the past period. Also the stratification variables used for the delineation of survey frames may change over time.

The frequency of SBR updating depends on data availability, mainly on the availability of administrative data. If data for SBR updating are only available on an annual basis, use of the SBR for monthly or quarterly surveys is hampered. Some administrative data are made available on a monthly or quarterly basis, others only on an annual basis. Also, data acquisition from the various sources occurs on different dates; some administrative data may be available quite early after the period to which they refer, others may be available much later. A further issue is the time lag between an actual business event and the date when it is recorded in the administrative or other data base.

SBR updating is usually a continuous process in order to minimise time lags in processing data.

- *Territorial area:* An SBR is established to cover all business units that are resident in a specific territorial area. The territorial area is normally the territory of a country. Therefore, an affiliate of a resident enterprise in a foreign country would not be recorded in the SBR of the country of the enterprise. This affiliate would also not be

covered in the administrative sources and therefore information for updating would not be available. However, information on the foreign affiliates might be collected by a specific statistical survey.

The backbone role of the SBR is not be easily achievable if a separate SBR is established and maintained for each region of a country. This would lead to various practical and conceptual problems. For example, an enterprise could be active in different regions. In order to avoid such problems it is highly recommended that there should be only one SBR in a country covering all respective entities resident in that country. If regional SBRs exist they must be coordinated at national level.

The restriction to the national territory has the consequence that multinational enterprise groups are not covered in their entirety, only the national parts of them. The coverage of multinationals would need a supra-national SBR with appropriate cross-country cooperation between the NSIs.

Maintained by an NSI. Administrative registers are maintained by administrative authorities for purposes of public administration and programmes. Their coverage, characteristics, methods of updating, etc. are designed to serve the specific administrative purposes. An administrative business register might not cover all respective businesses but only selected ones, for example those with a particular legal form (e.g. companies), or those with a particular type of economic activity (e.g. farms), or those in a certain region. An SBR has to fulfil the requirements for business and economic statistics. The coverage requirements are often quite different and more comprehensive than for administrative registers. Thus an SBR has to be established and maintained by an NSI, in fulfilling its task of producing official business statistics.

- *Statistical purposes:* An SBR is established in order to provide the frame population for business surveys, and for other statistical purposes. The requirements for an SBR can therefore be quite different from the requirements of an administrative register. Even if the SBR is based on information from administrative registers, the concepts, characteristics, methods of maintenance, etc. of SBR need to be based on statistical concepts, preferably on internationally recommended ones.
- There is also one other important feature in addition to the different nature of administrative and statistical business registers. The SBR is only to be used for statistical purposes. It may not be used for other purposes, especially for administrative purposes. The reason for this is the confidentiality of individual data collected by NSIs. Data on individual natural persons or businesses may not be disclosed to other institutions, to administrations or to the public. They are exclusively used for statistical purposes.

In satisfying the requirements implicit in the above definition, an SBR is a complex system both in terms of structure and in terms of technical implementation. It might seem that a quite simple system, such as a list of businesses with the appropriate variables, could serve to provide sets of units for survey frames for collection of data. This could be the case when only one type of unit, for example enterprise, was of interest, and there were very few surveys. However, even then such a list would have to contain many variables, including name, address, legal form, economic activity and other classification codes, and its handling and updating would be a difficult task. Where there are several types of units and many surveys to consider, a simple list is impractical, a database is needed.

A general principle is that a data update to the SBR should be time stamped to know, first, the date the change has been made, and, second, the date when the change becomes effective. For

instance, a merger of two enterprises might have taken place (be effective from) 1 January 2015, but the appropriate changes to the SBR might be made at a later date, say 1 June 2015, when the information becomes available for the SBR updating purposes.

A second principle is that SBR data should never be physically deleted. For example:

- an enterprise that has ceased activity will not be deleted, rather the *activity status* of the enterprise will be set to *inactive*;
- a change of the location of an enterprise will be recorded in such a way that the old address is marked as *no longer valid* and the new address is introduced with an effective starting date.

To manage this sorts of requirements a well-structured relational database is needed. They cannot be managed by as a list.

As for any other statistical product, quality of the SBR is of extreme importance, particularly as the SBR is not an end in itself, but an input to all the economic statistical products based on it. The quality of a SBR can be measured by the same criteria that are used for measuring the quality of statistical data: relevance, accuracy, punctuality, accessibility, comparability and coherence. More specifically this means: that the information provided to the users corresponds to their needs; that the business units covered are in accordance with the defined target population; that the data are as timely as possible: that business unit coverage and characteristics are as complete and accurate as possible: that the data are comparable over the activities, regions and internationally as a result of using the same concepts, methods and definitions: and that the SBR is coherent with other statistical datasets, where relevant.

1.3 Use of an SBR

The main purpose of the SBR is to provide sets of units referred to as *frames* to surveys. A survey frame is a list of units together with the characteristics of these units required in conducting the survey. By providing frames for all relevant surveys from a central source, the SBR is also acting as a coordinating instrument. As the SBR units and their characteristics are updated on a continuous basis, so survey frames can be derived from the SBR can be similarly updated. Updating centrally in the SBR is also much more efficient than if updates of survey frames were done by the different survey areas themselves. Thus, a single SBR is a good solution not only for harmonised statistics but also for cost efficiency.

However, the SBR can do more than simply provide frames. It can also be used to combine data from different sources to improve coverage. It can provide a mechanism for identifying units and avoiding double counting, it can facilitate central storage of metadata and it can coordinate various data streams. All of these help optimize the quality of statistical processes before direct data collection has even begun.

Variations in the coverage, content and quality of SBRs across countries make comparison of economic statistics more difficult. Therefore the coverage and content of the SBR should be harmonised as much as possible and international concepts and definitions should be used to the maximum extent possible. Economic statistics describe economic production processes and financial transactions in various statistical domains. Whilst each domain has its own specific peculiarities, the data collected should be comparable across domains. This is only possible if the populations for the various domains are derived and used in a similar and coordinated way during the course of statistical processing.

Using administrative data more extensively for statistical purposes is an important strategic goal for NSIs in their efforts to keep up with ever growing demands for economic data by governments, international bodies and researchers, in particular data for microeconomic

analysis of small areas, and for specific sectors of activity aggregates. Whilst administrative data may not include precisely the information that the statistician would like to collect and disseminate, they have the virtue of providing more (possibly nearly complete) coverage of a target population, whereas sample surveys cover only a relatively small proportion. The use of administrative data may therefore decrease or eliminate, sampling errors, significantly reduce or remove non-response, and may provide more accurate and detailed estimates for various sub-populations. Thus, to reduce both the costs of data collection and response burden, administrative data sources should be further integrated into the statistical processes used in collecting and compiling economic statistics. The starting point is using administrative data in the creation of statistical units in the SBR.

Use of administrative data also increases the range of tasks that can be achieved by an SBR. It becomes not only a database providing frames for survey purposes but also a powerful database that can be used for the direct production of statistics. Furthermore, the link to administrative data and registers facilitates the use of administrative data for supplementing or replacing survey characteristics.

New developments in the production of economic statistics will extend the role and use of the SBR. Globalisation, increase in the use of administrative data, industrialisation, and upcoming new data sources (for example big data, Internet) are some of the developments that have to be taken into account when looking at the future use of the SBR.

Within an NSI, the SBR should be the single, central place where statistical units are derived and maintained for economic statistics. In this way the economic behaviour of various populations of enterprises can be compared over space and in time. The SBR's coordinating role is even stronger when statisticians gathering or compiling statistics refer to, and use, the units provided by the SBR without further modification. However, this requires their satisfaction with the quality of the data provided by the SBR and hence with the administrative sources on which these data are based.

The demand for better, quicker and more detailed business statistics, and the need of policy makers for more comparable international economic data, is likely to force many NSIs to extend the roles of their SBRs and create a multipurpose system which supports statisticians in improving the efficiency of their processes and quality of their outputs.

1.4 Aims of SBR Guidelines

The main objectives of the Guidelines as mentioned in the Background and Acknowledgement were to:

- 6) Provide practical guidance on core issues of establishing and maintaining the SBR.
- 7) Clarify typology, concepts and definitions, including for statistical units.
- 8) Provide guidance on the use of administrative and other sources for the establishment and updating of the SBR.
- 9) Provide guidance on how to use the SBR in its own right for production of statistics and how information from the SBR can be combined with information from other statistical registers, administrative sources or surveys to produce new statistics.
- 10) Provide guidance on the role of SBRs in the modernisation of statistical production and services.

The Task Force has tried to ensure that the Guidelines satisfy these objectives. However, guidelines are never finished and these Guidelines will need to be updated from time to time to integrate guidance on national and international developments and to deal with new challenges that arise for the SBR.

The Guidelines are targeted at SBR management, SBR staff members, the staff of business statistics, and the staff dealing with respondent relations and with administrative authorities that deliver data to the SBR. They may also be useful for training purposes.

These guidelines are not sufficiently detailed and comprehensive to cover all the SBR issues, concepts, definitions and methods that are important in every country. They do not and could not take fully into account all the various national institutional structures and concerns. Like any other international manual, the Guidelines can only provide guidance in the form of a broad range of concepts and explanations that need to be interpreted within each particular national context, which is invariably different from country to country.

The main conceptual framework on which the Guidelines are based is the System of National Accounts 2008 (SNA2008) and the European version (ESA). The SNA2008 is the international framework for integrated economic statistics and therefore the natural reference for the concepts and definitions of the relevant terms of the SBR. International classification manuals, in particular the International Standard Industrial Classification of All Economic Activities (ISIC) Revision 4, are also referenced.

A variety of other international guidance manuals that deal with specific phenomena or domains should also be mentioned. They include manuals on the use of administrative data, on the informal sector, on business demography, on foreign affiliates statistics, and on integrated economic statistics, which also addresses the key role of the SBR. Other relevant international manuals are currently under preparation, such as the guidelines on measurement of global production.

All these international manuals and various other documents, papers and reports were consulted and utilized in developing and drafting of these guidelines. However, there are two international guidelines which directly focus on the SBR: the *Business Registers Recommendations Manual* of Eurostat (latest version 2010) and the *Guidelines for Building Statistical Business Registers in Africa* (African Development Bank 2012). These two manuals have been extensively used as it was a Guidelines development principle that already existing materials should be used whenever feasible and appropriate.

As already stated, international manuals and guidelines can provide concepts and definitions needed for harmonized statistics, and thus help developing the national implementation. However, exchange and share of experience between countries and bilateral co-operations are also recommended. There are two main international fora for exchange and discussion of national experience, country practices and international developments in the area of SBR. One is the Wiesbaden Group on Business Registers, which organizes biennial meetings and the second is the biennial, joint UNECE, Eurostat, and OECD Expert Group Meeting on Business Registers, which takes place in years where there is no Wiesbaden Group meeting.

1.5 Overview of the Guidelines

The Guidelines are structured into 11 chapters and 8 annexes. Each chapter can be read separately, according to the interests and needs of the reader. Cross-references are provided. In the following paragraphs short description of the topics and contents of each chapter and of the annexes are provided.

Chapter 2 on the roles of the SBR provides an overview of the various roles that SBR play in the context of the production of business statistics. In total eight different roles are distinguished and described. The first role is the central role of the maintenance of a set of statistical units providing a gateway between data from various input sources (mainly administrative sources) and statistical units. This means continuous updating of the content of

a *live register*. The second role is to provide *register snapshots* and *frozen frames*, from which *survey frames* are derived. These roles are the most central and important ones and focus on the support function of the SBR for business surveys. An additional role is direct support for the surveys by survey registration, monitoring survey response and measurement and control of response burden.

Chapter 2 also describes some additional new roles that have emerged in the last years and show the usefulness of SBR databases. The most important is the production of statistics directly based on the SBR. Examples are economic censuses and business demography statistics. In some countries individual data on enterprises can be made public and thus support directory/lists of businesses or georeferenced information systems. Due to the globalization, international comparability of statistical data becomes more important. A key enabler would therefore be the exchange micro-data between countries through the SBR, which would allow a significant increase in quality of global statistics. A related SBR role for the SBR is the integration of the SBR in the production process of business statistics.

Chapter 3 describes coverage of an SBR, examining the concept from the perspective of various institutional unit types and sectors of the economy. It notes that, in principle, the SBR should record all units in the national economy that contribute to gross domestic product (GDP) but that, in practice, this is not be achievable due to various constraints, which are described. The chapter deals with the relevant concepts, such as market and non-market producers, observed and non-observed economy, informal sector and illegal activities. It provides recommendations for a minimum coverage of an SBR.

Chapters 4 and 5 deal with the concepts and characteristics of statistical units. While chapter 4 introduces the various statistical units and their delineation, chapter 5 provides information on their characteristics that should be recorded in the SBR. The main statistical units discussed in Chapter 4 are the *enterprise group*, the *enterprise*, the *establishment*, the *kind-of-activity unit* and the *local unit*. The chapter also deals with various forms of legal and administrative units that are the main building blocks for delineation of the statistical units. Also the *units model* that indicates the definitions the various types of units and their relationships are described.

Connected with Chapter 5 are the Annexes C and D. Annex C lists and explains in detail the characteristics of the legal/administrative units and the various statistical units. Annex D introduces to the SNA classification of institutional sectors which in addition to the economic activity classification is the second most important statistical classification in the SBR.

Chapter 6 deals with the data sources for the establishment and maintenance of the SBR. It considers not only the various administrative data sources, which are the main sources of input information, but also all other sources. The chapter describes the general methods, procedures and issues in relation to the data sources, especially the co-operation with the administrative authorities. It gives advice on the use of administrative data and the problems that may occur when administrative data are used for statistical purposes. The chapter also gives an introduction to the record linkage methods that are used when the linking of individual units in the various data sources is not supported by common identifiers. Methods based on other data such as name and address have to be applied to identify the same units in the different sources so to correctly link them.

Related to Chapter 6 is Annex F3 which contains an excerpt of the administrative data source evaluation checklist developed by Statistics Netherlands.

The concepts of the update and maintenance of the SBR are the subject of *Chapter 7*. It provides the fundamentals and basic considerations for the establishment of a maintenance strategy that is based on the needs of the users, especially the survey staff. The chapter deals

with the recording of demographic events, such as enterprise births, the handling of the changes in the characteristics, the continuity rules and, last but not least, the treatment of errors. Annex G2 illustrates an issue arising in terms of time lags in the data sources used for maintenance.

Survey frame methodology is the subject of *Chapter 8*. Starting from the Generic Statistical Business Process Model (GSBPM), the specific roles of the SBR are described in relation to the survey requirements. The various population concepts lead to the discussion of frame design. The chapter deals with the effects of over-coverage and under-coverage errors and their correction, as well as other kinds of frame errors.

The emerging role of producing statistics directly based on the SBR is discussed in *Chapter 9*. As previously noted, the SBR may itself now also have a dissemination role, whereas in the past only survey results were published. The full coverage and good timeliness of SBR data are used to produce census type statistics and business demography statistics.

Dissemination also requires consideration of confidentiality rules, which certainly differ between the countries. The provision of micro-data certainly needs a specific legal basis. In some countries it is possible to provide micro-data to researchers or even to publish individual business information. In most countries it is not.

Annex G1 give an overview of the Italian register-based census as an example for statistics directly based on SBR.

Chapter 10 discusses SBR quality. As already mentioned, SBR quality is a crucial issue as it has a major impact on the quality of business statistics, especially in terms of completeness of coverage and accuracy of the characteristics (such as economic activity code) used for survey frames and sampling. Beginning with the definition of quality and the explanation of the quality dimensions related to the SBR, the chapter deals with the methods of quality assessment. It proposes quality indicators and provides the basic elements of a quality improvement policy.

Annex F1 describes the Italian SBR quality indicators and Annex F2 the Columbian experience in implementing quality processes.

Chapter 11 on key considerations in establishing an SBR is of particular interest for countries where an SBR is being developed or under review. It deals with three related areas: planning, governance and organisational and IT considerations. It covers the various aspects that need to be taken into account in the process of establishing an SBR (including long term scoping and modular approach). It deals with the legislative framework, funding and human resources, and it discusses and provides guidance on IT topics such as system architecture, database management, programming and software tools. A particular case of SBR architecture from Statistics Canada is described in Annex E3. Annex G3 illustrates the calculation of a check digit for an identification number.

Annexes not mentioned above include Annex A , which provides a list of references, Annex B which contains a glossary of the key terms used in the guidelines and Annex E which presents examples of SBRs in Costa Rica, Canada, Denmark, Georgia and Malaysia.

The final annex, Annex H, lists and briefly describes topics for further work and research as recommended by the Task Force. These topics were identified during the course of development of the Guidelines and are prompted by the need to deal with emerging issues that will pose major challenges in the future. The topics include the concepts and delineation of statistical units, the use of administrative data source, the exploration of new data sources, and strengthening the roles of SBR as the backbone for economic statistics and a core element of statistical modernisation role.

1.6 Terminology

Introductory Remark

As in guidelines of any sort, it is vital that everyone is on the same page as regards the meanings of the key terms used. The aim of the following paragraphs is to ensure all key terms are identified and defined, and that there is minimum use of other terms that are synonyms but could be interpreted as having slightly different meanings. In the case of synonyms only one term is used throughout the Guidelines, other term(s) being referenced the first time the selected term is used in the text. In the cases where the terms are not synonyms, the intended distinctions between them are made clear the first time they are used.

The definitions in the document and summarised below are based on terminology defined in SNA2008 and ISIC Rev 4, supplemented where necessary by terms from the Guidelines for Building Statistical Business Registers in Africa developed by the African Development Bank. More details are given in the Glossary

Terms for the various types of units

The terms *business, company, establishment, enterprise, unit, statistical unit, respondent, economic unit, economic organization, economic operator, economic producer, legal entity, legal unit, local unit, legal local unit* tend to be used rather haphazardly in the general literature. In the Guidelines they have the following specific meanings.

- *Unit* - single distinct part or object - as generally understood in the English language – needing further qualification to have a more precise meaning.
- *Entity* - a synonym for *unit*, not used except where it appears in a quotation.
- *Legal unit* - a unit that is recognized by law or society, independently of the persons or institutions that own it. The characteristics of a legal unit are that it owns goods or assets, it incurs liabilities and it enters into contracts. A legal unit always forms, either by itself or sometimes in combination with other legal units, the legal basis for a statistical unit.
- *Economic unit*- a *legal unit*, or part of a *legal unit*, with economic production as defined in the SNA2008.
- *Economic operator, economic producer, and economic organisation* are synonyms for *economic unit*, not used except where they appear in a quotation.
- *Administrative unit* – a unit defined by a legal unit for the purposes of conforming with an administrative regulation, for example VAT.
- *Operational unit* – a unit defined by a legal unit for the purposes of organising itself, for example a division, branch, workshop, warehouse, or outlet.
- *Local unit of legal unit* – an *administrative* or *operational unit* at a single location.

Standard statistical units

- *Statistical unit* – a unit defined for statistical purposes; the basic unit of observation within a statistical survey; the unit for which information is sought and for which statistics are ultimately compiled. There are four types of statistical unit defined in SNA2008 and ISIC Rev 4, as defined immediately below. In addition, enterprises may be grouped by ownership and control into an *enterprise group*.
- *Enterprise* - a *legal unit* (or the smallest set of *legal units*) that produces goods or services and that has autonomy with respect to financial and investment decision-making. An enterprise may be a corporation (or quasi-corporation), a non-profit

institution, or an unincorporated enterprise. An unincorporated enterprise is household or government unit in its capacity as a producer of goods or services.

- *Establishment*- 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.
- *Local kind-of-activity unit* – synonym for *establishment*.
- *Kind-of-activity unit (KAU)* - 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.
- *Local unit* - an *enterprise* or *a part of an enterprise* (for example, a workshop, factory, warehouse, office, mine or depot) that engages in productive activity at or from one location.

Other types of units

- *Business* – (loosely used) as a type of enterprise, namely a commercial enterprise or legal unit with commercial economic activity.
- *Corporation* - legal unit created for the purpose of producing goods or services for the market that may be a source of profit or gain to its owner; collectively owned by the shareholders who have the authority to appoint directors responsible for its general management.
- *Company* – synonym for corporation.
- *Observation unit* – a unit, *about which* data are obtained during the course of a survey; usually a statistical unit, or, if data cannot be obtained about a target statistical unit, then some other unit about which data can be obtained and from which data for a statistical unit can be compiled.
- *Reporting unit* – the unit *from which* data about an observation unit are obtained during the course of a survey. It may, or may not, be the same as the observation unit. An example of it not being the same is where an accounting business reports data on behalf of a client business that is the actual subject of the survey.

Data Elements associated with SBRs

- *Register* - as generally understood in the English language – database on items or events, often kept by official whose job is to do so - needing further qualification to have a precise meaning.
- *Business register* – database on businesses, or, more generally of any administrative or statistical units, kept for a commercial, administrative, or statistical purpose, including relevant characteristics of the business/units.
- *Statistical business register (SBR)* – database on statistical and other units kept by an NSI for statistical purposes.
- *Live register* – the part of an SBR that is being continually updated with new information about the units and their characteristics.
- *Register snapshot* – a copy of the *live register* as of a given point in time.
- *Register picture* – synonym of *register snapshot*.
- *Frozen register* - synonym of register snapshot.
- *Frozen frame* – *register snapshot* containing only active statistical units.
- *Common frame* - synonym of *frozen frame*.

- *Frame* – set of statistical units that forms the actual population for a survey.
- *Frame population* – synonym of *frame*.
- *Historical register* – capacity to view content of the *live register* at points in time in the past -can be via a set of consecutive *register snapshots*.

Miscellaneous other terms

The terms *enterprise* and *economic statistics* are used in preference to the narrower terms *business* and *business statistics* as an SBR may well include lists of government units and non-profit organisations, which are enterprises but not businesses, and which are the subject of some surveys (for example employment) for which frames are derived from the SBR.

The terms *characteristic*, *variable*, *property*, and *attribute* as applied to units in the SBR are considered as synonyms and the term *characteristic* is used in the context of the SBR and survey frames. The term *variable* is more appropriate in the case of surveys where sampling is involved and there is thus the notion of variability due the probability mechanism involved in selecting the samples.

Any type of survey conducted by SBR staff specifically to improve SBR quality is referred to as an *SBR improvement survey*. Synonyms are quality improvement survey, nature of business survey, control survey and SBR survey, SBR coverage survey, and SBR quality improvement survey.

As regards the organisations or persons using of SBR outputs, the term *user* is used, rather than *recipient*, *receiver*, *customer*. Most SBR users are internal to the national statistical institute, being the staff of surveys that draw frames from the SBR.

In the context of surveys (including SBR improvement surveys) the term *response burden* is used in preference to synonyms such as *respondent burden*, *reporting burden* and *administrative burden*.

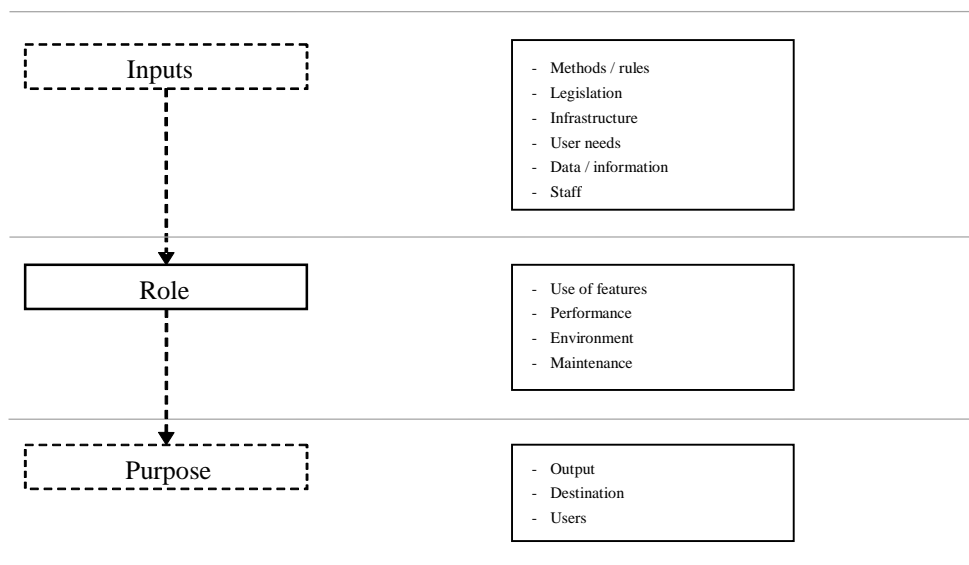
2 Roles of the SBR

2.1 Introduction

Concept of a Role

As discussed in Chapter 1, an SBR is a vital component of the core statistical infrastructure supporting collection of economic data and production of economic statistics. This chapter discusses the primary roles of an SBR in maintaining and providing sets of statistical units, in particular enterprises, and in providing frames for surveys. This primary role of an SBR is also referred to as the backbone role. This chapter also discusses various other SBR roles that are related to the primary goal and build upon the methodology and data bases of SBR. It draws on material from Eurostat working groups and international conferences and meetings and on experiences and procedures from the Dutch, Italian, Mexican and South African NSIs. In discussing an SBR “role”, information is needed about the inputs in terms of the relevant regulation and legislation, user needs, data sources, infrastructure, methods and staff. Each role of the SBR can then be defined as the set of features and functionalities to be used to perform the role for a specific purpose. It should also be clear who the users are, for what they will use the outputs, and the quality of the outputs. This concept is summarized in Figure 2.1 and is used to explain the SBR roles in the following sections.

Figure 2.1: Concept of a role



Summary of SBR Roles

The central all-encompassing role of an SBR is to serve as the coordinating mechanism for economic statistics. It should provide all the information needed by the statisticians who are responsible for the various economic surveys and related statistical processes and outputs. In fulfilling its central role, the SBR has seven specific roles as listed and briefly summarised in Figure 2.2 below and described in detail in the following sections. For each role the basic goal is noted in the figure.

Figure 2. 2: Overview of the SBR Roles

	Role	Goal
1	SBR Live Register	Provide the gateway between data from various input sources and statistical units
2	SBR Register Snapshot and Frozen Frame	Provide populations of statistical (and possibly administrative units) at fixed points in time.
3	SBR Survey Frame	Provide a set of statistical units for a survey, valid for a specified reference period, with all characteristics required.
4	SBR Survey Support	Monitor survey response and measure and control response burden
5	SBR Statistics	Produce statistics based directly on the SBR
6	SBR Information Source	Provide lists of enterprises and their locations and possibly other characteristics
7	SBR International Data Exchange	Promote coherence in international statistics
8	SBR in modernisation of statistical production and services	Promote integration of SBR within the production processes for economic statistics

The first role (*SBR Live Register*) of the SBR listed in Figure 2.2 is to maintain a set of statistical units as the foundation for creating frames for economic surveys. On a daily basis, the SBR is updated with new information from various sources. This information is used to create and deactivate statistical units, and to update their characteristics. This functionality of continuously creating, updating and deleting of the corresponding units is referred to as the “*live register*”.

The second role (*SBR Register Snapshot and Frozen Frame*) is to produce register snapshots and frozen frames from the live register as the basis for subsequent generation of survey frames. A *register snapshot* contains the set of statistical units valid for a specific reference period and hence represents a coordinated population of statistical units in space and time. A *frozen frame* contains all the active statistical units in a snapshot that have sufficient information associated with them for sample selection purposes. A snapshot (or frozen frame) may also provide the links between statistical and administrative units (for data collection) for the specific reference period.

The third role (*SBR Survey Frame*) of the SBR is to provide frames for surveys. For each survey the SBR provides the set of statistical units in scope for the survey (typically a subset of the frozen frame) together with the characteristics required for stratification and sampling, and, for the selected units, the information required to contact and communicate with them.

The fourth role (*SBR Survey Support*) concerns survey registration and survey control. It involves tracking the reporting statuses of enterprises and the response burden on enterprises imposed by surveys. Response burden should be minimised. The SBR supports this by storing information that can be used to optimize survey design as well as to monitor reporting statuses and to compile overall respondent burden.

The fifth (*SBR Statistics*) and sixth (*SBR Information Source*) roles are to provide macro-data (statistical aggregates), and micro-data (enterprise names, address and characteristics), respectively, to users external to the NSI. Such dissemination is constrained by the confidentiality provisions applying to the NSI, particularly in the case of micro-data.

The seventh role (*SBR International Data Exchange*) depends on current and upcoming developments in the production of economic statistics, in particular the exchange of cross-border information, also increase in the use of administrative data and new data sources.

The eighth role (*SBR in modernisation of statistical production and services*) refers to the SBR's role in coordinating, linking and connecting units that occur in all kinds of sources relevant for statistics. The SBR is a potential gateway for combining data from traditional sources and new media and may serve as a data warehouse providing integrated data to users. It can also facilitate new, industrialised and standardised production of statistics, at both national and international levels.

2.2 SBR Live Register

Statistical units in economic statistics are used to describe various populations of organizations with similar production processes or financial arrangements. An important role of the SBR is to maintain, and to keep track of changes in, statistical units and their characteristics that occur in the economy. 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 to be 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 communications with the owners of the sources 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.

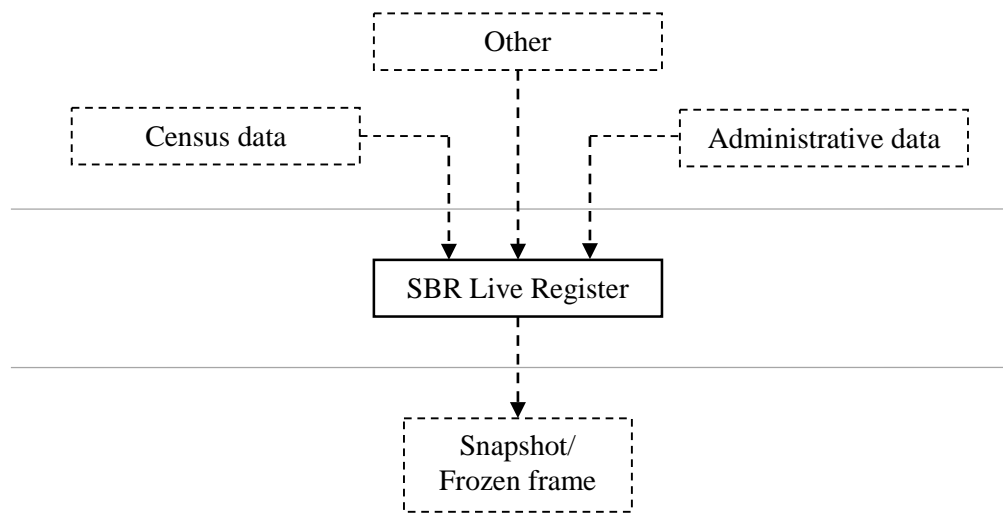
Statistical units are created in the live register. There are several types of statistical units (as further discussed in Chapter 4). The most important type is the *enterprise*, usually defined in accordance with an international definition (as further discussed in Chapter 4) though implementation may require some compromise.

An economic census used to be one of the main sources from which to derive a common frame in order to conduct and coordinate business surveys. Nowadays SBRs are usually based on legal and/or administrative units obtained from a network of administrative registers and feedback from economic surveys, as illustrated in Figure 2.3.

The processes for creation and maintenance of statistical units, in particular the set of enterprises, are the most important factors in determining the quality and the usability of SBR populations for economic statistics. A well-defined maintenance strategy is a key aspect of an SBR. The units in an SBR are maintained and updated with the most recent information available, as further discussed in Chapters 6 and 7.

Timeliness and linkage issues can pose problems when using multiple sources, but their use in combination allows an NSI to derive all relevant statistical information and hence to reduce response burden and create new and more detailed outputs.

Figure 2.3 SBR Live Register



In order to make full use of source information, the values of key characteristics of units in the sources are harmonised using well defined standards, for example, for dates, telephone numbers, addresses, legal form, economic activity, number of employees, etc., as further discussed in Chapter 5.

Monitoring the *continuity* of statistical units is also an important aspect of SBR maintenance. In the event of new information about a statistical unit, a decision has to be made whether the unit keeps its identity in the SBR, is deactivated in the SBR, or is registered as a new unit. This is further discussed in Chapter 7.

When updating administrative and statistical units and their characteristics, the existing and previous values should be retained in the live register if possible, or at least be retained in a *historical register*. This enables any change over time in any unit used for statistical production to be traceable back to the source, which may help understanding anomalies in statistical processes. Key SBR users, particularly those with direct access to the live register, must be trained so that they are aware why historical versions exist and why they may differ for certain units for different extraction dates.

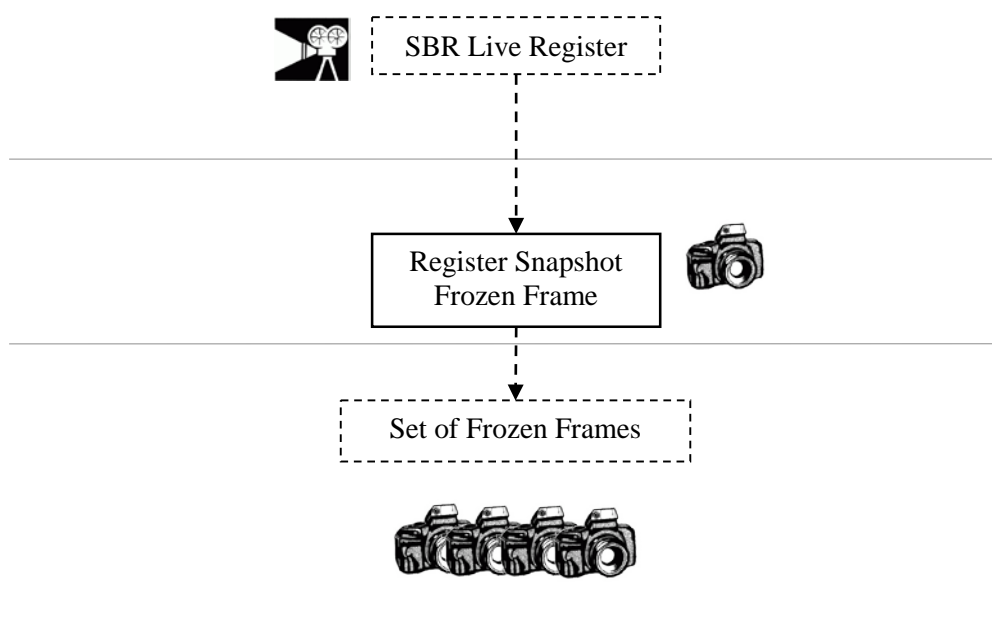
Units from different sources are linked at micro level. However, not all legal or administrative units lead to the creation of a statistical unit or are necessarily linked to a corresponding statistical unit. This can happen because a legal/administrative unit is dormant, or inactive, or the information to determine whether the unit is active is simply not available in the source data. Thus, some legal/administrative units are not represented in the frozen frame since it contains only statistical units and are not, therefore, used in constructing survey frames or in subsequent compilation of statistical aggregates. However these *floating* legal/administrative units and their data should remain available in the live register. Survey statisticians may need the information about such units to impute missing data, or to conduct surveys on specific groups of missing units, or to help in profiling.

2.3 SBR Register Snapshot and Frozen Frame

The most important role of an SBR is its coordinating task. This backbone role is best fulfilled if all statisticians in an NSI use the units delineated in the SBR as of specified reference dates. In other words, it must be possible for statisticians to view and retrieve a common set of units from the SBR as of specific reference date.

Thus, whilst the live register changes on a regular basis, statisticians need coordinated input for their processes. If surveys are conducted with different time lags from a given reference period, the composition and the quality of the population of statistical units in the live register at the times the survey frames are drawn will differ. This indicates the need to derive a “frozen version” of the population of statistical units - a composition of units that is valid and unchanging for a particular moment in time for use by all surveys. This is realised by taking a “snapshot” of the live register (which itself may be viewed as a *continuously changing movie*) as illustrated in Figure 2.4

Figure 2.4: SBR Frozen Frame



Each *snapshot* is extracted from the live register at a point in time. The set of all snapshots represents a set of coordinated populations of statistical and administrative units at specified time points, with a certain quality level. More precisely, a snapshot at date t consists of:

- statistical units (SUs) at t ;
- administrative units linked to SUs at t (or *links* to these administrative units); and possibly
- administrative units not linked to SUs at t , but in the live register at t .

The units in a particular register snapshot that are “active” as of the specified reference period together with all the information about their characteristics needed for survey sample selection and other statistical purposes, constitute the “frozen frame” for that period. More specifically, a frozen frame at date t consists of:

- SUs active at t , and with sufficient information for survey sample selection; and
- administrative units linked to SUs at t , or links to these units.

In some literature on this topic a frozen frame is called a *master frame*, or a *common frame*. This is because typically (as discussed below) several survey frames are derived from this single common extraction from the live register, thereby coordinating the survey frames at a moment in time.

The quality of a frozen frame depends on the quality of the information about the existence of units and their characteristics and on the reference period for which statistics are being compiled. This should be taken into account in trying to ensure that short term and annual indicators, which refer to overlapping reference periods, are coherent. It may lead to the revision of previously published short time indicators when annual indicators are produced. If a statistician needs more information about certain units than is available from the frozen frame, then this extra information may be obtained by linking back to the same units in the live register.

In summary, a snapshot is a copy of all the statistical units in the live register, and all administrative units or at least all links from statistical units to administrative units. It is an intermediary step between live register and 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 (as further discussed below). It may contain many inactive enterprises. A 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. Its aim is to include all units and all characteristics that are used by subsequent processes and nothing else. In other words it is a trimmed down version of the snapshot that is easier to manipulate because the (possibly large number of) inactive units are not there.

If the frozen frame for an annual survey for reference period t is also used to observe the short time indicators in year $t+1$, it will not include the information on newly active units for reference year $t+1$. Therefore some NSIs create new version of the frozen frame at later points in time for subannual surveys. Thus frozen frames may be produced annually, quarterly, or monthly. In all cases, it is important to control their use so that the resulting survey frames are as coherent as possible.

A set of frozen frames effectively constitutes a *historical register* which:

- coordinates populations of statistical units;
- provides links between statistical and administrative units;
- links the SBR to other sources (administrative registrations, survey samples) used for statistical outputs; and
- enables the reconstruction of the history of the units it contains.

Statisticians often compare particular populations of statistical units at different reference points in time and try to describe and explain the differences. The basic data for this sort of analysis is available from the historical register. The evolution of a particular population of enterprises may be presented in the form of enterprise demographic statistics.

Frozen frames for business statistics at Statistics Netherlands

The Dutch SBR is the container of statistical units, and of administrative units coming from the Trade Register and the Tax Office. This information is needed to describe financial and production processes of resident enterprises.

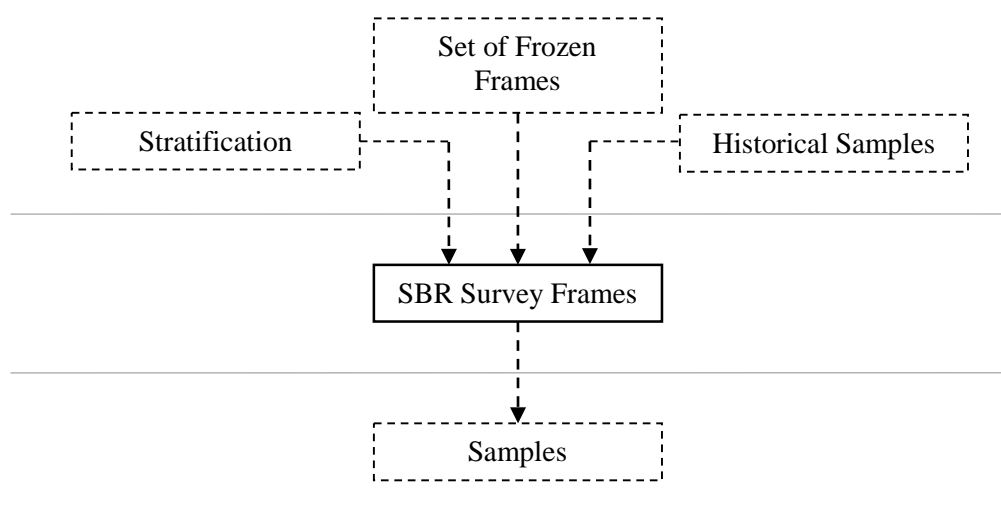
Statisticians within Statistics Netherlands are obliged to use the SBR in producing economic indicators based on coordinated enterprise populations. Every month, an actual frozen frame is derived from the live register. This frozen frame represents the coordinated population that is valid for that month. This coordinated population consists of enterprise groups, enterprises and local units, and their main characteristics. In addition the statistical units are linked to the legal and the administrative units valid for that month. Besides a list with frame errors referring to activity codes or size classes of statistical units, is kept in order to be able to generate several instances of each frozen frame with improved quality levels.

Each month the frozen frame is stored a separate environment, which can be consulted for statistical production. This environment includes all past frozen frames and coordinates the populations used for statistical purposes. From this set of frozen frames an authorized statistician can retrieve the micro data needed for a statistical activity, for example, linking tax data and survey data to a frozen frame.

2.4 SBR Survey Frame

A *survey frame* is a list of units from which a survey sample is selected, together with the characteristics required to draw the sample according the agreed design (i.e. characteristics required for stratification, sample size determination and sample selection).

Figure 2.5: SBR Survey Frame



The frame for an economic survey should be a subset of the frozen frame, comprising the set of statistical and (linked) administrative units that match the specification of the survey target

population and are active during the survey reference period, together with the characteristics that will be needed for the survey. Thus, for example, a survey of employment will include active 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 active units that have an ISIC code in the manufacturing group, whether they have employees or not. A survey of capital expenditure will typically include active units in all industries above a certain size. Thus, the frames for different surveys for a given reference period are different from one another but are extracted from the same common set of units, namely the frozen frame.

The primary benefits of survey frames that are extracted from a coordinated set of frozen frames are:

- harmonisation of the surveys;
- increased potential for integration of survey data;
- reduction of costs;
- prevention of double counting; and
- more coherence in the resulting statistics.

Of course, these benefits can only be realized when a single live register is maintained and used to derive the frozen frames used for all relevant surveys.

It may be that an optimized sample design cannot be applied because some important stratification variables (e.g. turnover) that correlate strongly with the key indicator to be compiled, are missing from the live register and hence from the frozen frame. In this case it should be possible to enrich the statistical units with characteristics available from other (independent) sources and satellites in order to complete the survey frame.

A sample design which uses a panel population may be very complex to apply. In this case information on units in historical samples for similar surveys should be made available.

Some NSIs exclude enterprises from the survey that have already participated in a previous survey, or are in another on-going survey, in order to spread the response burden more equitably across the survey population. This is sometimes referred to as giving the enterprises a *survey holiday*.

Example from Statistics South Africa (SSA)

The SSA live register has two parts:

- an administrative part, containing several sets of administrative units maintained from administrative sources; and
- a statistical part (called the business sampling frame), containing statistical units derived from administrative units and maintained from administrative data, survey feedback, profiling etc.

The SSA snapshot contains all statistical units as of the point in time it is created and included the values of all the characteristics that may be used for sampling. A snapshot is created, analysed and verified every quarter.

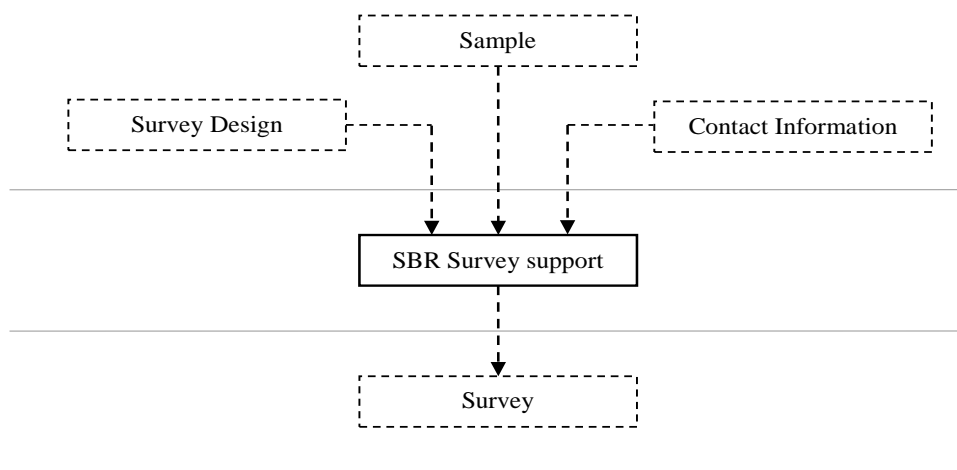
The SSA equivalent of a frozen frame is called the *Common Frame*. It is created, analysed and verified quarterly, but only used once per year as the basis for generation of survey frames. It contains all statistical units active with the characteristics (and only these characteristics) that are, or may be, used for sampling. Thus, statistical units that do not have valid values (industry code, size code, provincial code) for these sampling characteristics are excluded. Selected data are disseminated directly from the Common Frame.

Once per year the all survey frames are drawn afresh from the most recent Common Frame. Contact data for sampled units are subsequently extracted from the live register.

2.5 SBR Survey Support

The role of survey support is illustrated in Figure 2.6. In a typical enterprise survey, information is collected from *reporting units* who report on behalf of the *observation units* (which are statistical units) in the survey sample (as further discussed in Chapter 4). The answers to survey questions are recorded on questionnaires filled out by reporting units directly, or, less frequently, with the aid of an interviewer. To support this process, information on how to contact the reporting units and their relationships to the observation units is required. Whilst this information can be included in the Frozen Frame and hence in the survey frame, it is more efficiently extracted afresh from the live register for the sampled units only. Exactly what information is required depends on the mode(s) of data collection.

Figure 2.6: SBR Survey Support



Survey support is provided in three areas: data collection; survey registration and survey control.

Survey data collection

The primary activity in conducting a survey is data collection from reporting units. As previously noted, the information needed to contact the reporting unit depends on the mode used for data collection. In some cases a paper questionnaire is sent to the mailing address of the reporting unit. For a face-to-face interview a fieldworker has to visit the reporting unit to collect the information, so a physical address is needed. If the information is to be gathered

by telephone, a telephone number is required. It is also possible that a reporting unit may submit an electronic questionnaire by e-mail or via a web site, in which case an email address is required.

Survey registration

Units that are involved in different surveys, or on repeated occasions for the same survey, can be the subject of a policy to reduce and/or spread response burden. For example, it may be policy for any small enterprise to be observed only once every two years or to participate in at most one survey per year. Also, as previously noted, enterprises that have already participated may be given a *survey holiday*.

The survey registration process provides information about the observation units about which data are to be acquired and the reporting units that are to provide these data. The likelihood of contacting a reporting unit is improved when the most recent contact or communication information is available.

It is important to provide the reporting unit's preferred response mode for the survey. Response mode may influence response rate. Making available a range of modes (paper, telephone, face-to-face, web-based or other electronically) can improve response rates.

Survey control and response burden measurement

To help control the data collection process, reporting units should be monitored throughout the collection process. When frame errors are detected they should be recorded. The survey help-desk may be the mechanism by which information about errors is received.

In order to improve response rates, reminders are often sent. The administration of reminders can be a complicated process, especially as it depends not only on the official deadline for reporting specified by the survey but also on an estimate of the likely reaction time of the reporting units and/or of the quality of the answers they can provide. All attempts to contact reporting units have to be recorded, not only to ensure that follow-up is efficiently conducted, but to support subsequent enforcement of a response, as further illustrated by the following example from the Netherlands.

Response enforcement information in the Netherlands

Many countries have legislation requiring enterprises to respond to questionnaires from the NSI. In 2003, the Netherlands introduced a new statistical law. This law states that Statistics Netherlands (SN) is obliged to make maximum use of administrative data to derive its statistical output. Another of SN's obligations is to reduce the response burden on enterprises, for example by utilising administrative data sources. This implies that the units registered with these source need to be integrated into the SBR. Therefore SN has included fiscal information on the structure of enterprises (available from the finance department) in its SBR. However, not all information requested by European statistics by national agreements can be extracted from the administrative sources available. Therefore surveys in which data are directly obtained by questionnaire from a relatively small number of respondents are still needed.

By law SN has the right to require enterprises that do not respond to questionnaires to provide the requested information. In introductory letter that accompanies a survey, the enforcement procedure is explained. If an enterprise does not respond to an official questionnaire sent by the SN, then several (predefined) reminders are sent to the responsible contact person. If the enterprise refuses to respond, then SN has the right to fine the enterprise without any intervention of a third party. To assist this process a separate system has been created, containing all relevant SBR information and keeping track of the complete history of all attempts to contact the responsible person in the enterprise. This information can be used as evidence if the enterprise starts a juridical procedure against SN.

Reporting unit response rates (by mode in the case of multiple modes) and questionnaire item non-response rates are important in monitoring the quality of the survey outcomes (as further discussed in Chapter 10). Sample sizes and reporting unit response rates also used in calculating target and actual response burden.

Use of SBR for economic census

An economic census is a special type of survey in which the aim is to collect data from all statistical units in the target population. An SBR can be involved in variety of ways.

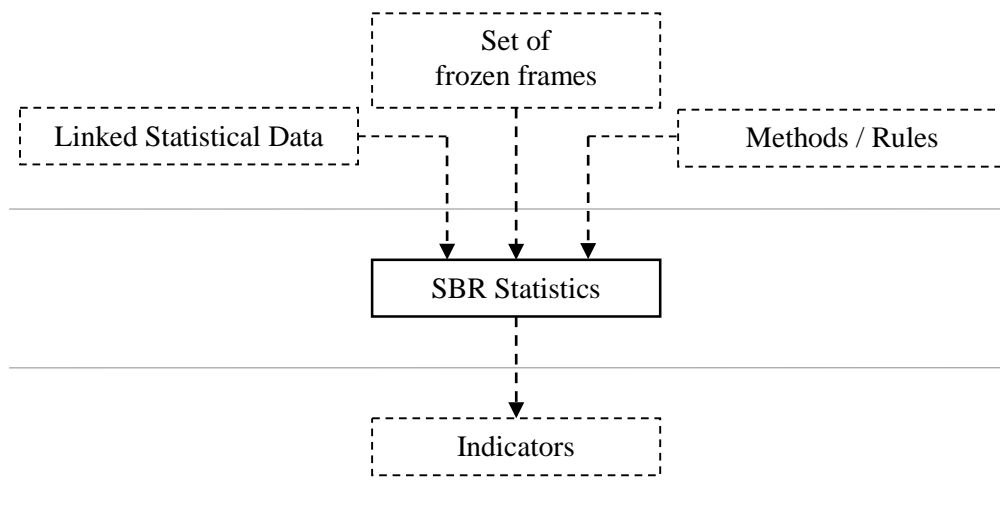
- ***SBR is not used.*** The economic census is conducted on the basis of comprehensive area enumeration. The resulting set of statistical units, whether local units or enterprises, may be used subsequently as the basis for creating or updating the SBR. This approach was used in the past but has become less and less common because (1) it requires extensive resources, (2) does not take advantage of existing lists of enterprises and data about them and (3) leaves behind a list of units that is impossible to maintain in its entirety because to do so would require an ongoing census.
- ***SBR provides initial frame.*** In this approach, which might be called a *register assisted census*, enumerators are provided with an initial list of statistical units (local units or enterprises) based on the SBR, and they correct and supplement this list. This approach makes better use of the SBR but does not assume the SBR is definitive.
- ***SBR provides the complete frame.*** In this approach the census is restricted to units that are in the SBR. Statistical data are collected about some or all of these units. For some units, data may be obtained from administrative sources instead.
- ***SBR provides the complete frame and links to administrative data.*** There is no statistical data collection process. The data are obtained entirely from administrative sources through the SBR. This may be termed a *register based census*. An example from Italy is described in Annex G1.

2.6 SBR Based Statistics

NSIs are confronted with increasing demands from governments, international bodies and researchers for business statistics on specific topics such as small areas or sectors of activity, and other particular aggregates. The information provided has to be consistent and comparable at country level, internationally, and for any other geographical area, also across different economic activities. At the same time NSIs are under pressure to reduce the data collection costs and response burden.

As illustrated in Figure 2.7, the SBR can play a key role in providing more information at less cost through direct use of SBR data. It can be a primary source of information on the structure and the demography of the population of enterprises. Also, in combination with other administrative registers or statistical data sources, it can produce additional statistical information for economic analysis.

Figure 2.7: SBR Statistics



Enterprise demography statistics

Data on births and deaths of enterprises, their survival rates and the role they play in economic growth and productivity as well as data for tackling social demographic issues are increasingly requested both by policy makers and analysts. As further discussed in Section 9.3, there are requirements for:

- data for sub-populations of small and medium size enterprises;
- historical data about enterprises; and
- identifying subpopulations of interest: for example, enterprises with high growth, or with large increases in numbers of employees.

Also, when calculating ratios, there needs to be conceptual consistency between the denominator and numerator populations.

The SBR can satisfy all these requirements because it covers the whole population of enterprises. Thus it can provide regional and small area data, basic characteristics like economic activity code, employment and/or turnover, and data on births and deaths. In summary it is an ideal source in terms of coverage and cost for the production and dissemination of enterprise demographic statistics.

However, use of the SBR for producing statistics raises similar issues regarding treatment of data over time as does its use in providing survey frames. The live register is constantly being refreshed. The updates applied over the time period (t) to (t+1) represent not only the actual economic changes that occurred during that period, but also adjustments in the coverage or characteristics of units that resulted from SBR maintenance procedures and that had nothing to do with changes in the real world. For example, based on the results of an SBR improvement survey during the time period, (t) to (t+1) economic activity classifications of enterprises, or their measures of size, or the dates on which they ceased or commenced activity, may be corrected from their former out-of-date, erroneous or missing values. In summary, SBR data cannot be simply regarded as the result of a point in time statistical survey, which collects information about the status of an enterprise at time (t) or changes in the enterprise over the period (t) to (t+1).

It may be advisable for statistics derived from SBR data to be based on a *satellite approach*. If parts, or the whole, of one or more frozen frames are extracted from the SBR and linked to data from other sources and if the resulting information is maintained outside and independent of the SBR environment, then the product is referred to as a *satellite*. A satellite approach (as further elaborated in Section 2.9.2) is a good way of addressing issues that may arise from corrections of classifications or reference periods as the base population can be well defined and coordinated. There may also be an organisational advantage to a satellite approach in the sense that the task of producing SBR based statistics from the satellite can be assigned to the organisational unit responsible for economic data production and dissemination rather resting with the SBR unit.

Integration with external trade register

External trade statistics do not present any explicit information on the characteristics of enterprises that undertake the trading. The concepts and classifications in external trade statistics differ from those in production statistics. Knowledge of a trading enterprise's profile - economic activity, size, location, and whether or not it belongs to an enterprise group - is important in the analysis of globalisation of the national economy and the determinants of this internationalisation. Coherent compilation of trade statistics by enterprise characteristics requires linkage of trade and business registers at micro level. If this can be achieved, the combination of the key enterprise characteristics and the trade data, such as product code and partner country, offers many opportunities for producing a more complete and diversified view of the structure of both trade and production.

The role of the SBR in the System of Environmental-Economic Accounting (SEEA)

The System of Environmental-Economic Accounting (SEEA) is an international statistical system that relates economic and environmental information in a common framework to measure the contribution of the environment to the economy and the impact of the economy on the environment. The SEEA is a satellite accounting system to the System of National Accounts which uses similar concepts and classifications (e.g. for economic activities and products) to those employed in the SNA. The SEEA comprises indicators on environmental topics like supply and use of energy and water, air pollution water emissions, solid waste, oil- and gas reserves, environmental investments, -subsidies and -protection expenditure, and so on. The results are also main input for publication about *Green Growth*.

The SBR supports these fields of interests by providing coordinated populations of statistical units in order to derive aggregates by industry which are then automatically related to the aggregates of an economic indicator related to the corresponding population of industries used by NA. Therefore high quality of the activity classification (ISIC, NACE) and the linkage of the SBR with other data sources are of great importance.

The enterprises needed to describe the Environmental Goods and Service Sector (EGSS) are stored in the SBR. To measure the developments in this sector it is necessary that the demographic events of the statistical units involved are uniquely defined in the SBR and that those units are traceable over time. In the SBR this is supported by the continuation-algorithm for statistical units.

2.7 SBR Information Source

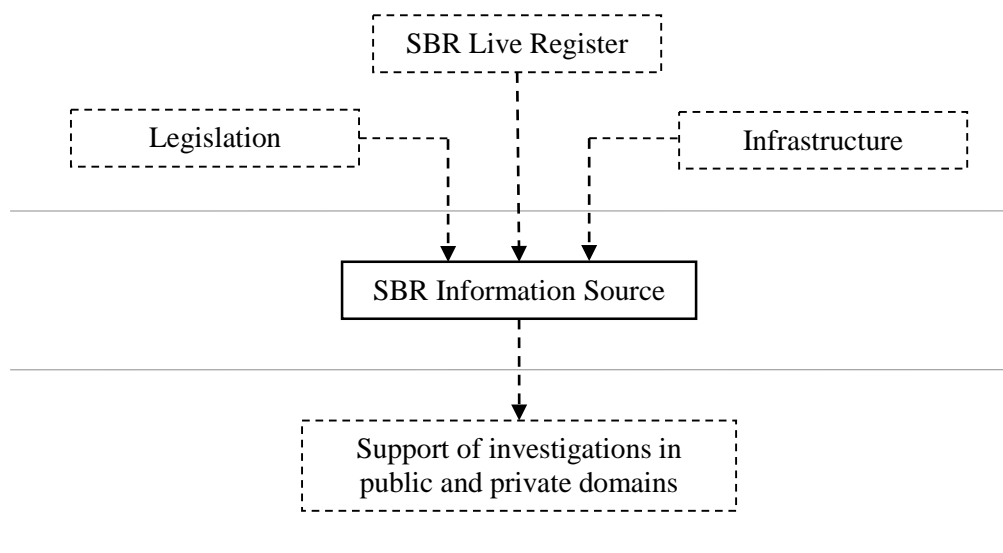
The roles discussed in Sections 2.5- 2.6 concern production and internal and external use of statistics (i.e. macro-data) based on SBR data. This section concentrates on the provision of SBR micro (individual) data as *open data*, i.e. the data that are publicly available for any individuals or organisations to use and that are licensed in a way that allows for reuse¹. Access may be made available to:

- *lists of enterprises* – for the use of other government agencies and market research companies so that they can conduct surveys;
- *ownership and control information* - “who-owns-who”;
- *information to support investigations* into government program performance or to support preparation or monitoring of local, regional or national economic development programs.

Furthermore provision of data may be through a geographical interface depending upon the availability of suitable geo-statistical infrastructure.

As illustrated in Figure 2.8, in determining the scope of this role, the NSI is likely to be significantly constrained by the confidentiality and privacy provisions embedded in the laws and regulations governing the dissemination of statistics as discussed in a following subsection. A supportive legal framework can open up the possibility of providing external users with access to business data that are already in the public domain, such as: trade name, ownership, economic activity, size, geographic location and contact data, as exemplified by the Mexican DENU system that is referenced in an example below and in Chapter 9.

Figure 2.8: SBR Information Source



¹ <http://www.theguardian.com/public-leaders-network/2014/apr/15/big-data-open-data-transform-government>

SBR regulatory framework

SBR data must be handled in accordance with the regulatory framework that governs the NSI. On the one hand privacy and confidentiality provisions under the statistical legislation have to be observed. On the other hand much of the data about enterprises in the SBR is information that the enterprises themselves make public and/or is generally considered (within the legal framework in a country) to be of public interest, i.e., information to which any citizen should have access.

In ideal circumstances a regulatory framework would allow dissemination of the following SBR data:

- identification and stratification characteristics: name of enterprise, denomination or corporate name, code and name of the economic activity class, size code based on number of personnel;
- geographical location characteristics: street, external and internal number, neighbourhood, postal/zip code, locality;
- geographical coordinates of the location: latitude and longitude;
- contact characteristics: phone, fax, e-mail, and web-page; and
- date of birth of unit in the SBR.

In practice, the ability to provide this information may require review and revision of laws, regulations, policies and standards governing both the NSI and its administrative sources.

Geo-statistical information and SBR data consultancy systems

In recent years, there has been a growing interest in the geo-spatial analysis. This interest can be attributed to significant increases in the use of geographic information systems (GIS), characterised by the capacity to integrate and use a wide variety of spatial information. Individual address records have become the standard level for spatial investigation in many socioeconomic and planning applications. Because of the increased level of user friendliness and accessibility of GIS packages, geographical coordinates can be assigned to the each address in the SBR. The geospatial dimension is important in building a geographic infrastructure data system for geospatial analysis of economic structure.

Ideally, the SBR should be able to provide information on the distribution of economic activity and enterprises/establishments by geography. A GIS can present this information as geo-statistical cartography. The use of GIS is fundamental in providing users with the necessary data to analyse economic activity in relation to geography. Through a GIS, it is possible to incorporate the many information layers as geo-statistical data, including settlements, street axes, block fronts, locality polygons, blocks, external numbers, urban services, reference elements, natural resources, geographical names, hydro graphic networks, routes of communication, territory images, relief, limits references, geodesic references.

Geo-statistical information in Mexico

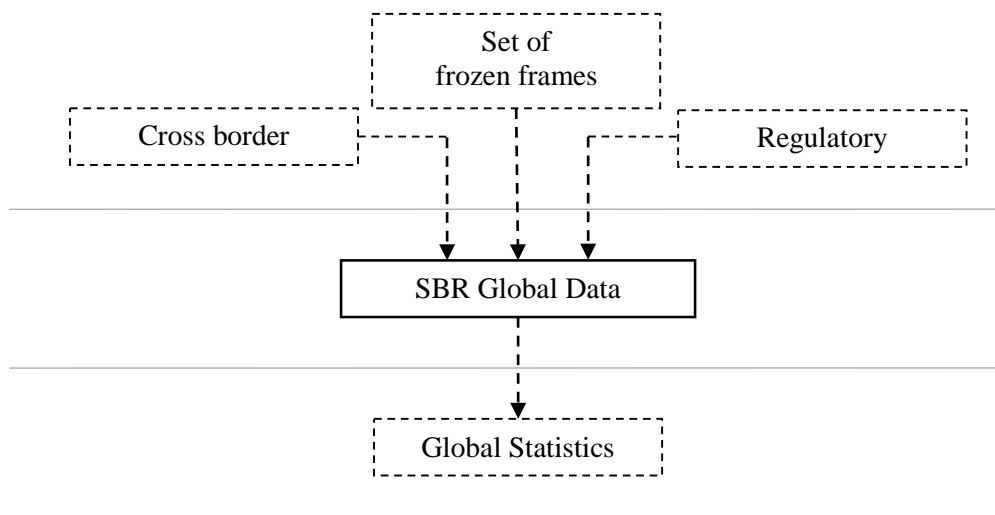
In Mexico, the *National Directory of Economic Units (DENUE)* is a part of the SBR that is available to any user through a free consultation system in the *National Institute of Statistics and Geography* (<http://www3.inegi.org.mx/sistemas/mapa/denue/default.aspx>). Its geo-statistical features have broadened its role as a tool supporting public and private policy developers and decision makers as well as professors and researchers.

2.8 SBR International Data Exchange

In a society that is more and more influenced by globalisation, international comparability of economic statistics is becoming increasingly important. Internationally defined indicators based on internationally harmonised populations of statistical units are desirable because the resulting statistics are more comparable when measuring globalisation effects. At the same time it is important for producers and users of statistics to understand the potential limitations of the SBR in the context of international comparability.

Much effort is spent in the SBR to guarantee coherence, in terms of concepts, methods and quality within an NSI and between NSIs. In the last few decades, thanks to the activities of supra-national and international organisations such as Eurostat, OECD, and UNECE, worldwide coherence in the definition and treatment of enterprises has increased considerably. This is important as SBR data are used in studies comparing (the structure of) different national economies and other international studies, as illustrated in Figure 2.9.

Figure 2.9: SBR Global Data Exchange



Key factors in improving the quality of global statistics are the exchange and coordination of micro-data between countries and the central accessibility of these data, for example, storing them in one central database or at least accessing them through a single gateway. In this context an SBR is a potential data provider and a network of SBRs can exist.

Data exchange is not only important for global statistics, but also in achieving consistency in statistics dealing with cross-border issues in different countries. For example a subsidiary of an enterprise in country X should in country Y be covered by inward FATS, and in country X by outward FATS.

A complex question for an NSI is the legal basis for sharing micro-data with statistical authorities of other countries or international organisations. International standards regarding the legal basis for micro-data sharing yet to be established. Exchange of data on global units can be seriously hampered by the requirements that, first, data are used exclusively for statistical purposes and, second, disclosure of data about individual statistical units is forbidden. The capacity to transfer the data to another country depends upon confidentiality requirements within the sending country. The receiving country is responsible for satisfying whatever conditions are imposed by the sending country and its own confidentiality laws.

Thus, such exchanges require close cooperation between the partner countries, comparable rules for the use and protection of confidential data, a solid agreement between the countries, and, -above all, trust that data are not misused.

Euro Group Register (EGR)

The EGR is a central business register maintained by Eurostat that supports production of micro based statistics on globalisation in Europe, both at country and European level. National data from various commercial and institutional sources are integrated into the EGR without questioning their quality. The EGR thus constitutes, in effect, an international network of registers. It provides access to integrated and up-to-date data on multinational enterprise groups that have statistically relevant (financial or non-financial) transnational operations in at least one of the European countries involved in the project.

European System of Business Registers (ESBRs) Project

The ESBRs project (2013-2017) is rationalising, strengthening and standardising all SBRs in the ESS with the ultimate aim of integration within an interoperable system. It is a continuation of the EGR Project and the ESS International Profiling Project. A core goal is improvement of the EGR so that it can provide better quality information on multinational enterprises (MNEs), with on-line access by NSIs. It incorporates a new model for statistical units and profiling that enables views of MNE groups in their entirety, colloquially referred to as *seeing the whole elephant*.

European Statistical System Network (ESSnet): Exchange of trade micro-data

Experts within the ESS EGR and International Profiling Projects are developing methodological guidelines for derivation of coordinated populations to support Foreign Trade Statistics and Statistics on Foreign Direct Investments at European level. These guidelines are a starting point for a worldwide discussion on the measurement of global indicators to describe globalisation effects and may well influence future exchange of data from SBRs.

2.9 SBR in Modernisation of Statistical Production and Services

Keywords in current attempts to modernise statistics are efficiency, coherence, interoperability and cooperation. In this context, the Generic Statistical Business Process Model (GSBPM) is a standard but flexible tool for describing and defining the common set of common business processes that typically constitute statistical production. It is being used in developing a common business reference architecture that has the aim of obtaining more coherence in official statistics. As mentioned before the SBR fulfils a backbone role in this respect by integrating several (mostly administrative) data sources and supporting the collection of economic data. This backbone role can be strengthened when the use of the SBR is seen to support the business processes as described in the GSBPM (as further elaborated in Chapter 8).

GSBPM Sub-process 2.4. Design frame and sample

This sub-process only applies to processes which involve data collection based on sampling, such as through statistical surveys. It identifies and specifies the population of interest, defines a sampling frame (and, where necessary, the register from which it is derived), and determines the most appropriate sampling criteria and methodology (which

could include complete enumeration). Common sources for a sampling frame are administrative and statistical registers, censuses and information from other sample surveys. This sub-process describes how these sources can be combined if needed. Analysis of whether the frame covers the target population should be performed. A sampling plan should be made.

GSBPM Sub-process 4.1. Create frame and select sample

This sub-process establishes the frame and selects the sample for this iteration of the collection, as specified in sub-process. It also includes the coordination of samples between instances of the same statistical business process (for example to manage overlap or rotation), and between different processes using a common frame or register (for example to manage overlap or to spread response burden). Quality assurance and approval of the frame and the selected sample are also undertaken in this sub-process, though maintenance of underlying registers, from which frames for several statistical business processes are drawn, is treated as a separate business process. The sampling aspect of this sub-process is not usually relevant for processes based entirely on the use of pre-existing sources (e.g. administrative sources) as such processes generally create frames from the available data and then follow a census approach.

The SBR can be linked to sources other than the administrative. This can be achieved by storing the identification numbers of units in other sources in the live register or by integrating special modules to the SBR to be used for matching with other sources. In this case the statistical units in the SBR can be linked to the units in these other sources, which means their data are available to assist in maintaining the units in the SBR and/or for use by surveys. To improve interoperability, the SBR may use web or other services developed by other countries for SBR maintenance, and it may create the conditions for exchange of coordinated statistical information among countries, supported by an international identification service.

2.10 Satellite Approach to Extend Functionality

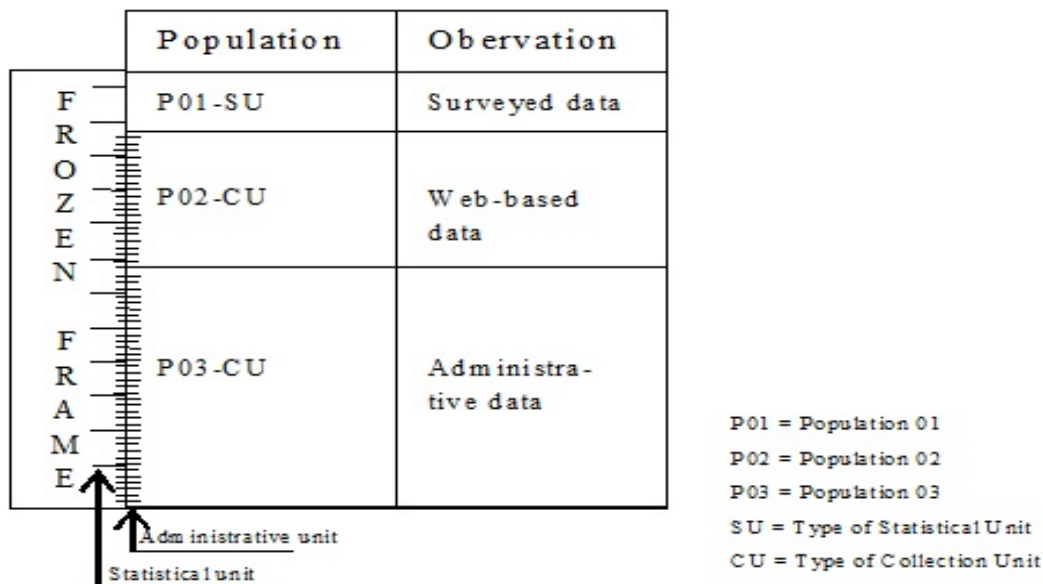
If all the roles described in the previous sections are supported by the SBR, the result may be quite a complex network of databases and functionalities. An approach that can be used to extend the functionality with minimum complication is to extract parts or the whole of one or more frozen frames from 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. Whilst this can adversely affect the coordinating role of the SBR, the big advantage is that much more additional information can be stored and managed without interfering with the basic functions of the SBR. Satellites can, for instance, support international trade statistics, or cooperation with central banks. In such cases anomalies that might result from integration within the SBR environment can be avoided.

The SBR may support the satellite approach in different ways. If the units available in one or more frozen frames are linked to data from sources other than those used in the SBR, then the combined result may be a semi-manufactured statistical product blending statistical and administrative data. The difference from a sample is that, in this kind of satellite, information is linked to all units in the underlying population. This approach is useful, for example, to provide economic indicators on self-employed (income-estimation) or to compile the regional distribution of production factors to support regional accounting information..

By linking information collected by different methods, for example by survey and from an administrative source to the units in one or more frozen frames, a satellite is useful in splitting off and controlling domains where secondary data collection based on administrative sources is not possible or is less efficient, as illustrated in Figure 2.10. This includes the traditional approach of integrating administrative data for small enterprises with data collected by survey from large enterprises.

Figure 2.10.



It is also possible to analyse and even compensate for over- or under- coverage of units in the live register by comparisons with other sources. Instead of including the other sources in the live register, it may be more convenient (and less complex) to maintain a satellite outside the SBR environment. This kind of approach may be appropriate for units in the governmental sector (S13), or units that belong to the financial sector (S12), or foreign units that belong to the sector representing the rest of the world (S2).

Another possible circumstance in which to set up a satellite is where different statistical units from those recorded in the live register are used to describe a phenomenon. There may be a case for maintaining descriptions of such units in a satellite to ease the workload on SBR staff and reduce the complexity of the SBR system, even though additional resources may be required to maintain two separate systems. For example, to produce tourism statistics, data are collected on the capacity and the occupancy of accommodation establishments used by tourists, for example, nights spent, arrivals, occupancy rates, number of bedrooms. For this purpose the data collection at the level of an enterprise unit with a main activity in the tourism groups would be insufficient as many tourist accommodation establishments are provided by enterprises or institutes with another main activity offering accommodation as a secondary activity. For this reason a special satellite linked to the SBR that records tourism establishments may be a better solution than attempting to extend the functionality of the SBR itself to include such units.

In conclusion, it must be reiterated that the main reason for setting up a satellite is to simplify otherwise complicated statistical processes by splitting them off from the SBR, with the aim of simplifying control and/or spreading responsibility.

3 Coverage of SBR

3.1 Introduction

This chapter describes the coverage of an SBR and examines the concept from the perspective of the various institutional unit types and sectors of the economy, and it flags key issues that NSIs need to consider.

In principle, an SBR should record all institutional units in the national economy that are engaged in productive economic activities, i.e., activities contributing to the gross domestic product (GDP). In accordance with System of National Accounts 2008 (SNA2008), the term “*economic production*” is understood as referring to “*a process carried out under the control and responsibility of an institutional unit that uses labour, capital, goods and services to produce outputs of goods and services*”², and institutional units engaged in production are referred to as *enterprises*

This chapter examines the three key aspects of SBR coverage:

- “*Completeness*” is the extent to which the SBR includes all institutional units within the SNA2008 production boundary.
- “*Coverage*” is the proportion of total national economic production that the units represent.
- “*Content*” is the set of characteristics (e.g., institutional sector, size, location and registration status) of the units contained in the SBR.

An SBR can be said to provide complete coverage and content if it includes all units producing goods and services together with all their required characteristics. In reality, complete coverage of all these units within the SNA production boundary is impossible to achieve. However, for the purposes of international comparisons, it is desirable that the coverage of an SBR should meet agreed standards, with completeness being the ultimate aim.

During any period units commence and cease operations. Thus SBRs are regularly refreshed, new units are created, and units that have ceased economic activities are deactivated. Construction and maintenance of the SBR are described in Chapters 6 and 7.

The SBR provides frozen frames, each of which contains all units in the live register for a given reference period. Frames for economic surveys are derived from frozen frames. Thus the coverage of a SBR for a *given reference period for a given survey* can be defined as the extent to which the survey frame from the SBR represents the survey target population.

This chapter presents considerations that are pertinent in formulating the SBR coverage and content decisions and it discusses specific issues for various segments into which the SBR may be divided. It is based on concepts articulated in the *SNA2008*, *ISIC Rev 4*, Eurostat’s *Business Register Recommendations Manual*, the African Development Bank’s *Guidelines for Building SBRs in Africa* and other sources as footnoted and listed in Annex A.

² SNA2008: 1.40

3.2 Key Concepts

3.2.1 *Production Boundaries*

The concept of *production* is defined in the SNA2008 together with descriptions of the production boundaries, which clarify the activities constituting production. The SNA2008 includes two types of production boundary: the general production boundary and the SNA production boundary.

The *general production boundary* defines production as an activity carried out under the control of and under the responsibility of an institutional unit that uses inputs of labour, capital and goods and services to produce outputs of goods or services. There must be an institutional unit that assumes responsibility for the process of production and owns any resulting goods or knowledge-capturing products, or is entitled to be paid or otherwise compensated, for the change-effecting or margin services provided³.

The *SNA production boundary* is a more restricted version of the general production boundary. It excludes activities undertaken by households that produce services for their own use, except for services provided by owner-occupied dwellings and services produced by employing paid domestic staff⁴.

3.2.2 *Market and Non-Market Producers*

The distinction between market and non-market production is critical for from the perspective of SNA2008 and hence important for the SBR. *Market producers* are enterprises that sell all or most of their output at prices that are economically significant⁵. Prices are said to be *economically significant* if they have a significant effect on the amount that producers are willing to supply and the amounts purchasers wish to buy. Market producers make decisions about what to produce and how much to produce in response to expected levels of demand and expected costs of supply and are exposed to the risks associated with this production. They adjust supply either with the goal of making a profit in the long run or, at a minimum, covering capital and other costs.

The core coverage of the SBR comprises units that are market producers and whose principal activity is the production of goods and services. Registered market producers are very important from an economic point of view and in principle are not difficult to cover based on available administrative registers.

Non-market producers

Non-market producers are not likely to respond to changes in economic conditions in the same way as market producers. Their economic behaviour is influenced by the receipt of material financial support in the form of transfers such as taxes, grants and donations. Government units and non-profit institutions that provide most of their services at prices that are not economically significant prices are considered non-market producers.

The recommended SBR treatment of non-market producers is covered in the discussion on government and non-profit institutions.

³ SNA2008: 6.24

⁴ SNA2008: 6.26

⁵ SNA2008:4.88

3.2.3 *Non-Observed and Observed Economies*

As noted in *Measuring the Non-Observed Economy A Handbook* (OECD et al., 2002), complete coverage of economic production is a vital aspect of the quality of the national accounts. However, it is hard to achieve because of the difficulties in recording certain types of productive activities. Activities that are missing from the basic data used to compile the national accounts are said to comprise the *non-observed economy (NOE)*. Productive activities may be non-observed because they are *informal, household production for own final use, illegal, underground, or due to deficiencies in the basic data collection system*. Making estimates for them in the national accounts is referred to as *measurement of the NOE*.

It is difficult to identify the enterprises that are part of the non-observed economy and thus they are usually defined to be outside the planned coverage of the SBR. Thus, their economic activities have to be measured in some way other than by a survey based on the SBR.

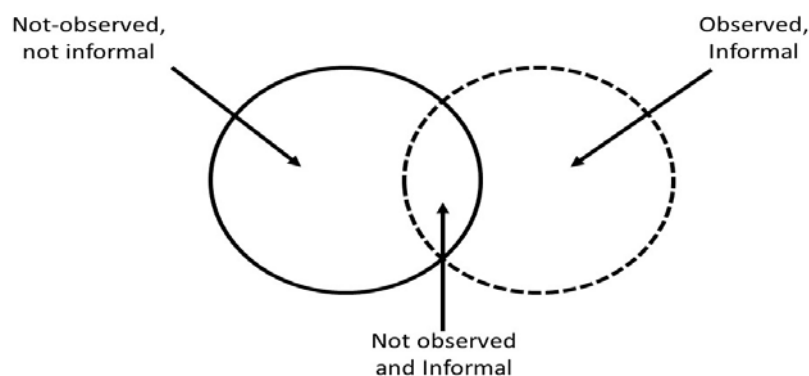
3.2.4 *Informal Sector*

A national economy can be divided into three parts: the formal sector, the informal sector, and household non-market production for own final use. The major portion of the difference between the total economic production of households and what is readily measured by enterprise surveys and administrative sources is the *informal economy*, i.e., production by enterprises that are considered to be in the informal sector as defined below.

The relationship between the non-observed economy and the informal sector is illustrated in Figure 3.1. The solid circle represents the non-observed economy and the dotted circle the informal sector. The relative size of the three segments shown in the figure varies from country to country⁶.

The informal sector represents an important part of the economy in many countries, playing a major role in production, income generation, and employment creation.

Figure 3.1: Non-Observed economy and Informal Sector⁷



The informal sector has been defined by the Fifteenth International Conference of Labour Statisticians in its resolution concerning statistics of employment in the informal sector

⁶ SNA2008: 25.1

⁷ SNA2008: Figure 25.1

organised by the International Labour Organization (ILO) in 1993. A resolution of the Conference (abbreviated ICLS93) describes the informal sector in broad terms reflecting the types of production units of which it is composed. It also provides a framework within which each country must formulate its own specific operational definition. Discussion of the informal sector in the SNA2008⁸ is consistent with the resolution.

The informal sector is broadly characterized as consisting of units engaged in the production of goods or services with the primary objective of generating employment and incomes for the persons involved in the production. These units typically operate at a low level of organization, with little or no division between labour and capital as factors of production and on a small scale. Labour relations – where they exist – are based mostly on casual employment, kinship, or personal and social relations rather than contractual arrangements with formal guarantees⁹. The informal sector thus defined excludes households producing exclusively for own final use.

More specifically, the ICLS93 framework requires that the informal sector be defined as a subset of household unincorporated enterprises, comprising those enterprises that:

- produce at least some output for the market; and
- are less than a specified size in terms of the number of persons engaged, or of employees or of employees employed on a continuous basis; and/or
- are not registered under specific forms of national legislation, such as factories or commercial acts, tax or social security laws, professional groups' regulatory acts, or similar acts, laws or regulations established by national legislative bodies.

These criteria provide a framework for a definition of the informal sector, *not a single, unambiguous definition*. Thus, there are variations across countries in how the informal sector is actually defined. The units included in the informal sector depend upon whether registration and/or number of employees is/are used as a criterion, and, in the latter case, on the size boundary for number of employees. It also depends upon whether or not agricultural units are included. In fact the ICLS93 recommends that household unincorporated enterprises that are involved in agricultural activities (ISIC section A) are measured separately from other economic activities to ensure international comparability and to facilitate the selection and application of appropriate statistical data collection tools and sample design.¹⁰ They are referred to in these Guidelines as *informal agriculture*.

Figure 3.2 shows how the three economic sectors of the economy relate to the institutional sectors defined in the SNA2008 and which household unincorporated enterprises are considered part of the informal sector.

The informal sector definition recommended in these Guidelines is the set of *non-agricultural household enterprises that have market production but that are not registered with the administrative source(s) on which the SBR is based*. This definition is in accordance with the ICLS93 framework and makes a very visible formal/informal boundary. It means that:

- the suite of economic surveys based on the SBR and on administrative data measure the formal sector; and

⁸ SNA2008: 25.1—25.27

⁹ ICLS93: 5(1)

¹⁰ SNA2008 25.46

- household market enterprises not included in the SBR constitute the informal sector and informal agriculture.

Figure 3.2: Identifying Units in the ILO Informal Sector

Institutional Sector	Sector of Economy		
	Formal	Informal*	Household Non-Market
Non-Financial and Financial Corporations	All	-	-
General Government	All	-	-
Non-Profit Organisations Serving Households (NPISH)	All	-	-
Households	Household unincorporated enterprise that are registered or have more than a given number of employees	Without employees (<i>informal own-account enterprises</i>)	Household unincorporated enterprises undertaking production only for own final use (including owner occupation of dwellings)
		With employees (<i>enterprises of informal employers</i>)	
*May be divided into <i>informal</i> and <i>informal agriculture</i> .			

3.2.5 Illegal Activities

There are two kinds of illegal production:

- production of goods or services whose sale, distribution or possession is forbidden by law;
- production activities that are usually legal but become illegal when carried out by unauthorized producers, for example, unlicensed medical practitioners.

Examples of activities that may be illegal but productive in an economic sense include the manufacture and distribution of narcotics, illegal transportation in the form of smuggling of goods and of people, and services such as prostitution.¹¹ Both kinds of illegal production are included within the production boundary of the SNA provided they are genuine production

¹¹ SNA2008: 6.43-44

processes whose outputs consist of goods or services for which there is an effective market demand.¹²

3.2.6 *Underground Activities*

Certain activities may clearly fall within the production boundary of the SNA and be quite legal (provided certain standards or regulations are complied with) but deliberately concealed from public authorities for the following kinds of reasons:

- to avoid the payment of income, value added or other taxes;
- to avoid the payment of social security contributions;
- to avoid having to meet certain legal standards such as minimum wages, maximum hours, safety or health standards, etc.;
- to avoid complying with certain administrative procedures, such as completing statistical questionnaires or other administrative forms.¹³

These are referred to as the underground activities.¹⁴

3.3 SBR Coverage

3.3.1 *Type of Institutional Units*

The SNA production boundary covers the productive activities undertaken by all institutional units. Thus NSIs should aim to cover as much national production as possible in their SBRs by including all types of institutional units engaged in production. However, as previously noted, complete coverage can never be attained in practice, and a more realistic aim is for coverage to meet agreed standards.

Groups of units should not be omitted without an assessment of their economic importance. SBRs should aim to record, at a minimum, all active economic units in the formal economy that are engaged in economic activities contributing to the gross domestic product (GDP).

Chapter 4 of the SNA2008 describes five types of institutional units which engage in production activity. These are *non-financial corporations*, *financial corporations*, *government units*, *households* and *non-profit institutions serving households*. As previously noted, in their role as economic producers, the SNA2008 refers to these institutional units as “enterprises”. The following paragraphs contain detailed discussion on the coverage of these various types of units in the SBR.

3.3.2 *Government Units*

SNA2008 describes *government units* as unique kinds of legal entities established by political processes that have legislative, judicial or executive authority over other institutional units within a given area¹⁵.

¹² SNA2008: 6.45

¹³ SNA2008: 6.40

¹⁴ Handbook for Measurement of the Non-Observed Economy 3.2.2

¹⁵ SNA2008 4.9

Data for general government units are often obtained directly from a single source within the government rather than requiring direct survey of individual government ministries, departments, and agencies.

The advantage of inclusion of government units in the SBR is that it enables a single system to provide frames for all surveys. Appropriate reporting arrangements can be recorded in the SBR to indicate whether or not the data for a particular survey can be collected from a single centrally located government source rather from the individual government units.

The identification of statistical units for government can vary from country to country. The same general principles can be used for determining the appropriate statistical units as for other large complex enterprises, as further discussed in Chapter 7.

3.3.3 Corporations

In the SNA2008 the term *corporation* covers legally constituted corporations and also cooperatives, limited liability partnerships, notional resident units and quasi corporations¹⁶.

Units that are market producers and whose principal activity is the production of goods and services constitute the core coverage of the SBR. The subset of these units comprising all registered profit oriented *corporations* is the most important one an economic point of view. Their inclusion in the SBR causes, in principle, no coverage difficulty since sources that identify corporations are usually readily available.

3.3.4 Non-Profit Institutions

In the SNA2008 non-profit institutions (NPIs) are legal or social entities created for the purpose of producing goods and services but whose status does not permit them to be a source of income, profit or other financial gain for the units that establish, control or finance them¹⁷.

Ideally all NPIs should be included in the SBR. However, some NPIs may not be included in the administrative sources used to maintain the SBR, for example because they are below employment and turnover thresholds.

3.3.5 Households

Based on the SNA2008 households involved in some form of economic production may be referred to as *household (unincorporated) enterprises*. As indicated in Figure 3.2, they may be divided into four categories:

- Household enterprises that are defined to belong to the formal sector, i.e. are registered, or have more than a given number of employees;
- Household enterprises without employees belonging to the informal sector (*informal own account enterprises*);
- Household enterprises with employees in the informal sector (*enterprises of informal employers*); and
- Household enterprises undertaking production only for own final use, including owner occupation of dwellings.

¹⁶ SNA2008: 4.7

¹⁷ SNA2008: 4.8

Ideally the SBR should include all household enterprises which sell or trade their output. In practice it is impossible to identify administrative sources that cover all such small units. Thus the recommended target coverage for each category of household enterprise is shown in the Figure 3.3. The shaded areas represent units for which coverage is desirable but for which no suitable source may be available.

Figure 3.3: Identifying institutional units in the ILO informal sector and their recommended coverage on the SBR¹⁸

Non-financial and financial corporations		Households							
Those that are registered or with greater than a given number of employees	Those that are not registered or with fewer than a given number of employees	Those that are registered or with greater than a given number of employees	Self-employed (Informal own account enterprises)				Unincorporated enterprise with employees (enterprise of informal employers)		Institutional households, households with no unincorporated enterprises, households only undertaking production for own final use (including owner occupation of dwellings)
			Market producers		Producers for own final use		Market producer	Producers for own final use	
			Self-employed professionals	Selling most or all production	Selling some production	Not selling any production	Selling most or all production	Selling some production	Not selling any production
Include in SBR	Include in SBR	Include in SBR	Include in SBR	Desirable to include in SBR	Do not include in SBR	Desirable to include on SBR	Do not include in SBR		

Enterprises that comprise the economic activities of households that are employers of domestic and other personnel should be included in the SBR if they are registered as employers. Their inclusion allows the SBR to be used in connection with employee or household-based surveys such as the labour force survey where employees of these units are likely to be included.

Some household enterprises are lessors, involved only in letting and/or operating self-owned or leased real estate. If they are not registered as employers they should be excluded from the SBR, first because they are part of the informal sector, and second because they occur in large numbers and would inflate SBR unnecessarily. If an enterprise's letting activities move to a larger scale, the enterprise may well decide to incorporate, in which case the resulting enterprise should, of course, be included in the SBR. It is recognised that in some countries the difference between household and incorporated lessors may be difficult to determine based on the administrative sources used in creating and maintaining the SBR. In this case, the NSI may be compelled to include household lessors (or some of them) in the SBR.

¹⁸ From SNA2008: Figure 25.4

3.3.6 Self-employed Professionals

Self-employed professionals who are non-employers, i.e. one person enterprises, are frequently under-covered but tend to play a very important role in several ISIC sections (e.g. section M and class 9000). If they are registered with an administrative source they should be included in the SBR. Typical administrative sources are social security files, income tax files, VAT files, files of chambers of commerce and of crafts, government registration files, and lists of doctors. Those that are not registered are part of the self-employed (informal own account enterprises) column in Figure 3.3.

Some countries may also, or as well, use area enumeration, for example during the course of an economic census, to detect these units, but this approach is not recommended as a source of SBR coverage as it is typically only periodic.

3.3.7 Own-use Production

The SNA2008¹⁹ specifies that the household production boundary includes the production of goods but excludes the production of most services for own use. The exceptions to the exclusion are services provided by owner-occupied dwellings and services produced by employing paid domestic staff.

However, household enterprises should not be included in the SBR if the goods and services they produce are for their own consumption, e.g. production from domestic gardens, or invested in the household, e.g. do-it-yourself activities, as there is no reliable means of maintaining such a list. Thus the SBR has no role to play in the measurement of household production for own final use. Data about these productive activities are best obtained through surveys of households, not of enterprises.

3.4 Special Cases

3.4.1 Agricultural Holdings

In this context, *agriculture* refers to activities defined in accordance with ISIC section A.

It is strongly recommended that agricultural holdings should be included in the SBR provided they are registered. Some countries manage a farm (possibly including forestry) register independently from the SBR. This practice can lead to inconsistency in coverage across economic surveys, as some units may be included in both the SBR and the farm register and others may be excluded from both. It may also result in less effective and efficient updating routines, and greater response burden for the agricultural holdings.

The inclusion of agricultural holdings is not only an SBR objective but it is also an important step towards the integration of agricultural statistics into a comprehensive and coherent system of economic statistics, which is highly desirable, especially for the purposes of national accounts.

¹⁹ SNA2008: 6.28-6.48

3.4.2 *Illegal and Underground Production*

SNA 2008 explicitly states that productive illegal activities should be included in national accounts²⁰. Illegal actions that fit the characteristics of transactions (notably the characteristic that there is mutual agreement between the parties) are treated the same way as legal actions. The production or consumption of certain goods or services, such as narcotics, may be illegal but market transactions in such goods and services have to be recorded in the accounts. If expenditures on illegal goods or services by households were to be ignored on the grounds of principle, household saving would be overestimated and households presumed to obtain assets that they do not in fact acquire. Clearly, the accounts as a whole are liable to be seriously distorted if monetary transactions that in fact take place are excluded. It may be difficult, or even impossible, to obtain data about illegal transactions, but in principle they should be included in the accounts if only to reduce error in other items, including balancing items²¹.

Some activities, for example the production and distribution of alcohol, or prostitution, may be illegal in one country but legal in another. Exclusion of illegal production would thus distort international comparisons. Likewise, it would give rise to distortions over time if some activities switch from being illegal to be legal, or the converse.

However, although there is general agreement on the correctness, in principle, of including illegal activities in the national accounts, *in some countries, the difficulties of capturing illegal activities may mean that they are either not well covered or deliberately ignored on pragmatic grounds*²². Thus, only a relatively small number of countries explicitly include illegal activities at the present time, and the methods available are still experimental. In particular, since 2010 the EU member states are required to include illegal activities in their accounts.

In summary, in countries where the value of illegal productive activities is quantitatively insignificant, it is almost certainly a poor use of resources to try to measure them precisely. In any case, whether or not illegal activities are actually included, an SBR based on administrative registration data cannot be expected to assist in their measurement. If an enterprise undertaking illegal production is registered, it will be included in the SBR but the enterprise will disguise its illegal activities. If the enterprise is unregistered, it will be not be included in the SBR. In either case *the SBR plays no role*.

The same applies to underground activities. Unless underground production is very widespread, it is almost certainly a poor use of resources to try to measure it precisely. In any case, the SBR has no role to play as, by definition, the enterprises involved are not registered with the sources on which the SBR is likely to be based.

3.4.3 *Extraterritorial Organisations and Bodies*

International organisations, embassies and foreign government representations can be split into two groups:

- Those whose sites are deemed to form part of the economic territory of another country (e.g. embassies, consulates, military bases). These units should not be included in the SBR of the host country. They should be included as units in the SBR of the country to

²⁰ SNA2008: 3.96

²¹ SNA2008: 3.96

²² SNA2008: 25.25

which they belong, where they would probably be classified to ISIC section U. For example, the German embassy in Canberra, Australia should be a unit in the German SBR.

- Those whose sites do not form part of the economic territory of another country. This group includes units such as international organisations (e.g. the United Nations and its agencies, European Communities, OECD, IMF, World Bank, etc.) that are not be part of the economic territory of the host country according to the SNA2008. However, for the sake of completeness they should be included in the SBR of host country, as they do not appear in any other country's SBR. The NSI in the host country can decide how to treat them for the purposes of its economic statistics.

3.5 Summary of Coverage Recommendations

Advantages flow from including as many categories of enterprises as possible in the SBR. More complete coverage of enterprises means that the SBR can, for example, provide frames for a wide variety of different surveys in different sectors of the economy. On the other hand, the more categories included the more sources that have to be used and the more maintenance activities. Thus, NSI decisions which enterprises to include should be based on the appropriate international standards, the availability of coverage sources, and the corresponding maintenance costs.

Figure 3.4 summarises the recommendations for SBR coverage for the various categories of enterprises. It is provided as a guide only. It can be used as a template by means of which an NSI can describe the coverage of its SBR. An example from Australia is presented in Figure 3.5.

Figure 3.4: Recommended Minimum SBR Coverage

<i>Enterprise Type</i>	<i>Within Recommended Minimum SBR Coverage</i>	<i>Additional information</i>
Government	Yes	Component of formal sector. Inclusion of units in the SBR is preferable but having links in the SBR to data for government units held in an administrative source is acceptable.
Financial corporations and quasi-financial corporations (including those controlled by government)	Yes	Component of formal sector.
Non-financial corporations and quasi corporations	Yes	Component of formal sector.
Non-profit institutions serving households	Yes	Component of formal sector.
Registered household enterprises (including agricultural)	Yes	Component of formal sector.
Informal own account*, non-registered non-agricultural household market enterprises excluding self-employed professionals	No	Component of informal sector. If significant, cover through two stage household-enterprise survey but do not maintain the enterprises in the SBR
Informal own account*, non-registered non-agricultural household market that are self-employed professionals	Yes	Component of informal sector. Depends on suitable source being available

<i>Enterprise Type</i>	<i>Within Recommended Minimum SBR Coverage</i>	<i>Additional information</i>
Informal own account*, non-registered agricultural household market enterprises	No	Component of informal agriculture. If significant, cover through two stage household-enterprise survey but do not maintain the enterprises in the SBR
Unincorporated enterprises with employees non-agricultural, market producers	No	Component of informal sector. Include if a coverage source can be identified
Unincorporated units with employees agricultural, market producers	No	Component of informal agriculture. Include if a coverage source can be identified
Household non-market enterprises	No	If significant, cover their activities through two stage household-enterprise survey but do not maintain the enterprises in the SBR
*Own account implies self-employed with no employees		

Whatever, enterprise categories are covered, the NSI has to make provision for measuring or estimating the economic production of enterprises that are not included in the SBR. For example, if, in accordance with the previous recommendations, informal sector enterprises are not included in the SBR but their production is thought to be significant, then it can be measured by means of a two stage household-enterprise survey. The first phase is a household survey in which households having an enterprise producing for the market are identified. The second phase is a survey of these enterprises. The same approach can be used for informal agriculture.

Figure 3.5: Example - Australian SBR Coverage

Enterprise Type	Included in SBR	Additional information
Government	Yes	Government units are eligible for selection in labour collections and other SBR based collections. There is a link between the administrative government finance units and government units on the SBR
Financial corporations and quasi-financial corporations	Yes	These units are eligible for selection in labour collections and other SBR based collections. A link between the surveys of financial information and financial corporation units in the SBR is under development.
Non-financial corporations and quasi corporations	Yes	Micro corporations are separately identified
Non-profit institutions serving households	Yes	Micro NPIs are separately identified
Registered household enterprises with greater than a given number of employees (including agricultural)	Yes	Micro units are separately identified
Registered household enterprises with	No	

Enterprise Type	Included in SBR	Additional information
less or equal to than a given number of employees (including agricultural)		
Informal own account*, non-agricultural household market enterprises excluding self-employed professionals	No	The coverage of the ABS Business Register is based on entities which have registered for an Australian Business Number. Registration is mandated for businesses above a size cut-off. So this sector is known to be very small
Informal own account*, non-agricultural household market enterprises that are self-employed professionals	Yes	
Informal own account*, agricultural household market enterprises	No	The coverage of the ABS Business Register is based on entities which have registered for an Australian Business Number. Registration is mandated for businesses above a size cut-off. So this sector is known to be very small
Unincorporated enterprises with employees non-agricultural, market	Yes	
Unincorporated enterprises with employees agricultural, market producers	Yes	
Household non-market enterprises	No	
* <i>Informal</i> implies non-registered; <i>own account</i> implies self-employed with no employees		

4 Units in the SBR

4.1 Introduction

4.1.1 *Content and References*

This chapter describes the different types of units to be maintained in the SBR from both a conceptual and an operational view. The following broad groups of units are distinguished.

- *Statistical units* – units defined for statistical purposes, for which information is sought and statistics compiled; there are several standard types of statistical unit.
- *Legal, administrative and operational units* – legal units are recognised by law or society; administrative and operational units are units that legal units define in order to satisfy administrative regulations (for example reporting units for accounting purposes) and/or to manage themselves; all these units are the basis for creating and maintaining statistical units.
- *Observation and reporting units* - observation units are the units *about which* data are actually obtained during the course of surveys, and reporting units are the units *from which* data are actually obtained.

The sets of statistical units in the SBR provide the basis for presenting the various aspects of the economy in an unduplicated and exhaustive manner. It is important to have harmonised definitions of each type of statistical unit so as to have coherent economic statistics with consistency across sectors, countries and geographic areas, and over time. Reduction in response burden achieved by collecting data from administrative sources also requires a common understanding of these units.

Usually, a statistical unit is also an observation unit and a reporting unit. However, the three types of unit may differ, especially, for example, when dealing with large and complex enterprises.

Links between the different types of units are important. Each NSI has to identify the differences between legal units, administrative units, statistical units and observation units and determine how it is possible to link from one to another.

The main characteristics of the various types of units and their links are presented in Chapter 5. The general structure of the SBR database required to support storage and access of units and their characteristics is presented in Chapter 11.

The main references for this chapter are Eurostat's *Business Register Recommendations Manual*, the African Development Bank's *Guidelines for Building SBRs in Africa, SNA2008* and *ISIC Rev 4*. Other sources are footnoted and listed in Annex A.

4.1.2 *Brief Presentation of Units and Problems Covered*

Statistical units

Statistical units (SUs) about which information is sought and for which statistics are ultimately compiled are the crucial units in the SBR. They should be economically significant, harmonised and have associated with them all the characteristics required for providing frames for economic surveys. They need to be harmonised so as to underpin the production of consistent statistics. They are not always directly observable.

The main SUs used internationally are:

- enterprise group;

- enterprise;
- establishment, also called local kind-of-activity unit (LKAU).

Other units with international acceptance are:

- kind-of-activity unit (KAU);
- *local unit*, of which there are two important types – local unit of enterprise and local unit of legal unit.

In the case of a *multi-national enterprise (MNE)*, the enterprise and the enterprise group to which it belongs cross national boundaries. Until recently individual countries have been interested only in those parts of enterprise groups and enterprises that operate within the national boundaries. These are referred to as “*truncated enterprise groups*” and “*truncated enterprises*” respectively, although these definitions have yet to receive international recognition. These Guidelines propose covering not only the truncated parts of enterprises but expanding coverage to multinational enterprise groups and to multinational enterprises.

Those enterprise groups comprising enterprises that are all contained within the national borders of a country are called *all-resident groups*.

Developments in definitions of enterprise group and enterprise in the European Statistical System (ESS)

The definitions of statistical units including *enterprise group* and *enterprise* are currently under revision in the ESS. The aim is to support development of statistics that take into account the globalisation of the economy and, more specifically, to establish a cross-cutting regulation suppressing sources of inconsistencies between countries and statistical domains and allowing integration of business statistics within EU Member States.

The revised definitions, as at mid-2014, are presented in Section 4.3.2 (for enterprise group) and in Section 4.4.4 (for enterprise). They take into account the outcome of two five-yearly work programmes launched by Eurostat and national members of the ESS. One is “*Profiling the large and complex multinational enterprise (groups)*”; the other is “*Consistency*” of business statistics between countries and domains through the “*Task Force on Statistical Units*”.

The revised definitions will be included in the Eurostat *Framework Regulation Integrating Business Statistics (FRIBS)*, 2014 for application in the coming years.

Legal, administrative, operational, and local units of legal units

The SBR records not only data about the SUs but also data about all the other types of units that are needed in practice for the creation and maintenance of the SUs. These types of units include:

- *legal unit* – a unit that is recognised by law or society independently of the persons or institutions that own it; this unit is pivotal in linking administrative units and statistical units, because, among other reasons, most administrative registers are based on legal units;
- *administrative unit* – a unit specifically designed for the purposes of conforming with an administrative regulation, for example VAT or Social Security;
- *operational unit* – unit defined by a legal unit for the purposes of organising itself, for example, a division, branch, workshop, warehouse or outlet;

- *local unit of legal unit* – an operational unit of a legal unit at a single location.

Observation and reporting units

In addition, for the collection of economic data, two more types of units are defined:

- *observation unit* - a unit, usually a statistical unit, about which data are obtained during the course of a survey. If a target statistical unit is not directly observable, the corresponding observation unit may be a legal unit or administrative unit linked to the statistical unit. Mostly observation units are legal or administrative units but they may also be operational units;
- *reporting unit* - unit from which data about an observation unit are obtained during the course of a survey.

Example: European survey on outward foreign affiliates

In this survey:

- the *statistical units* are enterprises operating abroad (foreign parts);
- the *observation units* may be the subsidiaries abroad;
- the *reporting units* are often the global group heads or the highest level consolidation units, if different.

(All these types of units are described later in the chapter.)

Sub-populations of units

Sub-populations of the above mentioned units are of great interest for economic or structural analysis. They are typically created according to their characteristics, such as economic activity code, size code, and/or institutional sector, whether they are a market/non-market producer, and/or region of operations, as discussed in the Chapter 5.

4.2 Types of Statistical Units

As noted above, a *statistical unit* is a unit about which information is sought and for which statistics are ultimately compiled. It is the unit that provides the basis for statistical aggregates and to which tabulated data refer. Data for a statistical unit may be directly *observable*, as for example in the case where a statistical unit coincides with a legal unit, or may be *derived* by splitting or grouping data from observation units²³ using statistical estimation methods, or in some cases may be provided directly by the respondent²⁴.

The various types of statistical units are not independent, but are linked to each other forming a *statistical units model*, that is, a conceptual framework of statistical units and their relationships. A statistical units model provides a simplified conceptual abstraction of the rather messy and complicated set of legal, operating, and administrative units that exist in the “real world”, i.e., the economic world that the NSI is measuring through statistical units.

²³ Extension of definition in §5.27 of UN Working Paper on Statistical Business Registers, which considers splitting units but not grouping them.

²⁴ This is often the case for enterprises in a large enterprise group that have been profiled for SBR purposes (as further discussed in Section 6.6.4) and that may operate on a pure “*gentleman’s agreement*” basis.

In defining standard statistical units, the two most authoritative international sources are the SNA2008 and ISIC Rev 4. As described in Chapter 3, the SNA2008 introduces the notions of institutional unit and enterprise, and of statistical units formed by partitioning an enterprise.

- An *institutional unit* is defined as 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 entities²⁵. On the basis of their principal functions, behaviour, and objectives, institutional units are grouped for national accounts purposes into five mutually exclusive institutional sectors²⁶, namely non-financial corporations, financial corporations, general government, non-profit institutions serving households (NPISHs) and households.
- An *enterprise* can be regarded as core statistical unit of the SNA2008. It is defined as *the view of an institutional unit producing goods and services*²⁷, as further discussed in Section 4.4.

The SNA2008 acknowledges that a large enterprise may be engaged in several different types of economic activities, at several different locations. The treatment of such an enterprise for statistical purposes as having a single type of activity at a single location would distort the resulting statistics. Thus, the SNA2008 recommends that large enterprises be partitioned into smaller, more homogeneous producing units that can be more precisely classified and that collectively represent the enterprise as a whole as follows²⁸.

- The type of unit resulting from partitioning an enterprise by its different economic activities is a *kind-of-activity unit (KAU)*.
- The type of unit resulting from partitioning an enterprise by its different locations is a *local unit*.
- The type of unit resulting from partitioning an enterprise by both its economic activities and its locations is an *establishment*, also called local kind of activity unit (LKAU).

This set of statistical units is illustrated in Figure 4.1. The enterprise is the basic unit; all other units are defined by partitioning the enterprise.

Figure 4.1 SNA2008/ISIC Rev 4 Statistical Units Model

		Partition by location	
		No	Yes
Partition by activity	No	Enterprise	Local unit
	Yes	Kind -of-activity unit	Establishment (Local kind - of-activity unit)

However, despite the widespread agreement on the use of SNA2008 and ISIC Rev 4, this model has not been universally implemented. In practice, there are deviations from the conceptual framework and not necessarily all types of statistical units are used.

²⁵ SNA2008: 4.2

²⁶ SNA2008: 4.17

²⁷ SNA2008: 5.1

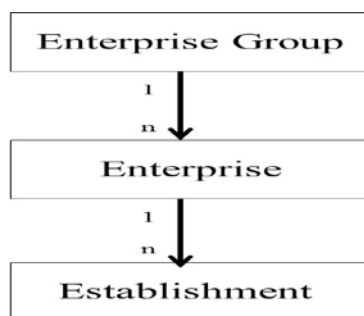
²⁸ SNA2008: 5.12-14

Recommended basic statistical units model

The basic model described in these Guidelines (in Sections 4.3-4.5) is the one recommended in the AfDB SBR Guidelines. It is a simplification and extension of the SNA2008 model and is loosely based on a 2008 European Commission regulation²⁹, which requires that four units – namely, *legal unit*, *enterprise*, *local unit*, and *enterprise group* - need to be maintained in the national SBRs of the EU member states.

The model comprises three hierarchically organized statistical units, as illustrated in Figure 4.2, namely *enterprise group*, *enterprise* and *establishment (local kind of activity unit)*.

Figure 4.2 Recommended Statistical Units Model



The enterprise and the establishment are essentially the same two statistical units as defined by SNA2008. The enterprise can be used at global level for multinational enterprises which, although relatively few in number, may represent a significant part of a national economy. The need for an enterprise group is recognised in the SNA2008 even though it is not defined. The SNA2008 notes that large groups of corporations, or conglomerates, may be created whereby a parent corporation controls several subsidiaries, some of which may control subsidiaries of their own, and so on. For certain purposes it may be desirable to have information relating to a group of corporations as a whole³⁰. The notion of an enterprise group is essential in SBR construction and maintenance as it is the basis for delineating and recording the enterprises that belong to a conglomerate, and it is vital for structural statistics as it facilitates linkage of national operating enterprises to the foreign enterprises that control them.

The way in which the model fits with the economic and administrative worlds it aims to represent is illustrated in Figure 4.3.

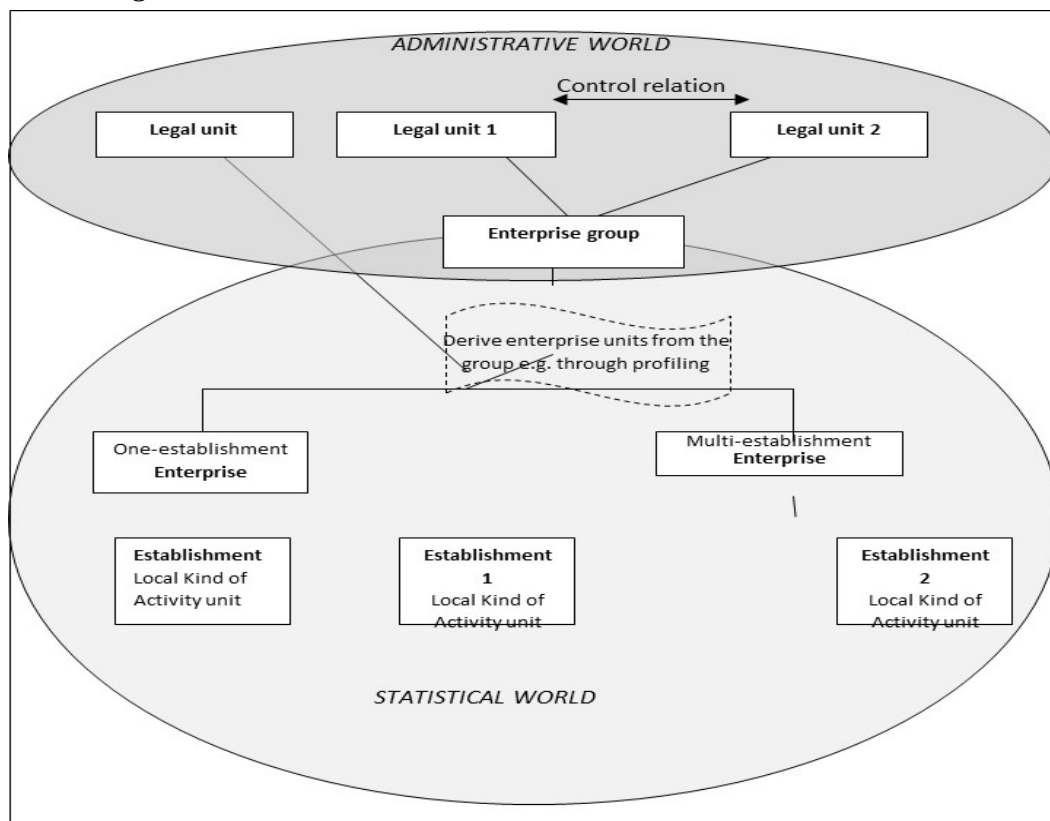
- The linkage between an administrative unit and a statistical unit is typically through a legal unit but in more complex cases may be through an enterprise group.
- Delineation of an enterprise does not depend upon geographical territory or country.
- By contrast, an establishment is always located in a specific country or a specific region within a country.
- Even though an enterprise is usually equal either to a legal unit, or to the combination of a number of legal units, the definition allows a legal unit to be split into parts, each part being an enterprise.

²⁹ Business Register (BR) Regulation (EC) No. 177/2008

³⁰ SNA2008: 4.51

If required, the basic model can be extended to include *local units* in accordance with the SNA2008 model³¹.

Figure 4.3: Links between Statistical and Administrative Units



4.3 Enterprise Group (and Truncated Enterprise Group)

4.3.1 Introductory Remarks

Statistics based on enterprise groups are presently not very far developed, but with increasing globalisation, they are progressively becoming of more interest. Statistics using enterprise groups as statistical units are also useful for the analysis of market concentration.

Enterprise groups are also essential for delineating enterprises consistently and by a top-down process. They are needed for compiling data on the impact of multinationals on an individual country's economy, presently mostly in terms of inwards and outward foreign affiliate statistics (FATS) and foreign direct investment (FDI) statistics.

Ideally, delineation of the enterprises within an enterprise group should be done using a *top-down approach*, starting from the enterprise group level, in consultation with the *global group head* or the *global decision centre*. (These terms are defined shortly.) This is easiest when the global group head is resident in the country where the delineation is initiated (the initiating country). In cases where only a part of an enterprise group is resident in the initiating country, cooperation and data exchange between the SBRs in all relevant countries are required in order to delineate all the statistical units belonging to the enterprise group and

³¹ SNA2008: 5.13

the corresponding truncated enterprise groups and enterprises in a coherent way. However, such cooperation is not easy to achieve. This means that the top-down approach may not be feasible and thus a *bottom-up approach*, starting from legal units belonging to the enterprise group resident in the initiating country, may be the only way to delineate the statistical units.

The interest in enterprise groups is not related to the degree of development of an economy but to the shares of the main aggregates (such as production, value added, and foreign trade) that these enterprise groups control, whether the control is national or foreign. In consequence it is very important to identify enterprise groups, using appropriate methods, even if they are few in number.

The concept of enterprise group is supplementary to the SNA2008 model. In fact, consolidation of enterprises into groups is an item on the SNA Work Group research agenda. As the SNA2008 does not provide a definition of enterprise group, EU regulations and research outputs are the best sources.

European Treatment of Enterprise Group

The enterprise group is one of the units selected for establishing a common framework for SBRs in Europe. As previously noted, *Regulation (EC) 177/2008*, the so-called “*BR regulation*” includes a simplified model containing just four units, namely, legal unit, enterprise, local unit, and enterprise group.

The *Eurostat BR Recommendations Manual (21.21-21.26)* states that:

“Whereas the enterprise is an actor in the economy at the level of the production process, with relative autonomy with respect to the allocation and use of its current resources, ***the enterprise group*** is an actor at a more strategic level taking strategic decisions on behalf of its constituent enterprises (e.g. on product policies, on major expansions, etc.). Since the enterprise group as a whole is subject to demographic events, the statistical unit ‘enterprise group’ in the register must be maintained according to special continuity rules for the global, truncated and all-resident groups.”

The characteristics of an enterprise groups represent properties of the group at global or national (truncated or all-resident) level. Determining whether an enterprise group refers to consolidated properties or non-consolidated properties is recommended.

4.3.2 Definition of Enterprise Group (and Truncated Enterprise Group)

In the absence of an internationally agreed definition of enterprise group the EU definition is relevant. However, the definitions in the European 1993 Statistical Units Regulation are under revision³² and the revised definitions not yet officially agreed. Thus, the definition of enterprise group and the associated operational rules proposed below are drawn from documentation agreed by the Eurostat Task Force on Statistical Units in the Spring, 2014.

³² The Eurostat Framework Regulation Integrating Business Statistics (FRIBS) project is establishing a cross-cutting regulation to integrate business statistics. An annex will provide a list of statistical units.

Definition and operational rules

An enterprise group is a set of legal units bound together by legal and/or financial links and under the same control.

Operational rule 1 – Control

Control over a legal unit is defined as the ability to determine general corporate policy. It can be exercised by (a) owning more than half of the voting shares, (b) having the right to appoint or remove a majority of the members of the management, (c) having the right to exercise a dominant influence over it (d) controlling more than half of the shareholders' voting power of another legal unit directly or indirectly, or otherwise (e) proving that there is de facto control exercised. Indirect control refers to controlling a legal unit via another legal unit. This includes also cumulative control, i.e. controlling two or more legal units that together own more than half of the voting shares of the legal unit in question.

An enterprise group is controlled by its *global group head (GGH)*. The GGH is defined as the unit (legal or natural person) which controls all legal units of the group and is not controlled by any other legal unit.

Operational rule 2 – Management and Control

An enterprise group is always controlled by a single GGH. Typically a GGH controls just one enterprise group, however may control more than one. Consolidation of the accounts at a level below the GGH and the existence of several consolidated accounts is an indication that a GGH controls several enterprise groups.

The unit carrying out the actual management of the enterprise group is named the *global decision centre (GDC)*. It is not necessarily identical with the GGH.

Operational rule 3 - Consolidation and Control

Shares of affiliates' undertakings have to be listed in the balance sheet of a company³³.

All the legal units forming the enterprise group have to be consolidated in its accounts. If the difference between the consolidation perimeter and the set of legal units concerns statistically non-significant legal units, the consolidation perimeter is used for statistical purposes.

If the GGH does not produce consolidated accounts and reports, the unit below the GGH which has to produce consolidated accounts and annual reports is called the *highest level Consolidation Unit* and determines the consolidation perimeter of the enterprise group.

4.3.3 Link between National and Global Views: Truncated Enterprise Group

In practice, an enterprise group may be an *all-resident* enterprise group (otherwise known as *domestic* group or *national* group) or *multinational* enterprise group (either domestically controlled or foreign controlled). For the purposes of a national SBR, it is useful to define the concept of a *truncated enterprise group* of a multinational enterprise group.

Definition

The part of a multinational enterprise group that comprises only the legal units resident in the particular country is defined to be a “truncated enterprise group” within that country.

If parent is non-resident it is possible that the truncated enterprise group consists of several seemingly unlinked units and subgroups.

³³ Following the EU Seventh Council Directive and its amendments.

A truncated enterprise group may comprise just one unit, a parent or subsidiary, in the national territory. Such truncated enterprise groups may be difficult to identify based on national information only.

4.3.4 Identifying Enterprise Groups in Practice

Group boundary

The usual starting point in establishing the *boundary* (also called *perimeter*) around an enterprise group is to consider the consolidated accounts it contains. In order to delineate enterprise groups based on mutually exclusive accounting groupings, the following rules are applied.

- Consider accounting groups at the highest consolidation level possible (often the GGH).
- Include in an enterprise group the units for which the accounts are entirely integrated in those of the consolidating company, and not the units whose accounts are treated differently (by the *equity method* in which only the profits and the value of the owned share of equities are consolidated).
- Add majority-controlled units with accounts not included in the overall consolidating by virtue of application of one of the criteria allowed by the seventh Directive³⁴, i.e. difference in the type of activity or small relative size.
- Discount temporary links of less than a year (which means not including temporarily affiliated units in the enterprise group).

Special cases of global group heads

The GGH of an enterprise group is the enterprise corresponding to the parent legal unit, i.e., the legal unit that is not controlled either directly or indirectly by any other legal unit. The enterprises it controls are referred to as affiliates. Affiliate enterprises of an affiliate enterprise are considered to be subsidiaries (sub-affiliates) of the parent enterprise. However there are some forms of cooperative or mutual associations where the ownership is reversed, the “parent” enterprise being actually owned by the units of the group.

When the GGH does not manage the enterprise group, usually the decisions are taken by the affiliate called global decision centre (GDC). The SBR has to collect the identification numbers and characteristics of the GGH and of the GDC in order to differentiate them clearly when relevant.

When the GGH and the GDC are not located in the same country, the enterprise group may be counted twice in supra-national aggregates, that is, in both the GGH country and in the GDC country. Thus, coordination between countries is useful. Such coordination exists in the EU as discussed in EuroGroup Register example below. The enterprise group is considered to be a resident of the GDC country and not of the GGH country, when the two countries are different.

The case becomes more complex when the GGH is a natural person living in a country different from the GDC country. Conceptually, the natural person should be registered as GGH but the treatment of this case is an issue because the legal ability to register a natural person as GGH differs from one country to another. However, the SBR can make provision to record the natural person as the GGH even if that person is not legally registered as such.

³⁴ EU legislation based on Article 54(3)(g) of the Treaty on consolidated accounts

Conglomerates and Sub-groups

An enterprise group may correspond to a *conglomerate* held together by a network of complex relationships and then it frequently covers a very wide range of activities. *Subgroups* can often be identified within enterprise groups.

It is useful to identify all (majority and minority) links between the GGH and a controlled legal unit via the network of subsidiaries and sub-subsidiaries. This allows the group's entire organization to be depicted.

Multinational Groups

The major challenge is that NSIs in general have only very limited information on the non-resident part of multinational enterprise groups. Statistical surveys and commercial data providers may assist in providing some information. However, it is difficult from a national perspective to “*see the whole elephant*”³⁵ or for all the NSIs involved to have the same view of a multinational enterprise group. Nevertheless this should be the aim of each SBR.

Evidently there is more chance of this being possible for a multinational enterprise group that resides entirely within EU member states where there are programs for collaboration than for a group involving countries where no such collaboration exists.

4.4 Enterprise (and Truncated Enterprise)

4.4.1 Introductory Remarks

The enterprise is the core standard statistical unit maintained in the SBR. According to the SNA2008 and ISIC Rev 4, it is *the level of statistical unit at which all information relating to 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*³⁶.

The following sections provide a definition of enterprise based on this approach and discuss the ways in which the definition may need to be changed to deal with globalisation and better representation of multinational enterprises (MNEs).

4.4.2 SNA2008 and ISIC Rev 4 Definitions of Enterprise

SNA2008 Definition

An ***enterprise*** is the view of an institutional unit producing goods and services. The term *enterprise* may refer to a corporation, a quasi-corporation, a non-profit institution (NPI), or an unincorporated enterprise³⁷.

In this context, as previously noted, an ***institutional unit*** is an economic entity that is capable, in its own right, of owning assets, incurring liabilities and engaging in economic activities and in transactions with other entities. There are two main types of units in the real world that may qualify as institutional units, namely persons or groups of persons in the form of households, and legal or social entities.³⁸

³⁵ “Seeing the whole elephant: A proposed experiment on measuring the activities of multinational enterprises”; Richard Barnabé; Statistics Canada 2003

³⁶ From ISIC Rev 4 Introduction

³⁷ SNA2008: 5.1

³⁸ SNA2008: 4.2-4.3

The main attributes of institutional units are described as follows.

- a) An institutional unit is entitled to own goods or assets in its own right; it is therefore able to exchange the ownership of goods or assets in transactions with other institutional units.
- b) It 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.
- c) It is able to incur liabilities on its own behalf, to take on other obligations or future commitments and to enter into contracts.
- d) Either a complete set of accounts, including a balance sheet of assets and liabilities, exists for the unit, or it would be possible and meaningful, from an economic viewpoint, to compile a complete set of accounts if they were to be required.

ISIC Rev 4 Definition

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. 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 an institutional unit - a household or government unit - only in its capacity as a producer of goods and services.³⁹

The limitations of these definitions are that they do not take into account the relationship between enterprises and the enterprise groups to which they may belong. Nor do they deal with MNEs.

4.4.3 Current European Definition of Enterprise

The following European definition⁴⁰ seems more operational, but has not ever been fully applied and is under revision.

Definition

An **enterprise** is the smallest combination of legal units that is an organizational unit producing goods or services which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources.

An enterprise carries out one or more activities at one or more locations.

An enterprise may be a sole legal unit.

The enterprise thus defined is an economic unit that can, in certain circumstances, correspond to a grouping of several legal units. There are two main cases.

- Some legal units perform activities exclusively for other legal units and their existence can only be explained by administrative factors (e.g., tax reasons) as they have no economic significance. Typically, the activities of these legal units should be seen as *ancillary activities* of the parent legal unit they serve, to which they belong, and to which they must be attached to form an enterprise. Many of these legal units either have no direct employment, or are the employer of the whole workforce of the grouping.

³⁹ Explanatory Notes in ISIC Rev 4 Introduction

⁴⁰ 1993 Statistical Units Regulation

- The enterprise is autonomous as far as production (including allocation of current resources) is concerned. Financial and most structural decisions are taken at a higher level in the enterprise group to which the legal units belong.

4.4.4 Towards a New Operational Definition of Enterprise

Revisions to the European definition are being considered for two main reasons:

- primarily to take into account globalisation of the economy;
- also to better harmonise the way in which general government bodies are represented as enterprises.

The need to develop statistics that take into account the globalisation of the economy, including the impact of enterprise groups and of multinational management of production, implies the use of a top-down approach in delineating statistical units. Such an approach starts with global analysis of the managerial, accounting and legal structure of an enterprise group with the aim of delineating the statistical units, in particular enterprises, within that group, and their links with legal units or parts of them. This may lead, in exceptional cases, to splitting a legal unit by allocating its parts to several enterprise units.

The legal unit structure is not often useful in delineating the statistical units within government bodies as the legal structure of public administration varies considerably between countries, and even within a country, depending on the sector. In particular, for general government bodies, legal units may be defined only at a very high level, or at much lower levels in the hierarchy. Hence, the top-down approach for defining enterprises is recommended. It may lead, in exceptional cases, to splitting a legal unit into several units in the SBR. Such a legal unit may, for functional purposes, be broken down into 'quasi' legal units.

In both the above cases, using the present definition that requires *the smallest combination of legal units for deriving the enterprise unit (and forbids splitting legal units)* may lead to an inadequate delineation of SUs and to lack of comparability within and between countries. To correct it, European experts are working on a new definition without any change to general principles.

New operational definition of enterprise proposed for the European Statistical System (draft as of Spring, 2014)

1. An *enterprise* is an organisational unit that has a sufficient degree of autonomy in decision-making and that can sell at its own will goods and services to a third party. It can be constituted by one legal unit, a combination of legal units or of parts of legal units. An enterprise carries out one or more activities at one or more locations.
2. An enterprise can correspond to:
 - a single legal unit not controlled by any other legal unit;
 - an enterprise group as a set of legal units under common control, or
 - an autonomous part of an enterprise group.
3. *Organisational* means that (for the economic activity in which the enterprise is engaged) a planned and formal structure is identified. This structure is able to govern its production processes and manage its factors of production. *Selling to a third party* means that exchanges are made with an independent buyer on the basis of commercial considerations.
4. Selling goods and services at its own will to a third party is a necessary condition to

identify the autonomous enterprise.

5. *Sufficient autonomy in decision-making* means that the enterprise may not be the owner from a legal point of view but has control of the use of its productive means, processes and outputs of the economic activities in which it is engaged.

A significant difference between this definition and the current European (1993) regulation is that in this definition *an enterprise is not, by definition, limited to a national territory*. Its boundary is delineated by a global top-down approach rather than by a national bottom-up approach based on groups of legal units within national boundaries.

In practice, an enterprise (or truncated enterprise) is often a combination of legal units and not always the smallest combination. In a few cases, legal units may be split and allocated to several enterprises.

Example from France

Some public companies or newly privatised companies (as in rail transport - SNCF - or in postal and financial activities - La Poste) are profiled into several enterprises that have different activities and that may belong to different institutional sectors, for example, financial and non-financial sectors. The GGH, which in these cases is also by far the biggest legal unit of the enterprise group, is organised into different business units (autonomous operational divisions) that are defined as enterprises. Thus, the legal unit has been split in line with the organizational criteria.

4.4.5 National and Global Views: the Truncated Enterprise

For the purposes of a national SBR it is necessary to distinguish an *enterprise* and a *truncated enterprise* in the case of a multinational enterprise. As noted above an *enterprise* is defined irrespective of national boundaries, and a *truncated enterprise* is its national component. A *global enterprise is the autonomous part of a multinational enterprise group irrespective of the national borders. A global enterprise may thus be acting in one or more of those countries where the multinational enterprise group has its resident legal units*. For an all-resident enterprise group, the truncated enterprise and the enterprise are the same unit. A truncated enterprise is the basis for usual national business statistics even though, because it belongs to a multinational enterprise it does not necessarily satisfy the criterion for autonomy of decision-making.

4.4.6 Simplified Correspondence between Legal units and Enterprises

Whilst it is recommended that all NSIs introduce enterprises as distinct from legal units into their SBRs, it is recognised that an enterprise as defined in the SNA2008, or in the current or proposed European definitions, may not be practically implementable in all countries. Data on the basis of which to combine legal units to form an enterprise may not be available, or the SBR system may currently not support the separate recording of legal units and enterprises as distinct units.

In the absence of data at enterprise level, it is acceptable that statistical data and demographic data are produced on the basis of legal units. However, in adopting this approach it is highly recommended that the notion of enterprise be kept separate from that of legal unit, leaving the way open in the future to be able to distinguish and treat individually the most important

cases where enterprises that are part of large enterprise groups (whether foreign or domestic) are not coincident with legal units.

AFDB SBR Guidelines recommendation for enterprise definition

The AfDB Guidelines recommend that legal units and enterprises are defined to be in one-to-one correspondence in the sense that there is one and only one legal unit per enterprise. The benefit of this approach is that it is simple and easy to operationalize. An enterprise is created for each active or potentially active legal unit.

4.5 Establishment⁴¹

4.5.1 Introductory Remarks

As discussed in Section 4.2 and illustrated in Figure 4.1, in partitioning an enterprise by economic activity and geography, the establishment (local kind-of-activity unit) incorporates both the kind-of-activity dimension and the locality dimension. For national accounts purposes, industries are defined in terms of establishments. An industry consists of a group of establishments engaged in the same, or similar, types of activity. In the national accounts, production accounts and generation of income accounts are compiled for industries as well as for institutional sectors.

According to the ISIC Rev 4 the establishment should be used for the analysis of transactions in goods and services and for the compilation of the production account. Furthermore, policy makers interested in investigating regional (state, county, local area) differences want demographic data at sub-national level, as provided by data at establishment (or local unit) level.

4.5.2 ISIC Rev. 4 Definition of Establishment

Definition

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 SNA2008 definition is essentially the same. Although it does not specifically include the *non-ancillary* criterion, this is implied.

Although the definition of an establishment allows for the possibility that there may be one or more secondary activities carried out, they should be small in magnitude compared with the principal activity. If a secondary activity within an establishment is as important, or nearly as important, as the principal activity, then the unit should be subdivided so that the secondary activity is treated as taking place within an establishment separate from the establishment within which the principal activity takes place.

⁴¹ As previously noted, some organisations and countries refer to an *establishment* as a *local kind of activity unit*.

In the case of most small and medium-sized businesses, the enterprise and the establishment are treated as being the same. The breakdown of an enterprise in two or more establishments is only required if it has locations situated in two or more different regions and/or if it has two or more different, economically significant activities.

4.5.3 Implementation Notes

Current European regulations do not require the establishment to be directly maintained in the SBR⁴². In practice, when the characteristics required to build national accounts are not available at the detailed local and activity levels, they should be estimated. Data for establishments may also be approximated from *local units of legal units* (as further discussed in the next section) maintained in, or by means of, administrative registers. Often, but not always, establishments and local units are in one-to-one relationship, in particular for small and medium-sized businesses.

Example from European 1993 Statistical Units Regulation: Definition of the LKAU

In Europe, the unit equivalent to the establishment is called a *local kind-of-activity unit (LKAU)*. It is the part of a kind-of-activity unit (KAU) that corresponds to a local unit. A local unit is an institutional unit, or part of an institutional unit, producing goods or services situated in a geographically identified place.

An LKAU groups all the parts of an institutional unit in its capacity as producer contributing to the performance of an activity at class level (four digits) of the NACE Rev. 2 and corresponds to one or more operational subdivisions of the institutional unit.

The institutional unit's information system must be capable of indicating or calculating for each LKAU at least the value of production, intermediate consumption, compensation of employees, operating surplus and employment and gross fixed capital formation.

An LKAU may never belong to two different institutional units.

⁴² “2008 BR regulation” (Regulation EC 177/2008)

Proposed new definition of LKAU in the European Statistical System (Spring 2014)

A local kind-of-activity unit (LKAU) is a KAU or part of a KAU that is situated in a single geographical location and engages in only one kind of productive activity, producing market output, or – in case of more than one productive activity - in which the principal productive activity producing market output, accounts for most of the value added.

The delineation of activities is based on the valid version of the NACE classification and the location is based on the local unit.

It should be possible, in principle, to derive a minimum set of economic indicators related to its production activity, in particular value of production, intermediate consumption (except overhead costs), compensation of employees and gross fixed capital formation of buildings and structures, machinery and equipment as well as employment.

(1) Operational rule: Cases in which LKAU may be delineated

Delineation of an LKAU should mainly be done for enterprises for which KAUs have been delineated and where these activities take place in different locations/regions. In the case where the prevailing activity in a location is not the main activity of any KAU of the enterprise, and this activity accounts for less than the threshold to form a KAU at enterprise level, an additional KAU may be delineated so that appropriate LKAUs can be delineated.

An LKAU is not strictly required to form part of a KAU corresponding to a single local unit. It is acceptable for an LKAU to comprise the sum of all respective activities performed within the lowest level of NUTS for which LKAUs are required.

(2) Operational rule: Estimates

In case that not all of the economic indicators are available from the respondents, they may be estimated by the national *statistical* authorities.

(3) Operational rule: Ancillary activities

In case that in a local unit only ancillary activities are carried out, this unit is treated as a LKAU and classified according to the activity of the KAU (for national statistics) as well as of the ancillary activity (for regional statistics).

4.6 Other Statistical Units

4.6.1 Introductory Remarks

The enterprise group, enterprise and establishment are the three most important statistical units recommended by these Guidelines. However, an NSI may wish to use either or both of two additional types of statistical unit defined by the SNA2008, namely the *kind-of-activity unit (KAU)* and the *local unit*, as previously discussed and illustrated in Figure 4.1.

The need for the establishment unit is based on the assumption that the aim of the statistical programme is to compile production or production related statistics that are classified both by activity and by geographical region. In circumstances where precision in both geographic and

economic activity dimensions is not required simultaneously, the KAU and/or local unit may be more appropriate.

4.6.2 Kind-of-Activity Unit

SNA2008 Definition

A kind-of-activity unit (KAU) 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 added⁴³.

The unit is characterized by homogeneity of activity. Compared with the establishment, there is no restriction on the geographic area in which the activity is carried out.

European Example of KAU

In Europe some conditions are added to the definition of a KAU.

- The delineation of the activities should be based on the most recent version of the NACE classification.
- The KAU is producing *market output*.
- For the KAU it should be possible, in principle, to derive a minimum set of economic indicators related to its production activity, in particular, value of production, intermediate consumption (except overhead costs), manpower costs and employment and gross fixed capital formation of buildings and structures, machinery and equipment.
- Ancillary activities should not be regarded as a KAU.

A new proposal also includes the following operating rules:

Operational rule 1. Delineation of a KAU may be restricted to enterprises that, because of their sizes (e.g. production values), have significant influence on the aggregated industry data at KAU level, and:

- at least one secondary activity of the enterprise accounts for more than 30% of its total production at class level of the valid NACE classification; or
- at least one secondary activity of the enterprise accounts for more than 20% of its total production at the division level of the valid NACE classification.

Operational rule 2: In the case of an enterprise that is not covered by rule 1, the KAU is considered to be equal to the enterprise.

Operational rule 3: In case that not all of the economic indicators are available from the respondents, they may also be estimated by the NSI.

⁴³ SNA2008: 5.12

4.6.3 Local Unit

SNA2008 Definition

A local unit is defined as an enterprise or a part of an enterprise (for example, a workshop, factory, warehouse, office, mine or depot) which engages in productive activity at or from one location. Location may be interpreted according to the purpose narrowly, such as at a specific address, or more broadly, such as within a province state county, etc.⁴⁴

The definition refers only to the geographic dimension; it does not refer to the kind of activity that is carried out. The local unit thus defined is conceptually different from a legal or an administrative unit (as described in later sections). However, it may coincide with such a unit, or be defined on the basis of such units.

Example: European (1993) statistical units regulation definition of local unit

Eurostat BR Recommendations Manual – Chapter 5c

The official name of the local unit is generally the same as the enterprise (more precisely: the legal unit) that controls it, with some additional part specifying usually location or activity. If there is only one local unit in the enterprise, a separate name may not exist.

The local unit must be linked to the enterprise to which it belongs. This link can be included in the SBR by adding the identity number of the enterprise to the local unit file (and vice versa). Other ways are conceivable, for example when the enterprise consists of one local unit only, another (simple) arrangement of the SBR is possible.

It is proposed that an *ancillary local unit* be distinguished when the activity carried out in the local unit constitutes an ancillary activity of the enterprise to which it belongs. This characteristic enables statistical analyses to reallocate the cost of ancillary activities to the activities for the benefit of which they are pursued. If the local unit has clearly been identified as a non-ancillary unit it should also be indicated. The absence of any mark should be interpreted as 'no knowledge' about this characteristic.

a single local unit may be spread over several adjacent administrative areas, in which case, by convention, the postal address is the determining factor.

3. The boundaries of the unit are determined by the boundaries of the site, for example, a public highway running through does not interrupt the continuity of the boundaries.

The definition is similar to the one in paragraph 101 of the introduction to ISIC Rev. 3 in that it concerns localization in the strict sense of the term, but differs from the definition in paragraph 102 in that this strict sense may not vary according to the statistics under consideration. In addition, the criterion of persons working in the unit is normally applied.

4. The ESA-REG (the regional application of the ESA) uses the same definition of local unit for regional accounts purposes. It interprets local as meaning smallest administrative area, for example, the commune in Denmark.

⁴⁴ SNA2008: 5.12

4.7 Legal Unit - Pivotal between Administrative Units and Statistics

4.7.1 *Introductory Remarks*

Although enterprises and legal units are usually closely related, the two concepts should not be confused. While legal units are independent in a legal sense, they may not necessarily constitute independent economic units with decision-making autonomy for their activities, i.e., enterprises. Data from legal units without decision-making autonomy may not be comparable with data from legal units with decision making authority.

Legal units play a pivotal role in the construction and use of an SBR. Administrative data about economic units are mainly available from legal units. Legal units are often the reporting units from which information is collected about enterprises (as contact and address information is usually available for legal units). Thus the links between the legal units and the corresponding enterprises should be always maintained in the SBR.

Whilst legal units are usually the building blocks used in defining enterprises *they are not themselves necessarily standardised and comparable across countries* since they reflect national administrative and legal requirements, which may differ across countries.

While the focus of the SBR is typically businesses, i.e., legal units engaged in commercial economic market-oriented production, the SBR may include other legal units, for example non-profit institutions (NPIs) and government units.

Administrative regulations usually apply to legal units. However, in responding to the requirements of a particular regulation, and in interacting with the corresponding administrative authority, a legal unit may register and report on the basis of *one or more parts of itself*⁴⁵, referred to in these Guidelines as *administrative units*. Administrative data are available for these administrative units.

Thus, although in most cases a legal unit has a single administrative unit for a single administrative source, in some cases, often the most important cases, a legal unit has several administrative units for a given administrative source. Furthermore, the sets of administrative units are typically different for different administrative sources. As administrative data are the major input for the SBR, understanding and recording the relationships between the legal units and the various sets of administrative units is vital.

4.7.2 *Definition of Legal Unit*

Definition

*Legal units are recognized by law or by society, independently of the persons or institutions that own them. The characteristics of a legal unit are the following: they own assets, they incur liabilities and they enter into transactions with other entities (contracts)*⁴⁶.

Legal units include:

⁴⁵ Local units of legal units are examples of such types of administrative units. Like legal units, they are pivotal units between the administrative world and statistics. They are further discussed in Section 4.9.

⁴⁶ Based on UN Working Paper on Statistical Business Registers 2011

- *legal persons* whose existence is recognized by law independently of the individuals or institutions which may own them, or are members of them; and
- *natural persons* who are engaged in an economic activity in their own right as independent self-employed producers.

Thus defined, a legal unit is very close to the SNA 2008 notion of *institutional unit*, the main difference being that the latter includes households containing natural persons rather than individual natural persons. The definition is also close to the European Statistical System 1993 Units Regulation, which makes explicit that the only natural persons of interest are those with economic production.

A legal unit can be a single entity, controlled or not controlled by another legal unit, or, in some countries, controlled by a group of legal units under common control.

Legal units are the core concept in the SBR in the sense that they:

- create the operational units by which they manage themselves;
- register themselves (and possibly administrative sub-units) in response to administrative requirements; and
- are modelled in terms of standard statistical units in the SBR.

A legal unit always forms, either by itself or sometimes in combination with other (parts of) legal units, the starting point for defining an enterprise.

4.7.3 Active Legal Units

Some legal units are created for taxation or legal reasons and not with the intention of conducting actual economic activities. Thus, a legal unit may never have any discernible production, or may cease to have production. From the perspective of economic production and the SBR, the legal units of interest, termed *active legal units*, are those that:

- are conducting economic production activities; or
- have indicated an intention to conduct economic production activities; or
- have recently conducted economic production activities.

4.7.4 Legal Form

The legal form (also known as legal status) of a legal unit is useful:

- in avoiding ambiguity and/or double counting in identification searches;
- in stratification for surveys; and
- in defining the institutional sector.

Whilst the range of legal forms varies from country to country, depending upon national legislation, there are some legal forms that can be found in most countries⁴⁷.

The tax regime applicable to a legal unit depends on its legal form, which means that taxation data supplied to an SBR depends on legal form. Some legal forms may be exempted from certain regulations (VAT for example) in which case the corresponding legal units are not registered by the corresponding administrative sources, adversely affecting the coverage provided by these sources.

⁴⁷ As indicated in the Eurostat - OECD Manual on Business Demography Statistics - 2007 Edition-glossary.

EU Member States typically support the following legal forms

- *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 company*: enterprise that is a joint-stock company, limited partnership with share capital, or private limited company. (Harmonized rules at European level governing the publication of accounts for these types of companies are laid down by the Fourth Council Directive.)
- *Co-operative society*: body established by law in a country, observing a number of general principles, for example may be entitled to provide services only to members, profits often distributed in proportion to members' dealings with the society, etc.
- Non-profit making body.
- Enterprise with *other form of legal constitution*: this group includes nationalized industries, publicly-owned enterprises, and state and local authority monopolies.

4.7.5 Market and Non-Market Legal Units

The legal form of a legal unit is not the only determinant of the institutional sector to which it belongs. For example, some legally constituted public sector corporations may be non-market units and hence be part of general government rather than publicly-owned corporations. Also some units that are part of general government may be classified as producer units for national accounts purposes and hence represented as enterprises even though they are not actually independent legal units. They are necessarily *recognised* as independent legal units. In this case, the corresponding statistical unit in the SBR contributes to the (national, regional) accounts, *but is not identified as a legal unit*. Thus, both the *market-non-market* characteristics of a unit, and its *legal form* (or quasi-legal form if not recognised as independent legal unit) are needed to classify the statistical unit to the appropriate institutional sector.

Annex D1 indicates the sector classification of enterprises based on legal form and market/non orientation.

4.7.6 Quasi-corporations:

Some *unincorporated enterprises* belonging to households or government units may behave in much the same way as corporations. Such enterprises are treated as *quasi-corporations* provided they have a complete sets of accounts, even though they are not *legal units* in the strict sense of the word.

4.8 Administrative Units

4.8.1 *Introductory Remarks*

Lists of legal units per se are typically not readily available for statistical purposes. To maintain their SBRs, NSIs tend to use data about units from administrative sources.

Definition

Administrative units are established by the relevant administrative authorities in conjunction with the legal units to which they refer in order to apply the corresponding regulations. An administrative unit may or may not correspond directly to a legal unit.

4.8.2 *Linkage to Statistical Units*

The identification numbering system for units in an administrative source often does not correspond to, or align with, that used in the SBR. Also, whilst the legal unit is usually the unit recognised by administrative authorities, legal units may be identified differently in the various administrative databases. Consequently, one of the prerequisites in using administrative data is to establish rigorous linkage between the sets of administrative units maintained by administrative authorities and the statistical units in the SBR. Such linkage is necessary to ensure that it is possible to match and combine information from various administrative sources and that there is no duplication in the resulting coverage.

As further discussed in Section 11.4.11 a unique identifier should be assigned to each statistical unit in the SBR. These identifiers should be linked to the various administrative identifiers. The goal is to ensure that during the creation and maintenance of statistical units they are correctly linked to corresponding administrative units.

Integration of administrative data into the SBR is greatly facilitated in countries where a single identifier for legal units is used by all administrative authorities in identifying administrative units.

4.9 Local Unit of Legal Unit

4.9.1 *Introductory Remarks*

The term *local unit* may be understood in two different ways:

- referring to an enterprise or a part of an enterprise that engages in productive activity at or from one location, as defined in the SNA2008 and described in Section 4.6.3;
- referring to legal unit or a part of a legal unit that engages in productive activity at or from one location, in which case it is referred to as a *local unit of a legal unit*, as further described below.

Thus, in discussing a local unit, it is important to specify whether the unit is a local view relative to an enterprise or a legal unit. In either case local units may be used to satisfy users' needs for a geographic dimension, for example, for administrative geographic delineations, for precise localisation for societal/environmental considerations, or for geographic zones for harmonised statistics.

Although both types of local unit are described in this chapter in practice, for most countries, one or other of these types in the SBR is sufficient. In most cases, there is only one local unit in an enterprise, and this local unit and the local unit of the legal unit (to which the enterprise

is related) are in a one-to-one relationship. However, in some cases (often cases where the enterprise is large) the two types of local unit are different.

4.9.2 Definition of Local Unit of Legal Unit

A **local unit of legal unit** is defined as a legal unit, or a part of a legal unit (for example, a workshop, factory, warehouse, office, mine or depot) that engages in productive activity at or from one location. There is no restriction imposed on the activity which is carried out at the location.

4.9.3 Local Unit of Legal Unit in Practice

As noted above, a *local unit* may refer to a *local unit of a legal unit* (with a precise address) or to a statistical *local unit* that may, in practice, be a set of local legal units at different addresses within some geographical area. Usually, it is a local unit of legal unit rather than a statistical local unit that is maintained in the SBR. In particular this is the case when a unique (possibly multi-source) administrative identification number, linked to a legal unit identification number, is attributed to it.

In some rare cases, a local unit of a legal unit may be shared by enterprises.

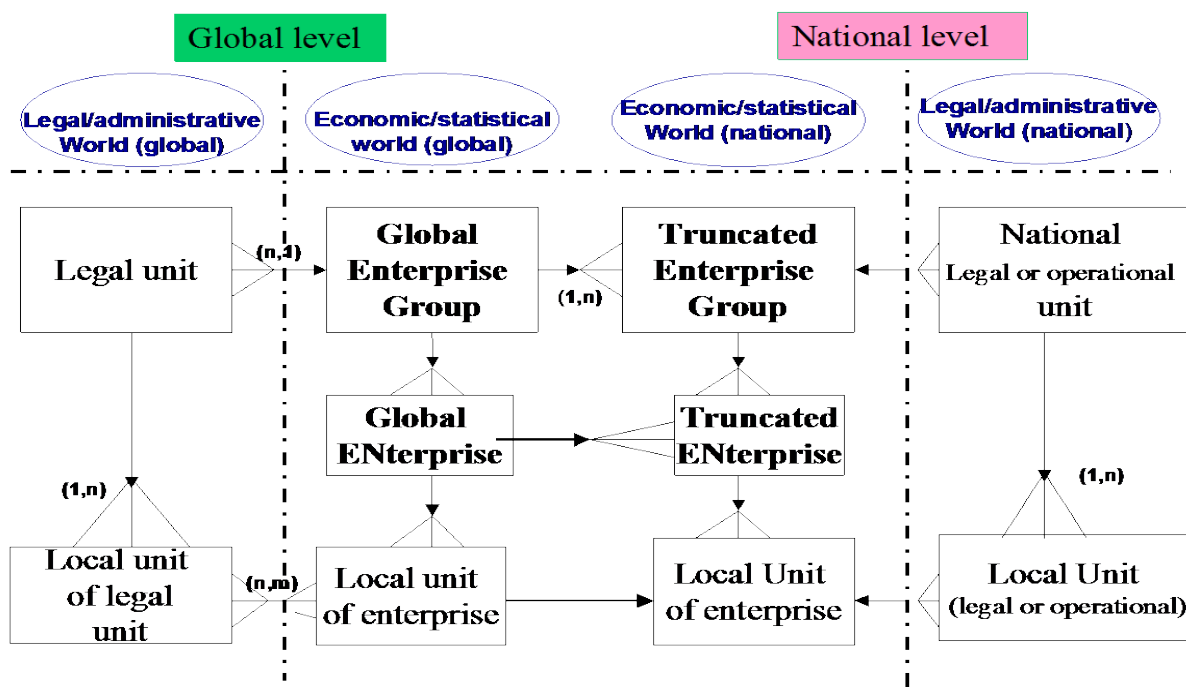
The actual definition of a local unit of legal unit is country dependent, depending on the local legislation.

Ancillary local units may be distinguished (as observation units) but they are consolidated with other statistical units in the statistical view.

4.10 Overall Model of Main Statistical and Administrative Units

The new definitions that the EU intends to approve, and that are in line with its research on profiling and consistency of the statistics in different domains, lead to the unit model illustrated in the Figure 4.4.

Figure 4.4: The Proposed EU Unit Model (Global and National Views)



This model provides a coherent set of units that can be used for compiling statistics in which data of MNEs are included in a consistent way among countries. For these units the delineation of the global enterprise (GEN) and its geographical (national) components, the truncated enterprises (TENs) is essential.

The GENs are the result of the economic and statistical analysis of the structure of the enterprise group at a global level, and each TEN is viewed purely as the trace of the GEN in a country in which the GEN is present. In this top-down analysis, a GEN *cannot have more than one TEN in one country*, just like a global enterprise group cannot have more than one truncated enterprise group in a country.

The standard statistical unit *establishment* does not appear in this figure.

4.11 Branches, Ancillary Activities, Special Purpose Entities, Franchises and Related Concepts

4.11.1 Introductory Remarks

This section deals with some issues that concern conceptual definitions of units and their practical delineation, that are under current review, scrutiny or research within the EU, and that have not been covered in earlier sections. Branches, ancillary units, and special purpose entities (SPEs), holding companies (HCs) and head offices (HOs) are discussed. Coordination is needed with the joint ECB/Eurostat/OECD task force on their statistical treatment. Descriptions of the related characteristics are in Chapter 5.

4.11.2 Branch

Definition

A branch is a local unit not constituting a separate legal unit in the country where it operates that is dependent on a foreign-controlled enterprise. Branches are treated as quasi-corporate enterprises.

Only branches with registered existence (as may be evidenced by a permanent address) are covered by foreign affiliate statistics. In particular, the Outward FATS statistical units are enterprises and all branches abroad that are controlled and managed by an institutional unit resident in the compiling economy (in the sense of their Global Decision Centre).

European example on branches

A new legal status has been created for the so-called “European companies” or “Europeae Societae”. This status allows the relevant legal units to operate in the whole of the EU without any legal registration in the Member States other than the initial one. It also allows a change of the place and country of registration through simply a decision of the board.

In consequence, according to EU methodology, locations of legal units in “other” countries are branches.

This status, which was rarely used in the recent past, seems to be now developing very quickly. It might create large problems in determining the enterprise group perimeter and country of residence.

4.11.3 Ancillary Activities

Recommendations regarding the treatment of ancillary activities and the units that perform them are given in the introduction of ISIC Rev.4 manual, as follows.

59. Principal and secondary activities cannot be carried out without the support of a number of ancillary activities, such as bookkeeping, transportation, storage, purchasing, sales promotion, cleaning, repair and maintenance, security etc. At least some of these activities are found in every economic entity. Thus, ancillary activities are those that are undertaken to support the main productive activities of an entity by providing goods or services entirely or primarily for the use of that entity...

62. If an establishment undertaking ancillary activities is statistically observable, in the sense that separate accounts for the production it undertakes are readily available, or if it is in a geographically different location from the establishments it serves, it may be desirable and useful to consider it as a separate unit and allocate it to the industrial classification corresponding to its principal activity. However, it is recommended that statisticians not make extraordinary efforts to create separate establishments for these activities artificially in the absence of suitable basic data being available.

63. Under the definition given in paragraph 59 above, the following activities are not to be considered ancillary:

(a) Producing goods or services as part of fixed capital formation. The type of units most affected are those doing construction work on the account of their parent unit. This approach is in accordance with the classification in ISIC of own account construction units for which data are available to the construction industry;

(b) Producing output which, although also used as intermediate consumption by the principal or secondary activity, is for the greater part sold on the market;

(c) Producing goods that become a physical part of the output of the principal or secondary activity (for example, the production of boxes, tin cans or the like by a department of an enterprise as packaging for its own products);

(d) Research and development activities, which are considered part of fixed capital formation in the context of SNA.

When a unit is to be classified by economic activity ancillary activities are not considered. For coding purposes only those activities that form the principal and secondary activities are taken into account. The costs for performing ancillary activities are allocated by the enterprise to its principal and secondary activities.

Usually an ancillary activity is not identified as a separate organisational unit or as a separate legal unit. However, the incidence of legal units that have been created only perform ancillary activities for the enterprises of an enterprise group is increasing.

The EU is also working on this topic, providing its most recent proposals on the treatment of “ancillary units”, via its task force of Statistical Units (Spring 2014) as detailed in the following example.

Complements to the proposed definition of the “enterprise” in Europe:

Operational rule n°7: Ancillary legal units/ancillary operating segments

If a legal unit or an operating segment performs one or more ancillary activities for other legal units or operating segments, it has to be considered as an ancillary legal unit or operating segment. In this case it is not considered an enterprise. The outputs of the ancillary legal unit or operating segment have to be considered as intermediate consumption for the other enterprises of the enterprise group and its data have to be consolidated within the enterprises which consume these outputs.

In case the output of the legal unit or segment, which performs one or more ancillary activities, is only partly consumed by other legal units or segments, and the legal unit or segment sells to the market on a regular basis, it should be treated as an enterprise.

A legal unit or a segment or part thereof located in one country may carry out exclusively ancillary activities inside an enterprise group and deliver its services to more than one enterprise of the enterprise group it belongs. If the enterprises that receive the ancillary services have locations in one or more other countries the legal unit or segment providing these services is by convention treated as an enterprise and is classified according to the activity it is performing.

(An operating segment as defined in IFRS is a component of an enterprise group that has discrete financial information available, and whose results are reviewed regularly by the entity’s chief operating decision maker for purposes of performance assessment and resource allocation. An operating segment manager is accountable to the chief operating decision maker for the results of the segment.)

Complements to the proposed definition of the “enterprise” in Europe (continued)

Operational rule n° 8: Serving legal units/operating segments

If a legal unit or an operating segment manages only factors of production such as land, buildings, equipment or staff for other legal units or operating segments it has to be considered as a serving legal unit or a serving operating segment. In this case it is not considered an enterprise. The data on the factors of production have to be consolidated within the enterprise(s).

In case data availability (including reliable estimations) does not allow for consolidating the output of serving units or segments within the enterprises, the serving units or segments have to be treated as enterprises and their output has to be valued at market prices.

In case the output of a legal unit or segment, which performs one or more serving activities, is only partly consumed within an enterprise, whereas the legal unit or operating segment also sells some part of its output to the market on a regular basis, it should be treated as an enterprise.

A legal unit or a segment or part thereof located in one country may carry out exclusively serving activities inside an enterprise group and deliver its services to more than one enterprise of the enterprise group it belongs. If the enterprises that receive the services have locations in one or more other countries the legal unit or segment providing these services is by convention treated as an enterprise and is classified according to the activity it is performing.

4.11.4 Special Purpose Entities

Special purpose entity (SPE) is the name used in the SNA2008 for identifying special cases of corporations. In various manuals they are also known as special purpose vehicles (SPVs), international business companies, financing subsidiaries, conduits, holding companies, shell companies, shelf companies, brass plate companies and so on. SNA2008, BPM6 and BD4 agree on the fact that there is no internationally standard definition of an SPE. Thus, the identification of institutional units that may be described as SPEs requires consideration of a number of characteristics summarized in Figure 4.5. Furthermore, a unit may be treated as an SPE even though it does not have all the features shown in Figure 4.5. For example ownership of non-financial assets may be allowed.

Different ways of treatment are expected in the BPM6 in where it is stated: “*Although there is no internationally standard definition of SPEs, in economies in which they are important they may be identified separately, according to either a national company law definition, or in terms of a functional definition, possibly referring to their limited physical presence and ownership by non-residents (4.87)*”.

In the SNA2008⁴⁸ a unit defined as an SPE is treated in the same way as any other institutional unit, being allocated to a sector and an industry according to its principal activity unless it falls into one of the three following categories:

⁴⁸ SNA2008: 4.58

- a) captive financial institutions,
- b) artificial subsidiaries of corporations,
- c) special purpose units of general government.

Figure 4.5 Characteristics of SPEs

SNA2008		BPM6		BD4	
Ref.	Characteristics	Ref.	Characteristics	Ref.	Characteristics
4.56	have no employees	4.50	have few or no employees	6.2 - 313	have little or no employment
4.56	have no non-financial assets	4.50	other parts of their balance sheets are claims on or liabilities to nonresidents		
4.56	they may have little physical presence	4.50	have little or no physical presence	6.2 - 313	have little physical presence in the jurisdiction
4.56	they are always related to another corporation, often as a subsidiary and are often resident in a territory other than the territory of residence of the related corporations	4.50	their owners are not residents of the territory of incorporation	6.2 - 313	their parent enterprises are typically located in other jurisdictions (economies)
4.57	Entities of this type are commonly managed by employees of another corporation which may or may not be a related one				
4.57	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.			6.2 - 313	They are often used as devices to raise capital or to hold assets and liabilities and usually do not undertake significant production
				6.2 - 313	no operations

ECB/Eurostat/OECD Task Force on Head Offices, Holding Companies and Special Purpose Entities (June 2013)

The Task Force identified 11 common types of SPEs, namely:

- 1-Holding company;
- 2-Shell company;
- 3-Unit for holding and managing wealth of individuals and families;
- 4-Securitisation company;
- 5-Conduit;
- 6-Royalty and licensing company;
- 7-Captive leasing company (including mobile equipment renting company);
- 8-Factoring and invoicing company;
- 9-Captive insurance company;
- 10-SPE carrying out other financial function;
- 11-Merchanting company

Aware of the difficulty of finding a common international definition, different task forces have tried to define decision trees to be used for allocating units between SPEs and normal units. In particular, it is possible to consider the following 4 main groups:

- Ownership of financial assets (Captive financial institutions : Sector S127):
 - Holding companies (ISIC Section K 6420);
 - Trusts, funds and similar financial entities (ISIC Section K 6430);
 - Securitization companies (ISIC Section K 6499 (Assuming that the relevant units pass the institutional test, they should be classified as part of S125 if they purchase assets on the open markets while raising funds on the open markets. If they do not operate in the open markets on either assets or liabilities, they should be classified in S127.))
 - Captive financial leasing companies (ISIC Section K 6491 (Financial leasing companies operating on open markets are to be classified under S125.))
 - Captive insurance and reinsurance companies (ISIC Section K 6512, 6520)
 - Invoicing companies (ISIC Section K 6499)
- Ownership of non-financial tangible assets
 - Renting of mobile equipment (ISIC Section N 7730 - Sector S11)
 - Merchandising companies (ISIC Section G – 46xx - Sector S11)
- Ownership of non-financial intangible assets
 - Licensing and royalty companies (ISIC Section M 7490 - Sector S11)
- Others
 - Offices of airlines in airport hubs abroad (ISIC Section H 5110 - Sector S11)

where Sector S11 is non-financial corporations, Sector S125 is other financial intermediaries except insurance corporations and pension funds, and Sector S127 is captive financial institutions and money lenders

ESSnet Project on Consistency

As SPEs are currently identified from a national perspective, the following proposals are to be further evaluated.

- *Terminology*: it is necessary to clarify the meaning of different labels and to use the same wording for the same things.
- *Characteristics*: a clear choice should be made among the terms “no”, “few”, “little”, and “often”.

The narrowest definition of an SPE could be based on the following features:

- no employees;
- no physical presence;
- no non-financial assets;
- always related to another corporation and their owners are not residents in the territory of incorporation; and
- no other subsidiaries in the country.

A SPE operating in the same country of another affiliate should be combined to form a single enterprise unless part of a different Global Enterprise (GEN) or out of the consolidation perimeter. A SPE operating in a country without other affiliates must be treated as an

institutional unit with its own characteristics and recorded in the BR with flags useful to the other statistical domains.

4.11.5 Distinction between Head Offices and Holding Companies

*“The statistical analysis of head office (HOs) and holding companies (HCs) and their delineation starts after the institutional unit test. Both types of units are often referred to as holding companies, because both of them have relations to other entities, their subsidiaries. However, the relationships are quite different: while a HO exercises managerial control over its subsidiaries, the HC does not undertake any management activities and its principal activity is simply owning a group of subsidiaries”.*⁴⁹

From a conceptual point of view the distinction between HOs and HCs seems clear. From an operational point of view, statisticians are not sure that the businesses such as multinationals are interested in this distinction for reasons other than legal. For those interested, mostly national accountants and the balance of payments statisticians, applying these concepts in practice is complex. HOs and HCs have a common operational characteristic: “Entity having at least 50% of its assets consisting of equity vis-à-vis its subsidiaries can be considered as a practical indicator for identification of such entities.”

Holding companies are one of the eleven common types of special purpose entities (SPEs) identified by the ECB/Eurostat/OECD Task Force on Head Offices, Holding Companies and Special Purpose Entities (June 2013), as noted above. The Task Force discussion on practical rules for distinguishing between HOs and HCs showed that information on characteristics like management control or auxiliary units are available for large units or large groups only. Wherever available, such data is especially important in view of the large proportion of the aggregate balance sheet that can be explained by these units. For units where such information is not available, or only available at great cost in practice, the distinction between HOs and HCs has to be based mainly on an employment criterion, pure HOs having no employment or a very low one.

Profilers, studying the structure of the large and complex multinationals, are more interested in the position of both HOs and HCs in the control tree than in the distinction between them.

ECB/Eurostat/OECD Task Force distinction between HOs and HCs

Head offices (HOs) are actively engaged in production. They may have noticeably fewer employees, and more at a senior (managerial) level, than their subsidiaries. However, having zero employment is a clear indication of not being a HO. On the other hand, holding companies (HCs) simply holding assets may do this with very few or without any employed personnel. However, HCs may have a limited number of employees for several reasons. For the delineation between HOs and HCs, national legislative requirements for the number of employees of HCs should be taken into account. In general, employment of three or more persons is a first indicator that a unit is an HO, rather than an HC.

ESSnet Profiling assessment in terms of control tree

When the holdings are situated upwards from the global decision centre in the control tree, they are not part of the profiled group, and usually belong to the financial sector.

~~When the holdings are located inside control sub-trees, they should be kept in the profiled group as part of the relevant enterprise.~~

⁴⁹ ECB/Eurostat/OECD Task Force on HOs, HCs and SPEs (June 2013)

4.11.6 Franchises

Franchise operators may or may not belong to the same enterprise group. The franchiser is always regarded as a separate enterprise.

Franchisees are deemed to be separate from the franchiser because they comprise a complete combination of factors of production, and they run the full entrepreneurial risk. Moreover, although the definition of the enterprise requires autonomy, it allows for this autonomy to be somewhat restricted (“a certain degree of autonomy” is required). Also, full accounts tend to be available only at the level of the separate franchisees.

Research is also taking place made on the concepts of *commercial networks* or of *networks of service producers* in order to be make it possible to compare fully integrated commercial groups (the cases where the franchiser and the franchisees have one and the same decision centre) and all other situations in which only a part of the network is fully integrated and the rest of it works under franchise.

5 Characteristics of Units

5.1 Introduction

This chapter describes the characteristics that should be recorded in the SBR for the various types of units defined and discussed in Chapter 4. Collectively, these characteristics identify the units, indicate how they can be contacted, and provide economic and geographic information, all time-stamped. They also contain links between different types of units, both within the SBR and across multiple sources of updating information. Attention is paid to the metadata documenting the characteristics and to the time stamping mechanisms (reference dates and updating dates).

The characteristics are grouped and described in accordance with the following classification:

- *identification and contact characteristics*, including not only names, addresses and communication means (telephone numbers, mail addresses, websites etc.) but also specific and unique identifiers (even internal identification numbers that enterprises use in monitoring their affiliates and that are not intended for the general public);
- *demographic characteristics*, including unit activity start and end dates, current status and continuity markers;
- *economic/stratification characteristics*, i.e. economic variables and classifications, used for sampling survey frames, grossing up sample data and/or presenting results;
- *relationships (links)* between units of different types within the SBR;
- *relationships (links)* with units in other registers and data sources, including their specific identification numbers.

Sections 5.2-5.6 provide general descriptions of the primary characteristics in each of the five groups. Precise details of whether, and if so how, each characteristic applies to each unit type are in Annex C.

The importance of a characteristic is summarised in terms of its *core/non-core* status.

- *Core* characteristics are those that are indispensable in maintaining and/or using the SBR. Most characteristics described in this chapter are core characteristics.
- *Non-core* characteristics are those that are optional, that can be added in a later stage, such as geographical coordinates to supplement addresses.

Identification numbers are very important characteristics because they are essential in merging micro data from various sources, including from administrative files.

Some characteristics need more explanation and guidance regarding use than others. These include economic activity, sector and regional classifications, and the size classes. Not only do these basic classification concepts have to be defined, but also the processes by which their values are acquired have to be described.

Chapters 6 and 7 detail the data sources used to establish and maintain the values of the characteristics. For every value, the associated metadata (including source, date of update, and reference date) should be recorded, as further discussed in Section 10.6.

The primary reference source for the chapter is Eurostat's *Business Registers Recommendations Manual*.

5.2 Identification and Contact Characteristics

Type of unit

Purpose: To identify the type of a unit.

Definition: Chapter 4 lists and defines the main types of units that may be included in the SBR. They are *enterprise group, enterprise, establishment, kind-of-activity-unit, local unit, legal unit, administrative unit, and local unit of legal unit.*

As noted in Chapter 4, some enterprise groups and enterprises are operating and managed globally. For purely national purposes they should be truncated to the national territory. Identification of truncated enterprise groups and enterprises does not involve anything more than splitting the multinational enterprise groups and enterprises on a national basis. Thus, double coding, i.e., coding at both national and global level, of some of the characteristics described below, in particular of principal economic activity code, is recommended.

Type of enterprise group

Purpose: To identify the type of enterprise group.

Definition: The type options are *all-resident, multinational domestically controlled, and multinational foreign controlled.*

Role of enterprise within enterprise group

Purpose: to identify the role of an enterprise within an enterprise group

Definitions: The type options are *management/control unit, global group head (controlling unit), global decision centre (managing unit), highest level consolidation unit, and other.*

Identification number(s)

Purpose: To identify the unit and to enable it to be linked to other units in the SBR and in administrative and other statistical sources.

It is preferable that no information on the content of the unit is embedded in the identification number, as further discussed in Section 11.4.11.

Name(s)

For legal persons: legal name, plus trading name(s) and acronym(s).

For natural persons: family name(s), names normally used and possible pseudonyms, forenames, personal number (if authorized by national law), gender.

For statistical units: trading name(s) and acronym(s).

For administrative units: names as registered.

Address

Physical address, including postcode and official geographic code at the most detailed level.

*Communication characteristics*⁵⁰

- Telephone and fax number(s).
- Electronic mail address(es).

⁵⁰ The communication characteristics of the unit may be different for different surveys. They may include the names and addresses of the corresponding contact persons. Also, as previously noted, the contact persons might be reporting units such as accounting or tax consultants hired by the unit.

- Information on sources of quick updates about the unit characteristics (such as web sites).
- Mailing address (if different from the legal one).

Purpose: These characteristics provide contact information for the unit. Telephone number and e-mail address are very important.

In the event that units from different sources do not share a common identification system, their names and addresses can be used for matching units across sources.

5.3 Demographic Characteristics

Date of incorporation of unit (for units that are legal persons)

Date of official recognition as an economic operator (for units that are natural persons)

Purpose: One or other of these characteristics should be recorded at the time of inclusion of a new unit, whether this is the result of real birth or of another demographic event that results in a new unit, for example the incorporation of an unincorporated enterprise.

Definition: The *date of official recognition* may be the date on which an identification number is assigned, or the date in which the legal existence is accepted, be it through a trade register number, the assignment of a VAT number, or other form of administrative registration.

Date of commencement of economic activity

Purpose: The characteristic is used in deciding if a new unit is economically active.

Definition: The *date of commencement of economic activity* should be the date on which the unit actually starts its economic activity.

As soon as a unit invests it is deemed to have economic activity even though production may not have started. If this were not the case, then investments, in particular gross fixed capital formation, could be heavily underestimated.

The actual *date of commencement of economic activity* is often difficult to know. A proxy for this date can be the date declared in advance by the unit at the time it is officially registered and an identification number is assigned. Another proxy is the date when its legal existence is accepted.

Date of final cessation of economic activity

Date on which a legal unit ceased to exist

Purpose: These characteristics are needed to record the *permanent inactivity* and/or *death* of a unit.

Death is a difficult status to establish with precision as liquidation can last a long time after economic activities have been stopped, particularly when lawsuits are in process.

Date of merger, take-over, split or break

Identifier of unit (if any) that is continuing the economic activity after the concerned unit has ceased to exist

Purpose: These demographic characteristics (further discussed in Sections 7.3 and 9.3) are needed to record the restructuring of units and to relate the resulting live units to the previous ones that have been ceased and might be considered dead.

5.4 Economic/Stratification Characteristics

Legal form

Purpose: Legal form is a key stratification characteristic. Legal form is primarily used in classifying legal persons, and hence the enterprises with which they are associated. For unincorporated enterprises there are different types of legal forms, including sole proprietorship and partnership.

Definition: The possible types of legal form differ across countries according to the particular legislation in each country. A typical set of legal forms for EU countries is listed in Section 4.7.4.

Institutional sector and sub-sector

Purpose: The institutional sector classification has at least five important functions:

- in the national accounts, for compiling data for individual sectors;
- in business statistics and business demography, for separating market and non-market activities;
- in separating data for public enterprises from data for private enterprises
- in finding the share of employment, or value added, in a country's economy of which the country itself is the ultimate controller;
- in distinguishing foreign controlled enterprises; and

The set of institutional unit types is described in Chapter 3 and the institutional sector classification is given in Annex D.

Principal economic activity

Purpose: The principal economic activity code of a unit is a key stratification characteristic.

Secondary economic activity (activities)

Practical guidance in assigning and maintaining principal and secondary economic activity codes is provided in Chapter 7.

Ancillary unit

Purpose: Some units perform activities exclusively for other units without autonomy. As a consequence, they have no economic significance per se. They should be seen as ancillary activities of the parent legal unit that they serve, to which they belong, and to which they must be attached to form a unit significant for economic analysis⁵¹.

Market orientation (market, non-market)

Purpose: Both the *market/non-market* characteristic of a unit and its *legal form* are needed to classify a statistical unit to the appropriate institutional sector⁵².

Turnover

Flag indicating consolidated turnover

Purpose: The SBR should record an estimate of economic size for every unit. For production statistics, *value added* is considered the most appropriate characteristic, but it is difficult to

⁵¹ Derived from Eurostat Business Registers Recommendations Manual: §4.11.3

⁵² Eurostat Business Registers Recommendations Manual: §4.7.5

measure and is often replaced by proxy measures. Because it is relatively easy to collect, the most commonly used monetary proxy is *turnover*. A commonly used physical proxy is *number of employed people*.

There are a number of drawbacks to using turnover due to the non-cancellation of internal flows between statistical or reporting units (cancellation that is part of the consolidation process).

- Turnover is not additive when changing the level of observation, for example from establishment to enterprise, or from enterprise to enterprise group, or from the usual observation unit (which is typically a legal unit) to enterprise or enterprise group. Consequently, the unit being measured must be very precisely specified.
- Turnover is not additive when changing the geographical level; it is different at global enterprise level than at truncated national enterprise level.

There are different definitions of turnover according to the economic activity of the unit being classified. For example, there are differing definitions for agriculture, forestry and fishing, manufacturing, finance and insurance, and economic activities of households as employers.

If the consolidated turnover of a statistical unit is not available it is better to use as proxy the sum of the turnovers of its constituent units than to have no measurement at all. As turnover of legal units is used in most administrative sources, it should be recorded for consistency purposes.

Number of persons employed

Number of employees

For all units the SBR should record 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)*. (FTE employment is defined as total hours worked divided by average annual hours worked in full-time jobs⁵³.)

Purpose: The main uses of these characteristics are in stratification for sampling, analysis and dissemination purposes.

- Number of persons employed is preferable for stratifying survey samples in the case of very small units.
- Employment data direct from the SBR may be used in compiling employment statistics. They are especially useful for small area statistics, where the SBR may be the only comprehensive source.

Number of persons employed is a core characteristic. It 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.

Number of employees is a non-core characteristic. It is defined as those persons who work for an employer and who have a contract of employment and receive compensation in the form

⁵³ SNA2008

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.

For stratification purposes the *number of employees at year end* and/or the *annual average number of employees* may be used. An annual average can also be calculated over a specified period within a year for seasonally active units.

The number of employees in FTEs may be calculated for the period during which the unit is active and/or for a full year. For stratification purposes the active period should be used while for annual economic statistics the whole year is more appropriate.

Proxies: In some countries all these data can be obtained directly from administrative sources while in other countries administrative sources may provide only number of paid employees. In the latter case, the number of persons employed can be estimated according to legal form and activity. For example:

- for sole proprietors, number of persons employed = number of employees+1;
- for two person partnerships, number of persons employed = number of employees+2.

Sources: Administrative sources, surveys, and SBR calculations.

Comments: The reference period used for the measurement of employment in business demography is a calendar year. Thus, in principle, a labour force estimate should be an annual average. If a unit operates only during a part of the year (as in the case of a seasonal enterprise or a new enterprise), the average for size measurement should be based on that period. However, for annual employment data the average should be calculated for the whole calendar year. How the annual average is calculated in practice depends on the updating frequency of the SBR and on the sources used⁵⁴. For example, the value may be approximated by using the number of persons employed at a specified moment during the year if this is all that is available.

Head counts and FTEs each have certain advantages, so both should be recorded if possible. Head count is the number of physical persons, full-time and part-time, employed by a unit. FTE is a more accurate measure of labour input but depending upon the origin of the data, it may not be possible to calculate FTEs. Also, as the concept of *full-time* may vary from one country to another, the FTE data may not be fully comparable across countries.

Another possibility is to use *hours worked* directly as a measure of work input. This measurement is gaining favour in employment statistics, but data are not often readily available.

5.5 Relationships and Links between Units in the SBR

There are two main reasons for registering relationships between the various types of units in the SBR

- for management and the direct use in the SBR itself and by surveys using frames derived from the SBR – relationships are required to link together the units that belong to the same enterprise or to the same enterprise group, and to ensure correct

⁵⁴ Fiscal sources are related to the country fiscal year which can be different from the calendar year. Their average employment refers then to a different period.

transformation of data collected from administrative units, or about observation units, to data for the corresponding statistical units;

- for potential uses of the SBR - relationships assist in statistical and economic studies that depend upon bringing data from different sources together.

The most evident need concerns links between statistical units and administrative units required for creation and management of the SBR, and between statistical units and observation units. Depending on the sources used in building the SBR, and on its priority uses, there is also a need to register links between administrative units and legal units, and between observation units and reporting units.

Each link should include the date when the relationship started and (if applicable) the date when it stopped.

In its simplest form a relationship is a link including only two identification numbers (one for each unit linked) and the date(s). In some situations more complex relationships are recorded. (For example, a single unit of one type of unit may be linked to n units of another type, and the relationship may indicate the percentages to be applied in splitting data for the one unit across the n units.) This depends on how the SBR is organized. In general it is best to keep the relationships as simple as possible.

It is important to recognise the difference in the treatment of administrative registers and the treatment of the SBR. In an administrative register units and their values may be changed without notice to the NSI and without continuity from a statistical perspective. It is thus essential to keep copies of the administrative registers used as sources by the SBR, and, where possible and relevant, to record the relationships between the units in these registers and legal units and statistical units.

Links should also be kept between the observation units and the corresponding reporting units. For example, in some cases a single lawyer or accountant may report on behalf of several enterprises. These links can be maintained in the SBR or in survey specific data collection systems.

5.6 Relationships and Links with Other Registers

Whether the SBR contains data for all the different unit types in a single database, or in separate satellite registers, such links are very important given the emerging role of the SBR in connecting various data sources and thus integrating statistics.

Links are expressed in terms of identification numbers for the external sources and the identification numbers of the corresponding units in the SBR.

Ideally, the links should be organized in such a way that all administrative units are linked to the legal unit to which they belong. For example a legal unit may have several administrative units for pay-as-you-earn (PAYE) purposes, and several more for social security contributions, as well as having one or more VAT accounts and an account for payment of income tax.

Depending on national legislation, it is conceivable that a single administrative unit is linked to more than one legal unit.

The sorts of relationships that exist and links that are possible depend upon country legislation and regulations. Examples are:

- links from legal unit to value added tax (VAT);
- links from legal unit to trade register;

- links from legal unit to social security;
- links to balance sheet data (for units required to publish accounts);
- links to balance of payments register or foreign direct investment register;
- links to the farm register;
- links to other satellite registers, for example, shop registers, tourist establishment registers, transport registers, educational services registers, and health services registers.

5.7 Characteristics by Unit Type

Figure 5.1 summarises how the characteristics described in general terms in the previous sections are applied to each of the types of statistical units defined in Chapter 4:

- ✓ indicates where the unit (in the column) should have the characteristic (in the row);
- ✓? indicates where it is desirable but not compulsory that the unit has the characteristic; in other words, the unit may or may not have the characteristic;
- x indicates where no characteristic can exist

Comprehensive details are provided in Annex C.

Figure 5.1 Summary of Characteristics by Unit Type

Characteristics of the unit	Enterprise group / truncated enterprise group.	Enterprise / truncated enterprise	Establishment	Local unit (of enterprise)	Legal unit	Local unit of legal unit	Other observation unit ⁵⁵
Identification and contact							
Type of unit	✓	✓	✓	✓	✓	✓	✓
Type of enterprise group	✓						
ID number	✓	✓	✓	✓	✓	✓	✓
Name, address	✓	✓	✓	✓	✓	✓	
Communication characteristics, contact details	✓	✓	✓	✓	✓	✓	✓
Administrative identification numbers (tax number...)					✓	✓	
Geographical code (zip code...)	✓	✓	✓	✓	✓	✓	✓?
Country code	✓	✓	✓	✓	✓	✓	✓
Demographic							
Birth date	✓	✓	✓	✓	✓	✓	

⁵⁵ VAT / Social security specific units etc.

Characteristics of the unit	Enterprise group / truncated enterprise group.	Enterprise / truncated enterprise	Establishment	Local unit (of enterprise)	Legal unit	Local unit of legal unit	Other observation unit ⁵⁵
Status of unit (active/inactive....)	✓	✓	✓	✓	✓	✓	
Restructuring code (merger, split...)	✓	✓			✓		
Restructuring date	✓	✓			✓		
Death date	✓	✓	✓	✓	✓	✓	
Economic/stratification							
Legal form					✓		
Institutional sector		✓			✓		
Principal activity	✓	✓	✓	✓	✓	✓	
Secondary activity (if any)	✓	✓	✓	✓	✓	✓	
Market oriented		✓	✓?	✓?	✓	✓	✓
Ancillary unit		✓	✓	✓	✓	✓	
Employment	✓	✓	✓	✓	✓	✓	
Consolidated turnover	✓	✓	✓				
Turnover	*	*			✓		
Links and external references							
Link between unit and enterprise group		✓			✓		✓
Percentage control					✓		
Percentage interest (ownership)					✓		
Type of management/control unit (within group)					✓		✓
Link between unit and enterprise			✓	✓	✓	✓	✓
Link between unit and establishment		✓		✓	✓	✓	✓
Link between unit and local unit (of enterprise)		✓	✓		✓	✓	
Link between unit and legal unit	✓	✓				✓	✓
Link between unit and local unit of legal unit		✓	✓	✓	✓		✓
Links between units involved in mergers, amalgamations, splits	✓	✓	✓	✓	✓	✓	✓
External register - name of register, identification number(s)			✓		✓	✓	✓

6 Data Sources for the SBR

6.1 Introduction

Context

This chapter provides guidance on the sources of data that can be used to construct an SBR and how they are used in combination. The main focus is on administrative data, but statistical sources and new potential sources are also discussed. The chapter defines what is meant by *administrative data* and describes main administrative sources that are used in many countries. The advantages and disadvantages of these sources are considered, also the legal aspects pertinent to their use.

Different data sources may be used in combination to construct and maintain an SBR with good coverage and content. The data sources must also provide information about changes in units to keep coverage and content up-to-date, as further discussed in Chapter 7. Unless the SBR is maintained on a regular basis, it quickly loses its value by becoming outdated and ceasing to adequately reflect economic activity in the real world.

Data sources may be grouped into three categories:

- administrative sources;
- statistical sources, including feedback from economic surveys and SBR improvement surveys; and
- other sources, for example data from private data suppliers, telephone directories and the Internet.

The Guidelines recommend that SBRs be created and maintained primarily using administrative sources. Typically they enable good coverage and stability. This approach is in line with Principle 5 of the United Nations Fundamental Principles of Official Statistics, which states “*Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents*”.

However, the best choice of data sources depends on the specific situation in any given country, including the availability of administrative data and the scope and complexity of the national statistical system itself.

Administrative sources

A commonly applied definition in Europe is: “*Administrative sources are sources containing information that is not primarily collected for statistical purposes*”. Another EU definition⁵⁶ defines administrative data as: “*data derived from an administrative source, before any processing or validation by the NSI.*” These definitions allow for the wide interpretation of the data that might be included under the heading *administrative data* in these Guidelines. A more traditional and narrower definition requires administrative data *to be collected by government bodies for the purpose of administering taxes, public pension funds and other regulations*.

Eurostat collects information on administrative and other data sources used for SBRs in its annual business registers questionnaire. The same questionnaire is also used by the UNECE

⁵⁶ From ESSnet on Administrative data, an EU funded project

and the OECD and so covers a wide range of countries. The most commonly used sources relate to taxation systems, such as value added tax (VAT) and personal income tax, and to compulsory business registration systems, often administered by chambers of commerce or government bodies and departments.

Statistical sources and survey feedback

Statistical sources refer to data collected by statistical processes carried out by the NSI and, in principle, by other producers of statistics within the national statistical system. The two primary groups of sources are, first, the economic surveys that draw their frames from the SBR and that can provide updates and, second, improvement surveys conducted by the SBR itself.

Commercial and other data sources

The Guidelines consider that commercial data available from private companies that manage utilities, for example, telephone companies, water or power supply, may be considered in the same way as administrative data. Use of such data can be a cost-effective way of improving the SBR since these databases are extensively researched and managed for commercial benefit by their owners.

Other sources include data collected by companies for their own commercial purposes, for example data owned by market information companies and companies that collect data and maintain directories and databases that they provide to their customers for commercial purposes.

New data sources are emerging and current discussion of *big data* may well highlight additional sources. Such sources may be used in SBRs, for example in helping to create enterprise groups.

Content of chapter

In the following sections each of the various types of sources is discussed in more detail and the final section describes the record linkage procedures that are required to combine data from them. Influential reference documents include Eurostat's *Business Registers Recommendations Manual* and the African Development Bank's *Guidelines for Building Statistical Business Register in Africa*.

6.2 General Methods, Procedures and Issues

6.2.1 Cooperation with Data Providers

Establishing and maintaining good relations with the owners of data sources, especially administrative sources, is vital, and can be complex and challenging. From an SBR perspective the purposes of cooperation are to understand the concepts, to ensure continuity of supply, and to ensure easy linkage of data from the various sources to the SBR, preferably through a common identifier. In other words, the aim is to promote the interests of the SBR and, more generally, of the economic statistics program. The purposes are the same whether the SBR is dealing with an administrative, statistical or commercial data source.

Although cooperation can take on many forms, it is recommended that it includes regular face-to-face contact and mutual visits. Contact should involve both management level staff and operational staff who are working with the data on a daily basis.

Through effective cooperation it is possible to be proactive regarding changes in data sources that might be detrimental to the SBR, and to limit undesirable changes, or at least to have early warning of upcoming changes that could disrupt, jeopardise or otherwise challenge

statistical production or accuracy. Cooperation should also be used to promote the benefits of using statistical concepts in administrative sources, for example, standard economic activity codes and standard definitions of business units.

To provide clear responsibility and working arrangements it is good practice to set up a memorandum of understanding (MOU) or service level agreement (SLA) with an administrative source. Whilst not often legally binding, an MOU or SLA provides an agreed framework and gives reassurance about the services to be received. Typically an SLA sets out the agreed data coverage and content to be supplied, delivery timetables, data security provisions, quality provisions, and response times within which to answer validation questions by the SBR.

6.2.2 *Common Identification Systems*

A common identification system eases the work in combining data from administrative sources. If there is no common administrative identification system, the NSI should use any opportunity to promote one and to draw attention to the advantages of linking data in terms of more accurate and less costly statistics.

Linking different data sources is not only a means of improving the quality of statistics but also of reducing administration costs and burden. If a common administrative identification system does not exist, the NSI should construct an internal linkage table (containing links between units in the various sources) in order to manage the SBR and to reduce duplication or omission problems.

6.2.3 *Identifying and Using External Data Sources*

The first step in using administrative data is to identify the most useful data source(s). Potential sources have to be evaluated in terms of their coverage, content and the costs to the SBR of acquiring the data. For evaluation purposes it is important to obtain all available information about each source, for example, the definitions that are used, the coverage, the updating methods, the frequency of updates, the time lag, and how frequently it is possible to get information from the source.

Data obtained from external sources should be stored in the SBR without change of content. Retaining the original data ensures there is a record of what was received, regardless of changes made during subsequent administrative source processing or by SBR updating processes. This is important for audit purposes and for communication with the administrative sources should there be any issues or concerns about changes that have been made. It also enables the SBR to be rebuilt if any system or processing issues occur during production and updating processes. Administrative data are often extracted from live databases, and cannot be replicated for the same reference period at a later date. So keeping the administrative data in its original state is good practice.

Storing metadata about changes in data values is essential. For example, if a unit changes address in an administrative register, the old address should be retained in the SBR but with a termination date indicating it is no longer current, and the new address should be stored with the commencement date and a blank termination date.

Following receipt and storage of administrative data, statistical units in the SBR are created or updated through a set of conversion rules and procedures, as further described below.

6.2.4 Combined Administrative and Statistical Register

Some countries have developed a single administrative register, with multiple functions, including that of an SBR. This has significant advantages in that statistical data requirements are within the core of the administrative systems. An example of this type of register is the Norwegian Business Register (as further described in Annex E1). Other countries have similar systems.

6.2.5 Linking and Matching

The SBR is more resilient with lower risk of error, if multiple reliable data sources are used. Some sources, such as VAT, provide extensive (but never complete) coverage of the economy, whereas others may only cover one sector, e.g. financial, but more completely. Thus there are significant advantages to using data from several sources. However using two or more data sources presents two types of linking and matching challenges (as further discussed in Section 6.8):

- first, where the sources have a similar (and thus duplicated) coverage of the economy, the challenge is avoiding duplication without incurring omission; and
- second, where each source covers only a limited sector of the economy, such as finance and public administration, the challenge is dealing with the sector boundaries associated with each source.

If there is no common identification code, a probabilistic approach to linking can be adopted, based on similarities in the name and address or other characteristics, such as legal form and economic activity code, of the units in the different sources. This process is generally referred to as *matching*. However, it can lead to units being linked in error, i.e. *mismatches*, (also called *false* or *incorrect matches*), as well as units not being linked, i.e., *missed matches*, which result in duplication. Experience over the years in several NSIs suggests that linking large datasets in the absence of a common identification code is difficult and requires substantial investment in software and systems. Matching is further discussed in Section 6.9.

6.2.6 Transforming Administrative Units to Statistical Units

After administrative data have been acquired, the next step is to identify the links between the administrative units and the corresponding statistical units (usually enterprises) in the SBR. This is relatively easy if a common identification number is available, and there are no errors in this number in either the administrative sources or the SBR. As discussed in Chapter 4, typically an administrative unit coincides with a single legal unit, the productive activities of which are represented by a single enterprise. However, there are occasions where a single legal unit coincides with several administrative units and/or where several legal units correspond to an enterprise. In the former case it may be possible to use administrative units as the basis for identifying the breakdown of an enterprise into local units or establishments

The matching results should be quality checked. Checking may involve comparison of the values of characteristics such as economic activity code, size or legal form across the linked units. If these values are consistent, it is more likely that the link is correct. If not, further clerical checks may be needed, particularly where larger units are concerned. Also if additional data, such as turnover and imports/exports are available from other administrative sources, these can also be used in the checking procedure.

It is also useful to periodically check administrative units that have not been matched to statistical units and to attempt to establish further links or to determine why they do not

match. If the non-matched units do represent active economic units, the failure to match may be due to timing or scope differences between the administrative source and the SBR.

6.3 Administrative Data Sources

6.3.1 Typical administrative sources

Some of the commonly used data sources are listed below. The list is not complete; other data sources may be available.

- *Business registration/licence register.* In some countries it is compulsory to register/license a business before trading. If available, this register can provide basic information on identification number, name, address and other contact information. This sort of register may be run by the tax authority, Chamber of Commerce, licensing office, or another public authority.
- *Tax registers.* Typically relating to VAT or employee income tax, tax registers may be a source of data on economic activity, turnover, and activity status.
- *Company/trade associations and chambers of commerce registers.* They can provide information on economic activity, legal form, and births and deaths.
- *Social security registers.* For businesses employing paid staff and making social contributions for employees, a social security register can provide identifying characteristics and economic/stratification characteristics such as legal form, and number of employees.
- *Labour and employment registers.* Such registers can provide additional economic and social information about employees.
- *Government units registers.* Such registers are maintained by government finance departments for financial management of the public sector. .
- *Non-profit unit registers.* Such registers are maintained by regulators, for example of charities and typically contain name, address, other contact information legal status and possibly economic activity code.
- *Industry association registers.* Such registers may contain name, address, other contact information and economic activity code. They are likely to be up to date, but only contain members of the association, so completeness may be an issue. .
- *Agricultural administrative registers.* Such registers may cover agricultural holdings as distinct from businesses. They typically contain name, address, other contact information and indicators of economic activity and possibly size.
- *Water supply and electric association registers.* Such registers are by public or private utility bodies. They typically contain name, address, other contact information and indicators of economic activity and possibly size.
- *Sector specific sources.* Sources include lists of schools from the education ministry, lists of hospitals from the health ministry, and lists of charities from regulators. Whilst coverage is limited to a specific sector, within that sector it can be very comprehensive.
- *Central banks.* Central banks often have information for the financial sector, and on units engaged in foreign direct investments, from supervisory authorities.
- *Published business accounts.* Data from financial reports are particularly valuable as they contain information on shareholders and subsidiaries that is essential in delineating enterprise groups. NSIs are encouraged to explore the possibilities of automatically

extracting data from internal financial or management accounting systems of businesses, for example, using XBRL.

Besides providing names and contact information, these sources may also contain data that indicate whether or not a unit is active, its principal economic activity, its size and some other variables relevant for the SBR, such as employment and turnover.

6.3.2 Advantages of Using Administrative Data

The advantages of using administrative data to supplement or replace survey data wherever possible are well known. They are particularly pertinent to SBR construction and maintenance, as outlined in the following paragraphs.

Coverage

Use of administrative data in place of survey data eliminates sampling error, removes or significantly reduces non-response and provides more accurate and detailed estimates for various sub-populations, e.g. small geographic areas. This is because administrative sources often give complete, or almost complete, coverage of a target population, whereas sample surveys often cover only a relatively small proportion directly.

Coverage is of great significance from an SBR viewpoint, given its aim of including all economically active units. Thus, the main advantage of using an administrative source is the level of coverage it provides, which is either complete, or, if not complete, at least well defined, so that it can be assessed against the target population.

Costs

Another advantage administrative data offer over survey data is comparatively lower cost. Surveys are expensive, particularly if they are conducted as censuses or involve the use of personal interviews. Administrative data are often available free of charge, or for the marginal cost of extraction, particularly if they originate from the public sector. Even if there is a charge, it is often cheaper to use administrative data than collect the same information by survey. Fewer staff are usually needed to process the data and there is no need for non-response follow-up.

The size and scope of an SBR makes it very difficult and expensive to populate and maintain solely by statistical data. A periodic economic census would be required to build an SBR, which would be very expensive, as would be the complementary intercensal maintenance procedures.

Response burden

Using data from administrative sources involves no additional response burden. Whilst businesses usually understand the reasons for supplying data for registration and taxation purposes, even if they do not like doing so, they may see statistical data requests as an extra, less necessary, burden. If they have already provided details to other government departments, they may become annoyed at receiving requests for similar information from the NSI.

An associated advantage is that the use of administrative data may, in some cases, allow statistics to be produced more frequently, with no extra cost to businesses.

Content and Timeliness

The use of administrative sources may increase the quality of the SBR by providing access to more up-to-date information on key characteristics, such as:

- name and address;

- births and deaths of units, and the dates of these events;
- economic activity code;
- location;
- size, in terms of number of employees and/or turnover.

As well as improving the timeliness of SBR data in the context of their use for survey frames, administrative data can improve the timeliness of statistics that are derived from the SBR. This is because surveys take time to plan, to design, to pilot questionnaires, to analyse the population and optimise the sample, to collect and process the data, etc. Access to a suitable administrative source via the SBR may provide a quicker and more efficient solution.

6.3.3 Disadvantages of using administrative data

The following paragraphs outline some of problems encountered in using administrative data for the SBR and methods that may be used to address them.

Administrative and statistical definitions differ

Administrative data are collected for a specific administrative purpose, and the corresponding needs and priorities are likely to be different from those of the statistical system. Thus administrative units may not coincide with statistical units, and their characteristics may be differently defined. For example, the tax authorities may permit to a single enterprise to have multiple value added tax (VAT) accounts. Turnover collected for VAT purposes may not include turnover related to the sales of VAT exempt goods and services, whereas the statistical system wants total turnover.

Similarly, the classification systems used within administrative sources may be different to those used by the NSI. For instance, a register within a food safety authority might classify a department store according to its food department. Even if the classification systems are the same, they may be applied differently, depending on the primary purpose of the administrative source.

Where classification systems are different, it is necessary to construct conversion matrices to map the codes in the administrative classification onto those required for the SBR. Such mappings may be one-to-one, one-to-many, or many-to-many. In the latter two cases, a probabilistic allocation can be used, but this is not desirable as it does not give accurate codes for the individual units, which is needed for the SBR in its primary role of providing survey frames

Timeliness

Another common problem encountered when using data from administrative sources relates to timeliness. Administrative data may not be available in time to meet statistical needs, or may relate to a period which does not coincide with that required for statistical purposes. For example, the income tax year may not coincide with the calendar year required for structural business statistics.

Furthermore, there is generally some sort of lag between an event happening in the real economic world and it being recorded by an administrative source. This is then followed by a further lag before the administrative data become available to the SBR. Lags in recording births and deaths of enterprises are a major source of SBR coverage errors. If these lags can be measured, allowance can be made for them in statistics based on SBR data.

Changes to administrative regulations or procedures

Public sector administrative sources are set up for the purposes of administering regulations, for example collecting taxes or for monitoring government policies. This means that they are susceptible to political changes. If a policy changes, administrative sources may be affected in terms of coverage, definitions, thresholds, etc. They may even be abolished completely. Such changes may happen quite suddenly, with little warning. Thus, reliance on administrative sources always carries a certain degree of risk in terms of continuity of supply. These risks can be mitigated by legal or contractual provisions, by obtaining early warning of impending changes through regular contact with those responsible for the sources, and by drawing up (and implementing where necessary) contingency plans.

6.3.4 Monitoring the Quality of Administrative Data

Create knowledge of the administrative sources

Although administrative data have many advantages, it is important to invest time in understanding and monitoring their quality. It can be useful to compare sources in terms of coverage of their units and accuracy of their characteristics. The closeness of administrative units and characteristics to what are required for an SBR is an important factor in determining the quality of an administrative source. An SBR improvement survey may be required to determine the values of certain characteristics for which values provided by administrative sources are not considered appropriate.

The starting point in assessing the quality of an administrative source is to build up a thorough knowledge of the source, including its primary purpose and the way the data are collected and processed. Thorough understanding of a source allows a more accurate assessment of its strengths and weaknesses. To help develop and document this knowledge, particularly for the benefit of future SBR staff, it is useful to develop some form of template to record information from the source on contacts, units, characteristics, quality and formats. As noted in Section 6.2.1, this information should be discussed with the administrative authorities and should be reviewed regularly.

Quality indicators

Regular monitoring of quality is important. A full range of quality indicators should be developed. Some examples of quality indicators are:

- the number and proportion of enterprises lacking a valid and complete economic activity code;
- the number and proportion of enterprises for which the activity status (active, dormant, dead, etc.) is unknown;
- the number and proportion of enterprises lacking a complete address.

A more comprehensive example of good practice is provided in Annex F3.

Such indicators may be compiled using feedback from surveys based on the SBR and/or from SBR improvement surveys.

Dealing with conflicts

Where an administrative source and the SBR do not agree, the reasons should be investigated with the aim of gaining a better understanding of SBR quality. Surveys may be used to investigate such discrepancies. They may be conducted, either specifically for this purpose or as part of some other data collection exercise. The investigations should help in determining the appropriate rules and priorities for updating from different sources.

There may also be conflict in data from different administrative sources, for example conflicting industry codes. Procedures and rules need to be developed to resolve these problems. This could involve verifying the data by contacting the enterprise, or undertaking analytical work to determine which source is most reliable. The goal is a set of general rules to deal with conflicts.

6.3.5 Legal Issues and Access to Administrative Sources

Access guaranteed through Statistics Act

The first step in use of administrative data is to ensure the NSI has access to the data. There are two aspects to gaining access: the first is the legal framework, the second is setting up and implementing the procedures for transfer of the data. Whilst the legal framework for statistics varies from country to country, the preferred approach is for the NSI's right of access to administrative data to be enshrined in a general statistics act. Transfer of data is then simply a technical issue, with strict rules on security to be followed.

In many countries, the NSI does not have legislated access to data. There may even be confidentiality legislation preventing the transfer of administrative data to the NSI, even for statistical purposes. This can prove to be a major problem in establishing a good SBR, and one that may not be solved quickly. Whilst the best approach is likely to vary according to the country, in any case the NSI should consider initiating and introducing changes to its statistical legislation to ensure access is guaranteed by law.

Formal agreements with administrative data suppliers

As previously noted in Section 6.2.1, whether or not the NSI has legislated access, the NSI should try to establish some form of formal agreement with administrative data providers. This can be a legally binding *contract* with a private sector supplier, or a *service level agreement* or a *memorandum of understanding* within a public sector provider. Such agreements should describe the rights and responsibilities of both parties, delivery flows, data confidentiality constraints, quality standards, frequency and format of data transfer, time frames for responding to queries and questions about the data, and procedures to follow in case of disputes.

Building relationships with administrative data suppliers

In addition to formal arrangements, good working relationships with administrative data providers should be developed. These can be achieved through regular contact, preferably face-to-face. It is usually worth devoting time to visiting suppliers to gain greater knowledge of their work, systems and constraints. This leads to a better appreciation of quality issues and can help to build goodwill and mutual understanding, which in turn helps to identify and resolve problems as they arise and before they escalate.

In some countries, administrative agencies may be willing to release data to the NSI. In other countries, there may be reluctance. It is very difficult to deal with such concerns, but possible approaches include the publication of clear limits and rules regarding the NSI's use of data, thus ensuring businesses understand that sensitive data will not be fed back to other parts of government (particularly tax agencies).

It is advisable that the NSI should be coordinated when it approaches administrative sources: It should not be the case that different departments of the NSI approach the same source in order to receive more or less the same data. There should be coordination within the NSI. In this context the SBR, as the primarily user of such data, has a specific role.

6.4 Using Administrative Sources in Practice

6.4.1 Introductory Remark

This section provides guidelines on the steps to be taken when creating and maintaining a SBR primarily using administrative data. The precise sources used depend on what sources are available and on data access. The following basic principles apply whatever the sources.

6.4.2 Keeping Administrative and Statistical Data Separate

Data about administrative units received from administrative systems should be stored separately and not mixed with data for statistical units.

Even in the case of combined administrative and statistical registers (see paragraph 6.2.4) it is important to maintain clear procedures on data sources and updates to enable data received from administrative records and any statistical transformations to be auditable.

Experience shows that it is much easier to discuss unusual/anomalous SBR data either with the enterprises themselves (who should know what information has been given to the administrative sources) or with the administrative sources directly, rather than to try and solve the problems within the SBR. Keeping information separate allows specific problems to be discussed and concrete examples to be provided if investigation is required. The particular sources of data that should be used in the SBR, both administrative and statistical, depend upon the outcome of such investigations and the feedback received from the relevant sources.

Another possible reason for separating administrative and statistical data is where administrative data are being used for other statistical purposes. For example social security data may not only be used for SBR purposes, but also provide employment data for a variety of statistics.

6.4.3 Establishing Unique Identifiers

The establishment of unique identifiers (as further discussed in Section 11.4.11) is essential for the accurate maintenance of the SBR. Use of administrative numbers for statistical units may lead to added complexity in SBR systems, and introduce risks such as duplication and omission of statistical units. Thus an identification numbering system should be created covering each type of statistical unit.

Where multiple administrative sources are used there may be several different identifiers, one for each source. In this case (as further discussed in Section 6.8) significant resources may be required to match and create unique enterprises using a combination of data from these sources. If the NSI has influence on the development of administrative data, it should promote use of a unique business number for all relevant administrative processes in the country. Whilst this may not be easily achievable in practice, there are countries where it has happened.

Administrative units – if they are not already in one to one correspondence with legal units – should first be linked to the corresponding legal units. The most complex situation is the situation where a legal unit may have several administrative units within the same administrative source and hence have several identification numbers in that source. Furthermore, the legal unit may close one or more of these administrative units or ask for more to be created. It is important to handle these situations carefully. Usually, it is possible

to identify the “main” unit in the administrative register as the one to be used to provide information about the legal unit.

It is important to see the difference between legal units that are economic producers and legal units that are non-productive but may own other legal units in an inter-corporate ownership structure. A legal unit is, by law, not legally responsible for another legal unit even though it may be part of the same enterprise group and control the other legal unit. Therefore it is vitally important to create all legal units within an enterprise group even though some of them may not be readily identified from administrative sources.

6.4.4 Using Sector Specific and Additional Sources

As previously noted, some trade associations, charity regulators and government ministries such as education and health may maintain data that are useful in providing coverage of certain sectors of the economy and/or additional content. However, care needs to be taken to ensure that coverage of the particular sector is complete, or if not, the gaps are understood, and that the possibilities of duplication and omission are carefully considered. Also, using multiple sources leads to challenges in maintenance and management. In some cases a source may be used to verify data from another source; in other cases a source may contribute directly to updating of statistical units.

Examples are:

- the characteristic *sales space* may be available from an industry association for retail trade enterprises;
- the characteristics *category/number of stars* and *number of beds* for hotels may be available from a tourism authority or bureau;
- import/export data may be available from the customs authority.

Such administrative sources may be used in data confrontation as well as in validating statistical units. They can increase sampling efficiency (i.e., reduce the number of units sampled for the same level of accuracy of survey results) in two ways:

- the SBR includes links to units in these sources that can thus be consulted and that contain additional stratification characteristics;
- SBR data are of improved quality thanks to cross-checking with these sources.

Use of additional sources can also reduce the response burden on enterprises through increased sampling efficiency or by removing the need to conduct a survey altogether. The sources may even include additional units not present in the basic administrative sources, thus extending the coverage of the population of interest. However, the priority rules for determining the values of those characteristics which are available from more than one source need to be determined.

To ensure that a source is sufficiently well aligned with the SBR to be easily usable, it may be necessary to introduce additional criteria for its use, e.g., common unit identifiers, common definitions and classifications. The greater the coherence, the more useful the source is likely to be.

The rights and mechanics of access to administrative sources are also factors in facilitating their use. Many potentially useful sources are likely to exist, but they are of no actual use if the legal and technical means to link them to the SBR are not in place. Thus, where the data are deemed to be of considerable potential use, the NSI should invest resources to enable the appropriate links to the SBR to be made.

6.4.5 *Thresholds*

It is important for the SBR to have as complete coverage as possible, so care needs to be taken to understand the impact of thresholds. For example an enterprise may not be required to register for VAT if its turnover is below a certain threshold. If the threshold is low then the enterprises that are omitted because they are below the threshold have only a small impact on the overall estimates and their omission from the SBR is not an issue. However, if the threshold is quite high, as it is in some countries, then its impact on coverage and estimates is likely to be large. In this case additional sources and methods may be needed to supplement coverage provided by the source. This typically requires matching and linking the additional sources to the SBR to avoid duplication and minimise omission.

6.4.6 *Handling Changes in Administrative Sources*

Changes in administrative sources, for example, changes in thresholds, definitions or computer systems, can lead to discontinuity in the data supply. Critical to managing this type of change is working very closely with the administrative source to gain a full understanding of the change in order to assess the impact on the SBR. If possible, it is good to measure the impact of the change, for example by processing the administrative data following the change in a copy of the SBR database so that the impact can be easily identified before regular processing takes place. Often there are unexpected consequences as a result of changes to administrative data, particularly when complex systems are involved.

Once an assessment of the impact has been made, the next step is to work actively with the survey staff who are the most important SBR users to assess the impact on their outputs.

Changes in tax thresholds are common and often just reflect the general inflationary environment in the country. This type of change may not have any real impact on coverage. Occasionally more radical changes in thresholds take place, reflecting a political decision, and the change may well be permanent. These types of changes have an impact on SBR coverage, particularly if a single source is used. It may be possible to estimate the impact on the population, for example by comparing unit birth trends before and after the change.

There are two possible methods of addressing changes so that they do not distort the statistics provided to users:

- make an appropriate transformation in the SBR to minimise the discontinuity; or
- let the survey staff deal with the discontinuity through their own systems.

The method used must be clearly linked to the nature of the change. For example, a change in the classification systems in administrative data can best be handled by transforming data in the SBR, whereas an expansion of the SBR frame population caused by improved SBR coverage procedures may have to be handled by the survey staff.

6.5 *Identifying Statistical Units*

6.5.1 *Identifying Enterprises*

In most countries with administrative registers, a legal unit of some sort is defined for SBR purposes. In many situations the legal unit is exactly, or a close approximation to, an enterprise. However, it is necessary, at least for the bigger units, to be able to create an enterprise in accordance with the enterprise definition, which does not require it to be in one to one correspondence with a legal unit.

Whatever the particular definition or profiling approach, the transformation from legal units to statistical units must follow well defined basic rules. The rules must make it clear how the administrative units from each source contribute to the updating of the statistical units, which source has data that may overrule or overwrite data from another source, and when the changes are deemed to have happened.

6.5.2 Identifying Local Units and/or Establishments

Although in principle there is nothing to stop an SBR containing both local units and establishments, typically, for resource reasons, it contains one or the other. Local units/establishments enable regional economic analysis to be conducted.

Many of the usual administrative sources, such as VAT, refer to a legal unit as a whole or a specific part of it, but not necessary at local unit/establishment level, so they cannot be used to populate local units/establishments in the SBR. In a few cases, local units are directly registered in administrative sources. Establishments never are. Therefore creating local units/establishment is a bigger problem than creating enterprises.

Some administrative sources are more useful than others. For example, employment tax systems, or social security databases, may hold information on the location of the employees – this information could be used to construct a local unit. Other sources may list business sites for business tax purposes. Whatever sources are used they have to be linked/matched to other data sources and/or the SBR.

As noted above, some administrative sources identify local units. If this is the situation, then it is important to establish whether these units are defined in the same way for all units covered by the administrative source and hence whether the data can be used as the basis for creating local units in the SBR.

In most countries there may be no suitable administrative data source and local units/establishments have to be established through a special SBR survey. As many small enterprises never have more than one site, it is best to stratify such a survey to ensure that all larger businesses are included whereas smaller businesses are sampled very lightly simply to estimate the probability of smaller businesses owning more than one site. The resulting data can be used to prepare a typical model for small businesses.

Even so, a survey is an expensive option and very burdensome for the enterprises, since a range of characteristics have to be collected for each local unit/establishment. To reduce burden and expense it may be possible to obtain the same information through an annual employment survey.

6.5.3 Converting from Legal Units to Statistical Units

The following paragraphs consider the processes that are necessary to transform data from administrative sources into creation and updating of statistical units and their characteristics in the SBR. The main processes are listed and explained briefly. They need to be adapted to fit the specific requirements of each country and source.

A principle adopted by several NSIs is that the SBR should serve as the gateway to administrative micro-data. This means that all administrative data at the level of individual enterprises (or local units) should be fed through, but not necessarily stored in, the SBR. This enables the units and characteristics in each administrative data source to be matched to those in statistical surveys based on the SBR.

Another approach is to use the administrative units as observation units and through the links in SBR to transform the administrative data into statistical data at the level of each individual enterprise or establishment.

The first step in processing generally involves checking the quality and coverage of the incoming administrative data to ensure some basic conditions are satisfied, for example that:

- the file is the expected size, i.e. it contains roughly the expected number of records and the required characteristics;
- the values of the characteristics are in valid formats and/or ranges, for example, dates are within a permissible range, text fields contain only text characters, numeric fields contain only numbers, and codes used are valid;
- there is good coverage of the main characteristics, for example, identity numbers, addresses, economic activity codes are present for all units.

If the preliminary analyses show that the incoming data are of sufficient quality, the next step is to transform them into updates to statistical units.

6.5.4 Transformation Rules

An administrative source is unlikely to use exactly the same definitions of units and characteristics as the SBR, especially in the case of complex businesses. The *transformation* of administrative unit data into statistical unit data may therefore involve several steps, and the creation of algorithms and/or look-up tables to convert characteristics and classifications. An example would be the conversion of administrative economic activity codes to the standard codes used in the SBR (ideally ISIC Rev 4, or equivalent). The resulting *transformation rules* are dependent on the circumstances in each country.

As already mentioned, it is important to determine whether it is necessary to put in place an intermediate step in the transformation from administrative to statistical data, by creating legal unit to which both administrative units and statistical units are linked. Even if this is done, it might be difficult to automate the transformation from legal unit to enterprise as, in certain cases, there may not be a one to one relationship.

6.5.5 On Line Delivery

In an ideal situation the SBR is updated with information on line, unit by unit. In this situation there is no validation of a file as such. However, individual units and characteristics still have to be validated and the same rules for creating and updating statistical units are still appropriate. Methods for identifying anomalies are dependent upon specific country processes and requirements and on SBR resources.

6.6 Statistical Sources

6.6.1 Economic Census

As discussed in Chapter 2 and Annex G1, an economic census can be conducted entirely independently of the SBR, or it can be SBR assisted. It can involve the use of some administrative data, or it can involve only administrative data. The description in this section refers to traditional economic census where trained field enumerators seek out each physically recognizable place of business and collect the necessary information by direct interview and observation.

Economic censuses are used in many developing countries and some developed countries, including the USA. An economic census is undoubtedly a very useful instrument when a country is initiating an economic statistics programme. It provides benchmark data. Also, in the past, especially in developing countries, it was a well-established method for the initial construction of an SBR. However, in this latter respect it has a number of drawbacks.

- In particular, it is a very resource intensive exercise and requires large inputs of manpower and time. Censuses, therefore, tend to be carried out infrequently, for example, once every five years.
- Intercensal updating of the SBR is thus required, which is itself costly. Once the census population of enterprises and establishments has been identified, additional data sources have to be utilised to maintain it. As some units have been found by enumeration and are not registered, administrative sources are not sufficient. There has to be continued enumeration in selected areas. This is particularly resource intensive in developing countries where high levels of births and deaths of companies may occur.

In addition to its high cost, the enumeration approach has the disadvantage of not being able to identify and document non-recognizable places of business, or enterprises without a fixed location, for example, web-based businesses, or individual entrepreneurs such as electricians and plumbers, providing services at locations other than their homes.

In summary, economic censuses are not recommended as a means of establishing an SBR. Rather the converse is true, an economic census should draw its basic frame from the SBR, possibly supplementing this by an area sample. However, if and only there is no reliable administrative source whatever, a periodic economic census is appropriate.

6.6.2 Feedback from Enterprise and Establishment Surveys

Feedback from enterprise surveys is a vital mechanism for updating the SBR as it provides information on changes in contact address, changes in the economic stratification characteristics, deaths, etc. Feedback from enterprise surveys has the advantage that it is available at statistical unit level, i.e. establishment or enterprise. Thus, there should be close contact between survey staff and SBR staff to ensure the SBR is updated with survey results. Survey staff should play an active role in SBR data quality management.

Survey feedback has its limitations. An SBR updated exclusively by feedback from a traditional sample survey would have serious deficiencies. First, it would lack new units as surveys are not designed to find births. Second, the population of enterprises would not be fully maintained as feedback would be coming only from the sampled units.

Furthermore, even for the sampled units, use of survey feedback from sample surveys introduces the possibility of feedback bias as SBR updates are provided only for the selected enterprises. There are no technical problems with using data for enterprises that have been sampled with certainty – typically the large ones. However, for medium-size and small enterprises that are sampled with probability less than one in repeating surveys, there is a potential for causing bias in future survey samples. In this situation, updates have to be very carefully applied.

For example, suppose that when a particular quarterly survey is first conducted, the sample is found to contain 30% dead enterprises (this is not an improbable figure). Furthermore, suppose that, based on this sample information, the dead enterprises are removed from the SBR, and that the survey sample for the next quarter comprises the 70% live units from the previous sample plus a replacement of the 30% drawn afresh from the SBR. This new sample will contain about 9% (30% of 30%) dead units. Thus, it will no longer be representative of

the population of dead enterprises on the SBR, which is still nearly 30%, assuming that the survey sample is a relatively small proportion of the population. There are proportionally too many live enterprises in the sample. If the survey weighting procedures do not take this into account (by making allowance for the dead enterprises that were originally found in the sample), the result will be an upward bias in the estimates. Furthermore, the bias will increasingly worsen with each survey repetition.⁵⁷

6.6.3 SBR Improvement Surveys

SBR updating information that cannot be obtained from surveys, or from the administrative sources on which the SBR is based, can be obtained by specific *SBR improvement surveys* (also termed *nature-of-business surveys* or *proving surveys* or *SBR control surveys*) conducted by SBR staff.

Given limitations in resources it is unlikely that all units in the SBR can be maintained equally. It is usually necessary to focus improvement surveys on specific strata to measure and improve coverage and quality.

One possible strategy for keeping the SBR up to date is to conduct SBR sample surveys every year in which the biggest, the medium sized and the smallest units are sampled 100%, 50%, and 10% respectively. This practice helps keep values of the characteristics of the units up to date in an efficient way. Such surveys may also be specially designed to measure SBR accuracy, for example, to measure errors in classification by economic activity, or by size, or to estimate the proportion of falsely active units, as further discussed in Chapter 9.

6.6.4 Profiling

Profiling is the practice of using company accounts, often accompanied by interviews with senior enterprise officials, to build and define the structure of enterprises, mainly those involved in large complex enterprise groups. The resulting *profiles* are used to produce a reporting structure appropriate for the surveys conducted by the NSI. Profiling usually involves establishing contact with the enterprise being profiled to develop a good understanding of its structure. It is possible, however, to complete smaller profiles simply using published accounts.

While profiling is not a primary source of data, it does provide valuable information on the larger and more complex enterprises that individually make a significant contribution to the country's GDP. It is especially important in identifying enterprise groups, as discussed in Chapter 4.3.

Profiling is often organized so that each individual SBR profiler has an assignment of large, complex enterprise groups for which he/she is responsible for reviewing and updating. When profilers react to signals from various sources on enterprise that may need updating this is referred to as *reactive profiling*. Profilers also periodically perform *proactive profiling* to augment reactive profiling. For example, they may routinely arrange to re-profile every large enterprise once every three years. Since large enterprises can grow and change dynamically, considerable resources may be needed to maintain the accuracy and relevance of the SBR in this respect.

⁵⁷ Example drawn from Section 9.4, Guidelines for Building Statistical Business Registers in Africa

The first step in profiling is to determine the criteria by which to identify the enterprises to be profiled. Profiling is usually focussed on large complex enterprise groups with multiple activities for which survey reporting is difficult. After determining the profiling criteria, a programme of regular updates over, say, a three to four year period, should also be established. There is a balance to be struck between the resources available and the amount of profiling that can be conducted, and this needs to be considered when determining the criteria. Resources should also be planned to deal with emerging issues that occur outside the routine reviews, i.e., reactive profiling. For example, a major merger might require a profile to be conducted ahead of the scheduled regular review.

The next step is to gather preparatory material. This should include records on all legal units within the enterprise group, the reporting history for surveys, and SBR data such as employment, turnover and classification, etc., for the enterprise group as currently defined. These data should be examined for consistency and to help identify reporting issues. As this process is labour intensive, it is best to create a standard template for the data required and to automate the extraction of this information from the SBR.

Further background information should be gathered by searching enterprise websites and examining annual accounts and reports. In simpler cases it may be possible to conduct a profile simply on the basis of this material, but for complex cases it is invariably necessary to meet with representatives from the enterprise.

Following this preparation, contact is made with the controlling enterprise in the enterprise group (the *global group head* or *global decision centre*) and a visit arranged if necessary. Through discussion it will be possible to identify the main trading (i.e., active) enterprises within the enterprise group and to agree a mutually acceptable reporting structure covering these enterprises. The aim is to settle on a structure with enough detail for the statistical surveys while minimising the respondent burden for the enterprise group. Specially trained staff are required to negotiate with the enterprise as they will typically be talking to top management in the enterprise – the chief accountant and/or the company secretary. Ideally, these staff will have a good understanding of the business processes and accounting practices, as well as the collection of economic data.

6.7 Combining Administrative and Statistical Sources

To build a comprehensive SBR, a combination of administrative and statistical sources is recommended. Administrative sources identify enterprises, but may not include all of the required characteristics. Statistical sources do not identify new units but provide additional or more accurate characteristics. A strategy of using administrative and statistical sources in combination should be developed and employed. For example, administrative sources can be used to identify legal units and transform their data to form enterprises, while local units or establishments can be identified by a survey of the enterprises, as can characteristics missing from administrative sources.

Using administrative and statistical data in combination is illustrated by the following examples.

- Suppose that an enterprise is thought to have only one establishment but employee data from an administrative source indicate that half of workforce live in an area far away from the identified establishment. This may be a signal that there is a second establishment, or that the employee data have been linked to the wrong enterprise.
- Suppose turnover from a VAT source does not correspond to the turnover for the same enterprise from an import/export administrative source. This may be a signal that some

part of the enterprise is missing, or the links to one or other of the administrative sources are wrong.

- Suppose there is feedback from a survey indicating that the responding (legal) unit no longer has any economic activity. This may be a signal that the activity has taken over by another legal unit.

Statistical sources can also be used to estimate missing characteristics. For example if an administrative source only contains employee numbers and industry code, it may be possible estimate turnover using turnover/employee value for similar units.

6.8 Record Linkage in Creation and Maintenance of the SBR

6.8.1 *Introductory Remarks*

Usually, the main sources for the maintenance of the SBR are administrative data, such as data from taxation, social security or other administrative sources depending on their availability and adequacy. If these administrative sources have separate identification systems and if no identifier known to the SBR is provided, then record linkage methods have to be applied in order to be able to use the administrative data in the SBR.

This section provides a guide to the theory and practice of record linkage. It is also important to note that there are commercial software products for record linkage. While these can be expensive, their use may be cost effective compared to the investment required to build sufficient functionality in the NSI.

Before any record linkage method is applied it is necessary to analyse the administrative source to determine the quality of the source and whether it is useful for SBR purposes.

6.8.2 *Basic Approach*

Record linkage is the linking of a data record in one data source with one or more data records in another data source. In the case of the maintenance of the SBR it is the linking of one data record in the SBR with one or more records in the administrative data/registers or vica versa. The challenge is to link records belonging to the same unit (whether this be a legal unit, enterprise, local unit, etc.) in different sources. If two records for the same unit are brought together this is called a *match*⁵⁸.

Where a common identifier is not available, identification of a match has to be based on a *similarity measure* that is computed using characteristics that are available both in the SBR and in the administrative source. The characteristics used are those that help identify a unit uniquely typically name and address, and possibly other characteristics like legal form. The choice of characteristics may be different for different sources. Only a limited number of characteristics should be selected, namely those that most help in identifying units and that are least likely to be in error.

58 “An alternative is known as a *statistical match* and is defined as the linkage of records for similar units rather than for the same units. Statistical matching is ordinarily employed when the files being matched are probability samples with few or no units in common; thus, linkage for the same unit is not possible for most units.” See Thomas N. Herzog, Fritz J. Scheuren, William E. Winkler: Data Quality and Record Linkage Techniques, Springer 2007, page 81.

It has to be taken into account that there may be errors in the data for units in either or both the administrative source and the SBR. For example, names may be misspelled, addresses may be out of date. Prior to matching, the quality of the characteristics of potential use for matching should be analysed in both sources.

Another problem stems from the fact that the data in the administrative source may not be recorded in a format that is standardized and comparable with that used in the SBR. The length of the fields in the data records may be different, abbreviations may be used or one characteristic may be combined with another one, e.g. the legal form of a unit may be part of the name or in a separate field. Therefore, parsing the format in the administrative source has to be done first and is a very crucial step before matching methods can be applied. If not, the quality of the matching results will be lower than it could be.

It should also be mentioned that data linkage is quite computationally intensive: each data record in one data source needs to be compared with all the data records in the other source. Even in case of only a few hundreds of data records in each source this results in tens of thousands of comparisons. Therefore a matching run may need hours of processing time. This can be drastically reduced if *blocking methods* are applied (as further described below).

6.8.3 Types of Matching

There are two basic types of matching: *deterministic* and *probabilistic*. The methods are very similar and are based on the computation of numerical values expressing the similarity of a record pair.

- In case of deterministic matching, a *similarity measure* (or *distance function*) is defined. A unit of a SBR and a record of an administrative source is said to be a match if the two records have a high similarity for each element within the collection of identifying characteristics called the *match key*.
- In probabilistic record linkage (formalized by Fellegi and Sunter⁵⁹) a so-called *likelihood ratio* is computed. The value obtained is compared with two thresholds: a *lower threshold*, below which all values are defined as *non-match*, and an *upper threshold*, above which all values are defined as *match*. The *grey area* between lower and upper thresholds defines *potential matches*.

Although probabilistic record linkage is mathematically very elegant its use in practice is limited since no implementation with a broad field of application exists in standard statistical software packages. Moreover studies⁶⁰ have shown that deterministic record linkage, which is easier to implement, can result in equal or even better quality than probabilistic linkage.

6.8.4 Standardizing

Matching should be an automated process as far as possible because of the large number of comparisons to be made. To achieve satisfying results with an automated matching algorithm the values of the characteristics being compared should not be formatted differently in the two sources. Before text characteristics are compared they must be standardised and parsed. Without standardization, many true matches would be erroneously designated as non-matches

59 Fellegi, I.P. & A.B. Sunter, „A theory for record linkage“ in Journal of the American Statistical Association, 64, pp. 1183-1210, 1969

60 Weghofer E.: Beurteilung ausgewählter Stringvergleichsalgorithmen zur Eignung für Record Linkage an Hand einer empirischen Datenbank. Diplomarbeit an der WU Wien. 2004

because the common identifying attributes would not have sufficient similarity due to differences in formatting. Standardization involves consistency of spelling, consistency of coding and elimination of obscure values.

Standardization of text is essentially a statistical process. It is usually done by computing the frequency of all words in sources. If the frequency of a string (like 'corp', 'inc', 'ltd') is very different in both registers, the string can be deleted from both sources, or abbreviated identically in both sources, or replaced by a synonym in at least in one source. An example of parsing is to ensure that all parts of a name are in the same order (e.g. family name, first name, title).

Before starting automated matching, the contents of each text field must be analysed and the following questions must be answered:

- Are different languages used in the different sources? Are there combinations of letters that have a special phonetic value?
- Does the name field contain personal names, company names or both?
- Does the address field contain street names, city names or both?
- Are there any other peculiarities in a text field, e.g., the string 'comp' precede every company name?

It is pertinent to consider how the data in the sources were obtained. Phonetic variations are likely to be found in data collected by phone, whereas typographical mistakes are probable in manually completed questionnaires. For standardizing purposes the following operations may be useful:

- deleting blanks at the beginning or the end of a text field;
- converting lowercase characters to uppercase;
- deleting village name from the field for street name;
- converting special characters;
- converting Latin numerals to Arabic numerals;
- moving the most crucial word to the beginning of the text field;
- if corresponding text fields (e.g. company name) have different lengths, truncating the longer to the shorter length.

6.8.5 Matching of Text Strings

Given the aim is to match two records (that do in fact refer to the same unit) by comparing text fields like name and address, if one or both of the two text fields has a typographical mistake, then the matching process will fail unless account is taken of the possibility of typographical errors. Thus, records and their textual characteristics (such as name and address) are considered as strings of alphanumeric characters and *string comparator metrics* are used to compare the two strings and to determine how much alike they are to each other. Usually the values of the metrics lie in the interval from zero to one, with one indicating perfect agreement and zero indicating high dissimilarity.

There are many different possible string comparator metrics. Jaro⁶¹ introduced a metric that accounts for the lengths of the two strings and for the types of errors that human beings typically make when writing or typing alphanumeric strings, like insertions, omissions and transpositions. (By *transposition* is meant that a character from one string is in a different position on the other string. For example, in comparing “company” to “copmany”, the “m” and “p” are transposed.) Winkler⁶² augmented Jaro’s metric by giving more influence to the characters in the beginning of a string. An extension of the Jaro metric is the Edit or Levenstein metric. It is computed as the minimal number of edit operations (substitutions, imputations and deletions) which are necessary to convert one string to another string divided by the maximum of the two string lengths.

Another important and widely used family of measures is based on N-grams. In this case the characters themselves are not compared but the sets of all N-grams of the strings. Usually either sets of all substrings of two (bigrams) or three (trigrams) successive characters are compared. The similarity measure is the total number of bigrams or trigrams that are in common divided by the average number of bigrams or trigrams in the two strings.

- For example: The bigrams of the string “Welcome” are “_W”, “We”, ”el”, ”lc”, ”co”, ”om”, ”me” and ”e_”, for “Welkome” the 4th and 5th bigrams are different, being “lk” and “ko”. Thus the measure is $6/8=0.75$. As an option, the first and last bigrams containing the blank can also be left out, giving a measure of $4/6=0.67$.

6.8.6 Blocking

If in matching the SBR with an administrative source the numbers of record pairs is so high that it is not feasible to compute string comparator metrics for all possible pairs, then consideration should be given to blocking. *Blocking* is a method that reduces the number of pairs of records that are examined. In blocking, the files from the two sources are partitioned into mutually exclusive and exhaustive blocks. Comparisons are restricted to record pairs within the same blocks. Blocking is generally implemented by partitioning the two files based on the values of one or more characteristics. For example, if both files have a postal code field, the pairs to be compared might be restricted to those records whose postal codes agree. This would be an example of regional blocking. The advantage of blocking is that the number of comparisons may be greatly reduced. The disadvantage is that record pairs disagreeing on postal code are automatically classified as non-matches. Thus, if a record in one of the files an erroneous postal code then it will never be matched even if there is a record for the same unit in the other file.

This situation can be mitigated by iterating the matching process with a different blocking scheme, or even a sequence of blocking schemes. Suppose the first iteration used postal code as the sole blocking item and the second iteration used company name then failure to match due to an erroneous postal code would be somewhat offset by the possibility of matching during the second iteration. Further iterations could be made until the analyst felt that it was unlikely that matches would be missed because of errors in the blocking fields.

61 Jaro, M.A., „UNIMATCH –a computer system for generalized record linkage under conditions of uncertainty“ Spring Joint Computer Conference, 1972, AFIPSL-Conference Proceedings, 40, pp.523-530, 1972
62 Winkler, W:E:, „String comparator metrics and enhanced decision rules in the Fellegi-Sunter model of record linkage“ Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 354-359, 1990

6.8.7 Computing Overall Similarity Measure

The calculation of a string comparator metrics for all compared pairs of records is usually done by a computer program. The metrics are computed separately and are combined afterwards, typically using a weighted sum, where the weights are defined according to the quality of the fields. The analyst has to define a lower and upper threshold for this overall similarity measure. Pairs of records with a value above the upper threshold are considered as matches, whereas pairs of records with a similarity measure below the lower threshold are considered as non-matches. When the metric lies between the lower and upper limit that pair has to be clerically reviewed to determine if it is a match or non-match.

- The higher the upper threshold the small the number of false matches.
- The lower the lower threshold the smaller the number of missed matches.
- On the other hand the further apart the thresholds the more manual work is required.

As manual work is time consuming and therefore expensive the choice of thresholds depends on balancing the impacts of false matches and of missed matches with the SBR resources available.

6.9 Other Data Sources

In addition to the data sources discussed in previous sections, there are commercial sources. Furthermore new sources are emerging associated with increased computer power and on-line form filling, or filing, of data to administrative systems, as outlined in the following paragraphs.

Telephone directories

Telephone directories or special listings prepared by telephone companies can be useful in adding or confirming SBR data. They should not be used as sources of new enterprises.

Payroll, taxation and accounting service providers

Payroll, taxation and accounting service providers provide enterprises with services that involve paying an enterprise's staff and/or making returns to the taxation authorities on its behalf and/or managing its accounts. It may be possible for the NSI to build agreements with such service providers and their client enterprises that allow the service providers to provide data for the enterprises directly to the NSI. This saves the enterprises from the burden of survey questionnaires and provides faster and more efficient data flow to the NSI. Identification of the relevant enterprises is required. However, it may still be necessary to contact the enterprises directly to obtain information that the service providers cannot provide, for example economic activity classification.

Internet search

The Internet may be useful as a supplementary source of data, but cannot currently be reliably used for identifying new enterprises. However, it can provide information on the economic activity, on the production profile, on up-to-date addresses, etc., so is becoming an important source

Commercial data providers

There are a number of commercial enterprises that provide global, regional and domestic company information, e.g. Dun and Bradstreet. These organisations serve businesses, allowing them to make informed decisions, e.g. credit decisions, marketing etc. Their data are

also valuable for maintaining an SBR, in particular by providing information on enterprise group structures.

Data from commercial data providers are, by definition, based on publicly accessible information and could also be obtained directly by the NSI. However, the equivalent collection, matching and processing within the NSI would require significant resources and may well prove to be more expensive than purchasing these tailor made data.

Big Data

Big data are defined as large, possibly unstructured datasets that are potentially available in real time. The opportunities for using big data are still being developed. So far no practical experience in their use in the SBR has been identified. In the future, however, big data may reveal opportunities for developing additional sources for the SBR. Accessing big datasets, which are often owned by private businesses such as mobile telephone operators or internet providers, might be a challenge. Structuring and editing the data might also require substantial investment but be worthwhile if accompanied by significant benefits in terms of coverage and timeliness.

7 Maintenance of the SBR

7.1 Introduction

Whereas Chapter 6 describes the data sources used in SBR *construction*, this chapter discusses *maintenance* of the various types of statistical units and their characteristics. The key objective of maintenance is to update the coverage and content of the SBR, taking into account continuity and stability rules, according to a well defined calendar, and in as timely a fashion as the information sources allow. The basic aim is to provide economic surveys with sampling frames that are accurate and as up to date as possible.

Section 7.2 presents key aspects of the SBR maintenance strategy. It discusses the data sources, in particular administrative sources, and how to deal with conflicting information. It describes the practice of allocating enterprises to *maintenance groups* for maintenance purposes, with different procedures for each group according to the size and complexity of the enterprises it contains. It discusses the timing of maintenance in relation to the production of survey frames.

The various types of changes that can occur to statistical units in the live register, and how they are handled, are described in Sections 7.3 and 7.4. The former section deals with *births*, *deaths* and *changes in linkages* of statistical units. The discussion is presented in terms of a typology of *demographic events*. The latter section discusses all other changes, i.e., changes in the *characteristics* of statistical units.

Continuity rules for determining whether a statistical unit is deemed to have *continued* through changes (for example, of ownership, size, economic activity and/or location), or to have died and been replaced by another unit, are discussed in Section 7.5, as are *stability rules* (sometimes called *resistance rules*) that restrict the speed with which changes of characteristics are applied in order to inhibit unwanted oscillations in values.

The final section introduces the distinction between *changes* that actually occur in the real economic world and *corrections of errors* caused by inadequacies in SBR procedures, and it discusses the sources and treatment of errors.

The primary references used in developing this chapter are the Eurostat (2010) *Business Registers Recommendations Manual* (Chapters 12-18), the AfDB *Guidelines for Building Statistical Business Registers in Africa* (Chapter 10), and the Eurostat/OECD (2007) *Manual on Business Demography Statistics* (Chapter 4).

7.2 SBR Maintenance Strategy

7.2.1 Introductory Remarks

“A country’s economy is constantly changing – new businesses are formed, existing businesses merge, change production activities or location, go bankrupt, etc. To ensure that enterprises (and other standard statistical units) in the SBR remain aligned with and representative of legal units and their productive activities, these changes have to be detected

and the SBR has to be correspondingly updated. This process is referred to as *SBR maintenance*⁶³.

Maintenance implies constant updating of the SBR in terms of coverage and content. New statistical units have to be identified and recorded without over-coverage, under-coverage, or duplication. Changes in the values of the characteristics of existing statistical units have to be identified and recorded. The dates of changes should be known and recorded. Ideally, the information should be collected just once.

The maintenance strategy should take into account the following aspects:

- the *sources of information* - administrative sources, feedback from surveys and SBR improvement surveys - and rules for dealing with conflicting information;
- the *maintenance groups* into which statistical units in the SBR are divided for efficiency and cost effectiveness of maintenance procedures;
- the *time dimension* - timing of the updates, the need for continuity and stability rules, and use of the historical register.

7.2.2 Use of Administrative Sources

In terms of number of updates to statistical units, administrative sources are the major contributors to SBR maintenance. It is recommended that one administrative source is identified as the primary source for construction and maintenance purposes. The source chosen should be the one that provides the best possible balance of coverage, content, timeliness and accuracy. In using additional (secondary) administrative source(s), care must be taken that no duplication of units arises. Duplication can most easily be avoided if the secondary sources use the same unit identification scheme as the primary source, or have no overlap with the primary source. An example of the latter case is where the primary source is a tax register and the secondary source is a list of government departments. In the absence of common identification scheme, or coverage that is known to be complementary, records in a secondary source must be matched to those in the primary source and links established between records referring to the same unit, thus avoiding creating duplicates of that unit

7.2.3 Use of Feedback from Surveys

Feedback of changes in frame data from previously conducted economic surveys is a valuable source of updates to statistical units. By design, the first few questions asked by any economic survey should constitute a check of the values of basic characteristics (such as name, address, contact information, and activity status) of the unit recorded in the survey frame. Inactive or dead units should be identified. Subsequent questions may collect updated versions of economic activity codes and size measures.

There are no technical problems in updating the SBR with data for any statistical unit that has been sampled with certainty. However, for medium-size and small enterprises that have been sampled with probability less than one, there is potential for causing bias in future survey samples, as described in Section 6.6.2. Thus, it is recommended that feedback from surveys about characteristics (such as economic activity and size) that are used for sampling statistical units should not be used to update units in the SBR that were not selected with certainty.

⁶³ African Development Bank, 2012, Guidelines for Building Statistical Business Registers in Africa, Chapter 5.7

Instead, the information can be used trigger further maintenance operations and/or can be taken into account in future repetitions of the survey, for example by not trying to contact units that are known to be dead.

7.2.4 Use of SBR Improvement Surveys

The objectives of *SBR improvement surveys* (also called *nature of business surveys*, *SBR proving surveys* and *SBR control surveys*) are to verify the current values of key characteristics and to obtain missing values. Such surveys are conducted by the SBR on a continually repeated basis, with sample sizes matching the SBR resources available. The aim in selecting the samples is to have maximum beneficial impact in terms of SBR quality improvement. There is no focus on estimation. The total sample size over a year is determined by taking into account the frequency with which changes occur in the economic world, the quality of data received from the administrative sources, the quality of data currently in the SBR, and the resources available.

7.2.5 Profiling

As discussed in Section 6.6.4, the aim of profiling is to produce an appropriate statistical and reporting structure for large and complex enterprises. Profiling can be *reactive*, i.e., in response to a signal from any of the various sources that there has been a change. It can also be *proactive*, typically involving a re-profiling program in which every large enterprise is re-profiled every so many years.

7.2.6 Dealing with Conflicting Information

As evidenced in the previous paragraphs, the SBR is updated from a range of sources, several of which may provide values for the same unit and characteristic. When this happens there may be a conflict in the values. This raises the question of which source and value to use. The answer requires a thorough understanding of each source - the methods of data collection and validation, the time of collection, and the relative importance assigned to that characteristic by the source. This understanding then allows the sources to be prioritized, either with respect to all units, or differently for various categories of units.

Once priorities have been determined for each characteristic, the next step is to apply them to SBR updating procedures. This can be done in several ways. The most reliable method is to record the source and date of recording of the value of a characteristic and to ensure that the updating algorithms specify what combinations of source and date can result in the value being overwritten subsequently. Date is important because it allows for the possibility that a new value from a lower priority source can overwrite the existing value from a higher priority source if the information is significantly more recent.

7.2.7 Maintenance Groups

The amount of maintenance effort devoted to a statistical unit should be in accordance with its size and potential impact upon published statistics, and should take into account its propensity to change, and the sources of updating information. Thus, statistical units should be divided by size and potential impact and propensity to change into maintenance groups, each of which is subject to a particular set of updating procedures. Three examples follow.

Australian Bureau of Statistics (ABS) Business Register: maintenance groups

- *ABS maintained enterprises.* These enterprises are maintained by the SBR staff by profiling, SBR survey, business survey feedback and ad hoc investigations, as appropriate according to their size.
- *Australian Tax Office (ATO) maintained enterprises.* These enterprises are currently maintained entirely by taxation data from the ATO. However, the capability to include ABS sourced information has recently been introduced into the SBR with a view to exercising this option to improve data quality in the future.

Swiss Business Register: maintenance groups

- *Enterprises subject to profiling.* All enterprises with more than 10 local units, or more than 100 employees, are included in the *profiling group*. The profiling staff contact them directly every quarter to gather structural information and employment data.
- *Enterprises with only one local unit, including small businesses.* These are maintained using updating information from administrative sources and economic survey feedback.
- *Enterprises subject to light profiling.* All other enterprises are included in the *light profiling group*. The profiling staff contact them annually by internet survey to gather structural information and employment data.

This maintenance strategy is dictated by the need to determine employment for all local units as regional distribution of employment is an essential component of Swiss business statistics. All enterprises with only one local unit are included in the SBR improvement survey, which also enables quality control of activity code for the whole SBR.

Statistics South Africa Business Register: maintenance groups

- *Enterprises subject to profiling.* This set of enterprises is maintained by profiling operations. Indications of changes from administrative sources or from business surveys are used as signals indicating the need for re-profiling. The enterprises are not included in SBR improvement surveys, nor are they subject to ad hoc investigations.
- *All other enterprises.* All other enterprises are maintained using updating information from administrative sources and economic survey feedback, which is supplemented as needed by information gathered by SBR improvement surveys and ad hoc investigations.

7.2.8 Updating Schedule

The SBR is updated from several sources of information, each having a given periodicity, coverage and content. This must be taken into account in building a coherent maintenance system. The schedule for maintenance of the SBR and production of frozen frames should be coordinated with the supplies of data from administrative and statistical sources, with the production of frames for the surveys using the SBR, and with the publication (if any) of data directly from the SBR. The schedule indicates to users the likely variations in SBR coverage and content over time.

The SBR should be updated on a daily basis. Potential updates cannot be allowed to backlog. It is for this reason often referred to as the *live register*. The timing of the supplies of updating data should be discussed and coordinated with the sources. The production of frozen frames can then be undertaken in a systematic and controlled way, taking into account the whole economic statistics program, and, in particular, the conflicting demands from structural surveys (that want the most up to date information available for the reference period for which the survey is being conducted) and sub-annual surveys (which want stability across reference periods so that the effects of maintenance do not appear as artefacts in survey estimates of change).

A frozen frame is never directly updated. A preliminary version of a new frozen frame is compared with the previous version. Unexpected changes are analysed. If they turn out to be due to errors in the SBR, then these errors are corrected and the production of the frozen frame is repeated.

Staff of surveys using the SBR as the source of their survey frames should be aware of the SBR production calendar and they should be informed about any particularly significant updates in SBR coverage or content occurring since production of the previous frozen frame that may significantly affect their survey frames.

Statistics Canada

Statistics Canada produces SBR reports and also allows survey staff to see preliminary versions of their survey frames after key update processes have been completed each month.

Swiss Federal Statistical Office

A working group on economic classifications produces reports for the users on important updates in economic activity classification.

7.2.9 Historical Register

As described in Chapter 2, the historical register is an SBR output that enables the reconstruction of the history of the units. It supports:

- obtaining information about births , deaths and continuity of units;
- obtaining information about size development (how fast a unit or group of units is growing, through what kind of economic activity);
- analysing changes in the characteristics of units, for example, in location, judicial and financial links, and economic activity;
- reconstructing the state of the SBR as of a past date or reference period.

As a prerequisite, the dates of all demographic events and changes in values of characteristics should be recorded. If this is not possible, dates for core changes should be recorded.

In principle, a historical register can be derived and maintained in three different ways

1. The SBR is considered to be established afresh at the beginning of each year. During the year, all changes that each unit undergoes and the dates and reasons for those changes are recorded. At the end of the year, the SBR is copied and stored. Together with the change records this copy becomes a historical register.
2. The snapshots of the SBR taken at regular intervals (typically quarterly or annually) provide a series of pictures of the structure of the register. From each of these snapshots, the numbers of units and their characteristics at a specific points in time can be obtained.
3. The SBR records and maintains complete information about changes as they occur on a continuous basis. Each change is date stamped with the date it was recorded in the SBR and the date it occurred in the real world. This allows populations of units and their characteristics to be constructed for any point in time. It also provides information about delays in registering changes. Storing the *reasons* for the changes is also useful but is much more difficult. As a compromise, some changes can be automatically identified and categorised on the basis of the automated rules and procedures by which they came about.

Statistics Canada's Business Register Journal

In order to keep track of the reasons for the changes, the *Business Register Journal* is used to store the reason for the change when a profiler, SBR user or updater performs an update. The Journal entry is created automatically with a link to the unit(s) being updated when the update is performed. The entry contains auto-generated scripts of the event or update, and allows for the profiler or updater to complement this information with additional information. All of the information in the Journal is in addition to the update log in which the following information is kept: previous information/value, new information/value, date and source of update, and for some events, the *effective date*, i.e., the date the event actually occurred in the real world.

7.3 Handling Births, Deaths and Linkage Changes of Statistical Units⁶⁴

7.3.1 Demographic Events

A *demographic event* is defined as an *event that has an impact on the existence of a statistical unit, or on links between statistical units.*

⁶⁴ This Section is based on the Chapters 12, 13, 14, 15 and 16 of the Business Registers Recommendations Manual (Eurostat, 2010) and on Chapter 4 of the Manual on Business Demography Statistics (Eurostat – OECD, 2007).

A demographic event is based on changes in the existence of *production factors*, or in their distribution, within and among statistical units. It may involve the continuity (survival) of a unit over time or its discontinuity (death). It may also be accompanied by changes to the values of certain characteristics, such as size or type of economic activity.

Figure 7.1 presents a general typology of demographic events relating to enterprises and enterprise groups. It is based on two distinct types of events:

- events involving *existential changes*, i.e. the emergence or disappearance of combinations of production factors: and
- events involving *distributional changes*, i.e. changes in the distribution of production factors between enterprises.

The typology reflects the importance of the enterprise as the core statistical unit.

Figure 7.1: Typology of demographic events

1			Changes of existence of combinations of production factors
	1.1		Emergence of combinations of production factors (birth)
	1.2		Disappearance of combinations of production factors (death)
2			Changes in the distribution of production factors
	2.1		Redistribution of the production factors within one enterprise
		2.1.1	Redistribution of production factors across local units
	2.2		Redistribution of the production factors of more than one enterprise
		2.2.1	Merger, takeover (Concentration of enterprises)
		2.2.2	Split, break (De-concentration of enterprises)
		2.2.3	Transfer of production factors between enterprises
		2.2.4	Enterprise restructuring
	2.3		Redistribution of the production factors within one enterprise group
		2.3.1	Redistribution of production factors across local units of more than one enterprise
		2.3.2	Redistribution of production factors across enterprises
	2.4		Redistribution of the production factors of more than one enterprise group
		2.4.1	Concentration of enterprise groups
		2.4.2	De-concentration of enterprise groups
		2.4.3	Transfer of production factors between enterprise groups
		2.4.4	Enterprise group restructuring

7.3.2 Demographic Events, Continuity and Information Sources for Enterprises

Information sources

Most events are detected after reception of *signals*, typically changes in the administrative data that are used to maintain the SBR. Additional information may be needed, either directly collected or from other sources, to decide what updates to enterprises should be made based on these signals. The updating procedures may well differentiate between micro, small, medium and large enterprises.

Creations and cessations of administrative units do not necessarily result in births and deaths, respectively, of enterprises. An enterprise may be born, or may die, without a change in the legal unit(s) that own(s) and/or control(s) it. Detection of enterprise deaths is more difficult than detection of enterprise births as administrative sources are quick to add new units but slow to indicate those that have ceased to have economic activity.

Definition of continuity

An enterprise is considered to *continue* through a demographic event if its production factors continue. It is considered to *die (discontinue)* if its production factors discontinue. In practice, continuity is interpreted as meaning that of at least two of three following characteristics remain essentially the same: *controlling legal unit, economic activity and location*.

An enterprise is deemed to continue if it resumes its seasonal activities, or resumes its activities within 24 months of stopping them.

Birth

The *birth* of an enterprise is the *creation* of a combination of production factors with the restriction that *no other enterprise is involved in the event*. Handling a birth in the SBR means creating a new enterprise unit, i.e. an enterprise unit with a new identification number. A birth is typically detected by a signal in the form of the appearance of a new unit in an administrative source. However, as noted above, not every new administrative unit results in the birth of an enterprise.

Death

The death of an enterprise is the dissolution of a combination of production factors. Handling a death in the SBR means giving the enterprise an activity status of *dead*, also described as *deathing* or *ceasing* the unit. Although this does not actually involve erasing the unit from the database, this is also sometimes referred to as *deleting* the unit. Deaths are typically detected through survey activities and through the disappearance of units from an administrative source.

Change of ownership

A change of ownership is where a new legal unit takes over the production activities of an existing enterprise. This event by itself does not affect the continuity of the enterprise and, therefore, should not cause the birth or death of an enterprise. However, a change in ownership typically results in the death of an administrative unit and birth of another one. Thus, in practice, for small enterprises that are maintained entirely by data from an administrative source, a change of ownership may well result in the death of an enterprise and the birth of another one (resulting from the birth and death of the corresponding administrative units). For a large enterprise, a change of ownership may be detected through profiling and thus the continuity of the enterprise can be detected and the enterprise maintained.

Restructuring within an enterprise

Restructuring within an enterprise, for example, creation of a local unit, does not affect the continuity of the enterprise, though it may be associated with changes in characteristics such as size or economic activity.

Concentration (merger, takeover)

In a *concentration* there is *more than one legal unit before the event and only one legal unit after the event*.

- If all the legal units before the event lose their identity, the event is a *merger*.
- If one of them retains its identity, the event is a *takeover*.

In principle, mergers and takeovers do not necessarily involve the deaths or births of enterprises as such, though some characteristics (such as size and economic activity) may change. However, quite often, an SBR is designed in such a way that enterprises are tightly linked to legal units. Thus, in the absence of additional information from profiling:

- in the case of a merger, all the enterprises corresponding to the legal units that existed before the event are ceased, and a new enterprise is created.
- in the case of a takeover, the enterprise corresponding to the legal unit that takes over the other legal units continues and the enterprises corresponding to the legal units that were taken over are ceased.

De-concentration (split-off, break-up)

A *de-concentration* is an event *involving one legal unit before and more than one legal unit after the event*.

- In a *break-up*, the original legal unit dies, all the legal units after the event are new.
- In a *split-off*, the original legal unit continues, the other legal units are new.

In principle, break-ups and split-offs do not necessarily involve the deaths or births of enterprises as such, though some characteristics (such as size and economic activity) may change. However, as noted above, quite often the SBR is designed in such a way that enterprises are tightly linked to legal units. Thus, in the absence of profiling information, the births and deaths of enterprises follow those of the legal units.

Reconstruction of demographic events from SBR information

The cause of a new enterprise record can be a birth, a merger, a break-up or a split-off. The cause of an enterprise being given the activity status of dead can be a death, a merger, a takeover or a break-up. To reconstruct a demographic event more precisely, it is necessary to determine all the enterprises involved and to have recorded their links to one another over time and the dates of changes in those links.

7.3.3 Demographic Events, Continuity and Information Sources for Local Units

Definition of continuity

The *continuity* of a local unit is defined in terms of continuity of its location and production factors, with an emphasis on those production factors that can be readily identified at the level of the local unit, i.e., land, buildings, and employment.

- *If the location remains the same*, the criteria for local unit continuity are (1) the continuity of production factors, in particular of employment, and (2) the continuity of the enterprise to which the local unit belongs. A local unit that does not change location is deemed *not to continue*, i.e. *to lose its identity*, if at least two of the following three

factors change: the enterprise identity, the principal activity, or at least 50% of the employment.

- *If a local unit changes location within a region (defined at the local level), it is deemed to continue, i.e. to retain its identity, if none of the three factors mentioned above change. Otherwise, it is deemed to discontinue.*
- *If a local unit moves outside the region it is deemed to discontinue.*

A local unit is deemed to continue if it resumes its seasonal activities or resumes its activities within 24 months of stopping them.

The birth of an enterprise that comprises only one local unit implies the birth of a local unit.

The death of such an enterprise implies the death of a local unit. However, a local unit belonging to an enterprise with only one local unit can discontinue and be followed by the creation of different local unit whilst the enterprise continues.

Birth or death

The birth of a local unit is the creation of a (partial) combination of production factors at a geographically identified place. A death is their dissolution.

Births and deaths are handled by creating new local unit records, and marking existing ones as dead, respectively.

Transfer

A local unit may be deemed to continue and be transferred from one enterprise to another. This may be handled by deletion of the link between the local unit and the enterprise to which it belonged before the event and the creation of a link between the local unit and the enterprise to which it belongs after the event.

7.3.4 Demographic Events, Continuity, Information Sources - Enterprise Groups

Typology of demographic events for enterprise groups

In the context of enterprise groups the focus is on events reflecting concentrations (mergers, takeovers) and restructuring, and their impacts, notably on employment at both national and international level. Births and deaths are far less frequent for enterprise groups than for enterprises.

There are certain differences in the recording of events and in the allocation of identity numbers according to type of enterprise group (multinational group, truncated group, and all-resident group). This is in order to ensure coherent handling of the continuity of multinational enterprise groups and their truncated parts.

The importance of specific events in enterprise group demography is very different from their importance in enterprise demography. Births, deaths, survival and employment changes in enterprises are very important in enterprise demography. Births and deaths are far less important for enterprise groups, where the focus is on events reflecting concentration (mergers, takeovers) and restructuring as well as their impact, notably on employment at both national and international level.

The categories of the general typology of demographic events for enterprise groups differ from those for enterprises in two respects.

- In the event of the death of an enterprise group, it loses its identity, whereas redistribution of production factors does not necessarily involve identity loss. This is because the (real) 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. The death applies only to all-resident and multinational groups, not to truncated groups, which may cease to exist through other events.

- The numbers of enterprise groups both before and after the event are decisive. For instance, the birth of an enterprise group and the concentration of two existing enterprise groups differ, among other things, in the number of enterprise groups involved. This is because births and deaths of enterprise groups are not in one to one correspondence with the creation and deletion, respectively, of identity numbers. There are two reasons why not. Firstly, births and deaths are events (concerning the real, observable world) whereas the creation and deletion of identity numbers are SBR updates. Secondly, a birth or a death involves only one enterprise group whereas other events, such as mergers, that involve many enterprise groups may also result in the creation and deletion of identity numbers.

The EU *Business Registers Recommendations Manual (2010 edition)* restricts births and deaths to the involvement of *only one enterprise group*. This convention is chosen because it is in line with the terminology of users who are interested in questions such as ‘What is the effect of enterprise group births on a certain economic activity?’ If the term ‘birth’ is used in the sense of these types of questions, enterprise groups emerging, for instance, from mergers or split-offs are not included, although their impact on concentration (or de-concentration) is relevant.

Existential changes (birth and death)

The birth of an enterprise group is the establishing of control link(s), direct or indirect, between two or more independent legal units, where no control link previously existed, and no other enterprise group is involved. Temporary links of less than one year are not taken into account. Birth applies only to all-resident and multinational groups (which are generally born as all-resident groups), not to truncated groups, which are created through other events.

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. Death applies only to all-resident and multinational groups, not to truncated groups, which cease to exist through other events.

Changes within an enterprise group

Changes may be divided into three categories:

- change of global group head (controlling unit);
- creation/deletion of a truncated group;
- restructuring within an enterprise group.

A change of global group head involves the controlling legal unit being replaced by another legal unit. Recording of the event should include the date when it takes place. This is very important for the statistics of foreign affiliates compiled according to the country of the ultimate controlling unit. Although the legal unit changes, this event should not in itself affect the continuity of the enterprise group. It does not result in the birth or death of an enterprise group. However, depending upon the sources used to create the enterprise group, the enterprise group may be defined according to the controlling legal unit and the source may issue a new identifier in the case of a change of controlling unit.

Creations and deletions of truncated enterprise groups are important special cases of restructuring within a multinational enterprise group. In general, they do not affect the

continuity of the multinational group, but they change its structure and possibly its characteristics. An important case is when there are several control links crossing national borders leading to seemingly more than one truncated group within the same multinational group. There are two possible options for handling truncated groups and their heads. The different groups of legal units thus formed within a national territory can be combined into a single truncated group or each group can be defined as a separate truncated group. As previously discussed in Section 4.3.3, the first option is recommended.

Concentration (merger, takeover)

As in the case of enterprises, if both enterprise groups lose their identity, the event is called a merger. If one of them retains its identity, it is called a takeover. It should also be emphasised that a takeover may lead to changes in some characteristics of the enterprise group that retains its identity. For instance, it may enter a different size class or have a different principal economic activity.

Mergers and takeovers only apply to all-resident and multinational groups. There are no similar types of events for truncated groups. If two seemingly separate parts of a truncated group that belong to the same multinational group become merged under a national group head, this is not a real merger. From a global viewpoint it is restructuring within the enterprise group. Whilst it is desirable to track and record such events, this can be difficult at national level and has to be coordinated at global level.

De-concentration (Split-off, Break-up)

A break-up is an event, where an enterprise group is divided in such a way that none of the resulting enterprises groups retains the identity of the original group. A split-off is an event where one of the resulting enterprise groups retains its identity.

As with mergers and takeovers, break-ups and split-offs apply only to all-resident and multinational groups. There are no similar types of events for truncated groups.

Complex Restructuring

Complex restructuring involves more than one enterprise group before and afterwards. An example is the transfer of an enterprise, parts of enterprises, or a number of enterprises between two or more enterprise groups. If the continuity of employment within an enterprise group is used as a criterion for the continuity, this may lead to discontinuity of an enterprise group even if no change of global group head has been detected.

Restructuring may also affect the characteristics of the enterprise groups that retain their identities. For example, their principal economic activity or size class may change.

The phenomenon of complex restructuring has been prominent in the economic and financial press for many years, so there is certainly a high demand for such information.

Translation of events into SBR updates

Births and deaths

In the SBR, an enterprise group birth is represented by creating a new identity number; and a death by marking the enterprise group as dead. The more difficult problem is determining the date of birth. In the case of all-resident groups, the date of birth is, in principle, the date when the control link is established between two or more legal units. In practice, as the smallest all-resident enterprise groups are not monitored, the date of birth is more likely to be the date when the enterprise group is created in the SBR. The same considerations apply to birth of a multinational enterprise group.

For a truncated enterprise group, the date of birth should, in principle, be the date on which the first unit is established in the national territory. In practice, the date is more likely the date of registration of a corresponding administrative unit with an administrative source.

Changes within an enterprise group

Changes may be divided into three categories:

- Change of global group head (controlling unit);
- Creation/deletion of a truncated group;
- Restructuring within an enterprise group.

A change in group head or restructuring within an enterprise group are events that should not have a direct impact on the demographic characteristics of the group. They are, however, important signals may trigger further investigation. In the case of a multinational enterprise group, they may have considerable impact at national level, even the creation or ceasing of a truncated group, and they may be reflected through changes in relationships or characteristics recorded in the SBR.

As previously noted, if there are several parts of a multinational enterprise group in the national territory, these should be combined and recorded as a single truncated group. This avoids creations and deletions of truncated groups when restructuring takes place within the multinational enterprise group. On the other hand, bringing together truncated groups that have several, seemingly independent parts makes SBR maintenance more complicated at national level.

Mergers and takeovers

In the case of a merger, all the identity numbers of the enterprise groups existing prior to the event are ceased and an identity number is created for the emerging enterprise group.

In the case of a takeover, the enterprise group that takes over the other group(s) retains its identity number, so no creation takes place. The enterprise groups that have been taken over are marked as ceased. The date on which the change is considered to have taken place is the date when both parties have accepted the merger or takeover or when it has been approved by the competition authorities (if this is required).

The creation and ceasing of identity numbers apply in the same way for all-resident and multinational enterprise groups. Truncated enterprise groups simply follow the events at multinational group level.

Break-ups and split-offs

In the event of a break-up, identity numbers are created for all the enterprise groups existing after the event and the identity number of the original group is marked as ceased. In the event of a split-off, a new identity number is assigned to the split-off enterprise group(s).

The creation and ceasing of identity numbers apply in the same way for all-resident and multinational groups. The truncated groups simply follow the events at multinational group level.

Restructuring involving more than one enterprise group

Complex restructuring may entail any number of SBR creations and deletions of different statistical units and, in practice, is very difficult to manage. For example in the European Union, restructuring of multinational groups should be coordinated by the EuroGroups Register whereas restructuring of all-resident groups takes place at national level.

Information sources

In practice, most events are detected as a result of signals from administrative or commercial sources (which may themselves depend upon public administrative sources) or from statistical sources. In some cases, additional data are needed, either collected directly or obtained from other sources.

In all cases, cost-efficiency must be considered. Thus, the SBR updating policy may differentiate between small and large enterprise groups. Small enterprise groups may be updated based on administrative sources, while for large enterprise groups, profiling is usually necessary.

Although administrative sources differ widely from country to country, it is generally the case that creations and cessations in administrative sources, or (especially) in commercial sources, do not necessarily result in SBR updates as they tend to follow entirely different continuity rules. Nevertheless, administrative information about new enterprise groups can be a good start for determining appropriate statistical updates

Reconstruction of demographic events from SBR information

The cause of a SBR creation of a new enterprise group can be birth, merger, break-up, split-off, or restructuring. The cause of a SBR ceasing of an enterprise group can be death, merger, takeover, break-up, or restructuring. Therefore, mere registration of creations and cessations in the SBR does not enable derivation of precise data on the underlying demographic events. More information is needed.

To reconstruct a demographic event, the enterprise groups involved have to be identified. In the cases of concentration, de-concentration and restructuring it is necessary to have links over time between the enterprise groups involved. For example, in the case of a merger, the original enterprise groups must be linked to the emerging group, and in the case of a takeover the enterprise group that is taken over must be linked to the continuing group. If such links are recorded with dates, all events can be reconstructed. This implies the need for a historical register as discussed previously.

Continuity in relation to the definition and use of the enterprise group

Continuity in theory

A general definition of enterprise group continuity based on its definition, uses and central strategic role concerning the units it comprises may be phrased like this:

- *if the enterprise group has continuity in its decision-making on its overall policy for production, sales and profits, financial management and taxation, and has a centralised strategy concerning the units it comprises, it is considered continuous.*

Production factors are important for enterprise groups, although less so than for enterprises. An enterprise group may often use subcontracting for much of its production, thus the physical production factors play a much smaller role, while the roles of strategic planning, research and development and intangibles are very important. This adds to the complexity as the physical production factors are generally much easier to take into account than intangibles.

Continuity in practice: basic rules

The discussion on continuity rules here concerns both multinational groups and all-resident groups, although the rules may need to be applied differently due to availability of information. The continuity of truncated groups in the SBR should follow the continuity of the multinational enterprise group to which they belong, in addition to the events discussed above. When information on the continuity of the multinational group is available, the

continuity of the truncated group can be decided accordingly. Only other events (creation, deletion, etc.) concerning the truncated groups themselves need to be handled separately.

Continuity of multinational groups should be coordinated at international level. For example a procedure in the EuroGroups Register allocates an identity number in a unique format to each multinational group. (An NSI can of course also assign an additional national identity number.) In the case of discontinuity, this identity number is changed. An indication of a discontinuity should come from the country of the group's decision centre when this is in Europe (which is the case for the majority of groups operating in Europe), or when there is a European group head. When the decision centre is outside Europe, a decision regarding continuity has to be based on the available information from (i) the country (if any) where a European group head is located, (ii) the commercial sources, and (iii) published information, for instance in the EU Industrial R&D Investment Scoreboard. The identity numbers and their changes are communicated to the NSIs where the enterprise group operates.

Continuity of all-resident groups can be decided according to cost-efficiency on the basis of different methods but according to agreed rules. For large groups, profiling may be used for the decision. For small groups automated procedures based on administrative and SBR information can be applied.

Only annual substantial changes are taken into account, not slow long-term developments. Changes are found by comparing the situation at a certain time in the year to the same time in the previous year. As in the case of large enterprises, sudden substantial changes in large groups are rare. One such event could be a possible change of group head. On the other hand, smaller changes within the group, or restructuring between groups, are very frequent. They may also occur sub-annually in which case they may not necessarily be taken into account in the annual monitoring of the group.

A key factor to take into account is how the enterprise group itself sees its continuity. There must be strong reasons to go against the group's own opinion, especially because discontinuity among the largest groups is rare.

In determining continuity the following factors are taken into account.

1. Global group head, composition and location(s):

- a) global group head;
- b) enterprises which belong to the group; and
- c) main location(s).

2. Economic activity:

- a) principal activity in NACE or ISIC;
- b) employment; and
- c) intangible assets.

Use of the continuity components

If all these factors change, discontinuity is obvious. In other cases, the changes need to be identified and weighted in order to make a decision. An important case, especially for smaller groups, is when a new group head appears and changes the whole profile of the group, both principal activity and employment, and possibly intangible assets. This event should be considered as the creation of a new group, even if the group head remains the same.

The main factors, namely 1a, 1b, 2a and 2b, may provide a suitable basis for the creation of an algorithm on which an automated decision about continuity can be based.

Change of global group head

The convention is that if there are no changes other than a change of global group head, there is deemed to be continuity of the enterprise group. Thus, change in group head is in itself not sufficient for discontinuity. However, it is an important indication of the need to check if there are changes in the other continuity factors.

Change of principal economic activity

The principal activity may be assumed to be positively correlated with the continuity of the production factors. However, this criterion is probably not relevant for multi-activity enterprise groups, where, relatively small changes to one part of the group may cause a major change to the economic activity code of the group as a whole. Also a gradual shift in activities may (after respecting an appropriate stability rule) eventually cause a change in economic activity code, resulting in a reclassification of the enterprise group. In such cases, there is continuity and the change of activity can be disregarded.

To use the principal economic activity in determining the continuity requires following the mix of economic activities of the group members in terms of value added, if available, otherwise turnover. Employment may also be used, especially in cases where no reliable turnover is available for each enterprise, only consolidated turnover for a VAT group. It has also been found that employment is also generally more stable over time than turnover.⁶⁵

For multi-activity groups, use of principal economic activity in determining continuity is very sensitive to the definition of the principal activity (especially at 4-digit level) and also to whether some pre-defined threshold (e.g. 50 %) is used. Thus, principal activity as a single criterion is not a good approach. However, especially for all-resident groups, a sudden change of principal economic activity when combined with some other changes, for example of group head or employment, is a strong indication of discontinuity. At global level, principal economic activity as a criterion is more difficult to apply.

Change of employment⁶⁶

If the employment of the legal units in the group remains much the same from one year to the next, the group may be considered as continuing.

Continuity of the national group head is strongly correlated with continuity of employment, but as it may change more easily, employment is a better suitable measure.

Using employment is a very practical continuity measure as it can even be totally automated, but it has weaknesses. One weakness is that changes may occur in the legal units due to restructuring within the group. The main weakness is that it is not clear how the approach can work in a global framework due to the lack of availability of the necessary data and their timing.

Change of main location(s)

The locations of the global group head and head office have a strong impact on the enterprises in a group, on the continuity of locations where the research and development are carried out, and on where the actual production takes place. However, unlike the case for

⁶⁵ See the French document presented at 1997 Roundtable:
http://circa.europa.eu/Public/irc/dsis/businesssurvey/library?l=/1997_tokyo&vm=detailed&sb=Title

⁶⁶ Change of employment is based on French document presented at 2005 Roundtable session 4:
http://circa.europa.eu/Public/irc/dsis/businesssurvey/library?l=/2005_cardiff/enterprise_measurement&vm=detail&sb=Title

enterprises, practical rules on how to use main locations as a criterion for enterprise group level continuity are difficult to define, and have not been used.

7.4 Handling Changes in Characteristics of Statistical Units

7.4.1 Changes in Identification and Contact Characteristics

Identification characteristics enable identification of units and their linkage to other units in the SBR and in other sources. They enable tracking of units over time, which is particularly important when there is conflicting information about the dates when changes occurred.

As previously noted, changes in contact characteristics should normally be reflected in SBRs as soon as they are detected as these data are used for mailing survey forms and for geographical analyses of SBR data. Changes to a unit do not necessarily imply changes to linked units.

Changes in legal form, for example from unincorporated enterprise to a limited corporation, may affect the consistency of survey populations and samples over time. Cross-checking name and legal form can help in identifying potential problems.

7.4.2 Changes in Economic/Stratification Characteristics

Economic/stratification characteristics are used in determining the probabilities of units being included in survey samples. It is therefore important to consider whether:

- to update these characteristics as soon as new information is received and risk increased volatility in survey populations, samples and the resulting estimates, or
- to hold the updates back until the point in the annual cycle of surveys when the impact of changes on the consistency of results is minimal, or
- to subject the updates to *stability rules* (also termed *resistance rules*) that inhibit changes, as further discussed in Section 7.5.

When the SBR is used to coordinate surveys, for example through the use of permanent random numbers, the impact of updates on coordination should also be assessed.

Changes to size characteristics (persons employed, number of employees, number of employees in full-time equivalent, and turnover) have significant effects on survey samples.

Large and complex units require more attention because changes to such units have more impact on statistics and there is an increased incentive to ensure the changes are correct. Potential updates should be checked by comparing data from a range of sources to see there is consistency, or by contacting the units concerned directly. As large units are often included in surveys with certainty (i.e., with probability 1), survey feedback is the most common source of updates. Also users of the SBR in particular managers of economic surveys and the national accounts should be consulted and/or informed about updates of large and complex units.

Potential updates to small units have less impact on the resulting statistics and for that reason the procedures for their maintenance can be largely based on automatic updating from administrative sources, without need to contact units to confirm changes.

Changes to geographical location codes for enterprises and local units have to be consistent with the corresponding continuity rules for these units.

Information received on changes in the various activities of a particular unit may prompt a change to the principal or secondary activity code for that unit. These changes may be

sudden, e.g. due to a change in management policy for the business concerned, or gradual, where the balance of activities has shifted over time. In all cases, use of *stability rules* is recommended in order to suppress movements that are no more than temporary phenomena or statistical artefacts.

- For large units, the stability rule may involve a case to case decision by a group of experts.
- For small units, automated rules may be used, for example, a change in principal activity may be made only after evidence for it has been sustained for two years.

For businesses engaged in certain activities, e.g. tourism, there is likely to be a seasonal pattern to the monthly or quarterly numbers of persons employed and turnover. Data for monthly or quarterly reference periods are useful for detecting short-term trends and turning points, however they may give a misleading view of the sizes of units from an annual perspective. It is therefore recommended that *averages over an annual period* are also held for such characteristics and used for sampling purposes.

When stratification characteristics are updated, it is useful to record at least the date of the change, the source of the new value, and the previous value. These are valuable in helping assess the quality of the values of the characteristics and in auditing change processes.

7.5 SBR Maintenance Procedures

Legal and operational units

Updating information about legal and operational units linked to large enterprises is obtained through profiling.

Administrative units

Information about administrative units is obtained from the corresponding administrative sources. The data received should be stored without change in the SBR. Thus in the case of:

- a new administrative unit appearing in incoming administrative data, a corresponding new administrative unit is recorded in the SBR;
- changes in values of the characteristics of administrative units appearing in the incoming administrative data, the values in the corresponding administrative units in the SBR are updated;
- an administrative unit being marked as cancelled, or simply not appearing in the incoming administrative data, the activity status of corresponding administrative unit in the SBR is changed to *dead*.

Statistical units

The maintenance procedures for statistical units depend on the relationships associating legal, operational, and administrative units with statistical units. Specification of updating procedures can be expressed in terms of a matrix with:

- one axis containing all possible types of legal and administrative *signals*, i.e., informative changes, that can be detected in incoming administrative files;
- the second axis containing the all the types of statistical unit change that can occur;
- the cell entries defining the rules for how each signal is to be reflected as an appropriate update to statistical units, taking into account continuity rules, stability rules, and rules for dealing with conflicting information.

Statistical unit birth, death, and continuity rules

Maintenance procedures must incorporate continuity rules identifying the types of legal and administrative signals that result in the births, deaths, or continuations of statistical units, as discussed earlier. These rules determine under what circumstances a unit is deemed to be born, under what circumstances it is deemed to have died and possibly replaced by the birth of a new unit, and under what circumstances it is deemed to continue but possibly in some new form, or under new ownership.

The rules have to cover every possible type of birth, death, or continuation of a unit that can occur, taking into account the source of the signal of a change, and the maintenance group to which the unit belongs. In order to be practically applicable, the rules have to take into account the ways in which changes can actually be detected.

Stability rules (resistance rules)

As previously noted, changes to the enterprise characteristics used for sampling must be subjected to *stability rules* (sometimes called *resistance rules*) that inhibit short-term changes thereby ensuring that changes are permanent before updates are made. These rules have to cover every possible type of change that can occur, and the characteristics involved. They must take into account how the changes can be detected, the types of change (permanent, temporary, seasonal), and the possible impacts on frames and published statistics.

Quality management considerations

When the values of a group of characteristics of statistical units are updated it is important to know how downstream statistical processes will be affected. The impacts of changes on survey frames, and on the statistics produced by surveys that use those frames, have to be considered. In some cases it may be desirable to hold back updates until a certain point in the annual cycle of surveys so that their impact on the consistency of survey estimates is minimised. For large units, such decisions should be made on a case to case basis, involving profilers and users. For small units, rules can be automated.

7.6 Treatment of Errors⁶⁷

7.6.1 Introduction

An error may be defined as "*a difference in the information presented in the SBR and the information as it should be, according to a chosen image of the real world produced and maintained by an accepted instrument and documented procedures*".⁶⁸

It is impossible to avoid errors in the SBR. Indeed, as discussed in connection with stability rules, it may not even be desirable for the SBR to be perfectly up to date at a particular point in time.

Changes to statistical units and their characteristics may be either a reflection of real world events, in which case they are regarded as *updates*, or an amendment of information that was previously wrong, in which case they are regarded as *corrections of errors*.

⁶⁷ This Section is based on the Chapter 18 of the Business Registers Recommendations Manual (Eurostat, 2010).

⁶⁸ Business Registers Recommendations Manual

This section deals with correction of errors and discusses how they should be applied using established and documented procedures.

In essence, there are three steps in handling errors:

- determine whether errors have occurred;
- decide whether they should be corrected; and if so
- decide how and when to correct them.

7.6.2 Taking Account of Different User Needs

The handling of errors should take into consideration the different types of surveys based on the SBR. Procedures appropriate for structural surveys may be different to those appropriate for short-term surveys.

Special provisions for the handling of errors in large and economically significant units may need to be negotiated with users, in particular with the managers of economic surveys and the national accounts. Errors in such units may have a considerable impact on statistics. If a special procedure is implemented to take account of the effects on estimates of corrections to these units, it should be fully documented. Special procedures can be applied by a dedicated profiling team or by appointing an authority (e.g., the SBR manager) who has the final say in making corrections.

7.6.3 Data about Errors and Corrections

As different surveys demand different approaches, it is desirable to record details of each error including source, type and dates of detection, occurrence and correction. This enables SBR users to access both corrected and uncorrected data according to their needs. In practice, depending on the information and resources available, it may be possible to record only some types of error and some dates, or to record these metadata only for core characteristics.

A database may be set up specifically to record the details of errors and associated corrections. The structure and functionality of such a database depends on the procedures agreed for handling errors. The database should be closely linked to the SBR or an integral part of it. Ideally the following metadata should be recorded:

- type of error;
- original value (the wrong one);
- new value (the correct one);
- date of occurrence;
- date of detection;
- date of correction;
- source of corrected value;
- mode of correction (interactive or automatic); and
- name of person making correction (if interactive).

It is also useful to record a *date of confirmation* for characteristic values that were thought to be in error because they looked unusual or out of date, but that have been investigated and confirmed as correct.

Although the inclusion of all this metadata makes SBR maintenance more complex, and results in even more opportunities to make errors, the benefits outweigh the disadvantages.

The information enables re-creation of the SBR at any given past moment in time, thus allowing users to ignore corrections if they need to do so.

7.6.4 *The Time Dimension*

The timing of corrections is important. Corrections of errors in characteristics not subject to stability rules, in particular identification characteristics such as names, addresses, and telephone numbers, should normally be carried out immediately. If not, data collection may be hindered and respondents upset because information they have provided is apparently not being taken into account.

Corrections of errors in economic/stratification characteristics such as economic activity and size are a different matter. If these corrections are made the moment the errors are detected, they may affect the consistency between statistics with different periodicities and they may be indistinguishable from genuine changes in the economic world. To avoid these potential problems, corrections can be stored until an agreed moment, for example once per year, when allowance can be made for them.

Whilst it would be useful to know the moment at which each error occurred, in practice, the date of occurrence is often not known. The date of correction, assuming a correction is made, may be considerably later than the moment of occurrence. If the dates of occurrence and of correction date are in fact recorded, it is possible to construct a population of units for at a point in time in the past that is more accurate than was the actual situation in the SBR at that past time point.

Another question is "How far back in time should corrections be applied?" The answer depends on the correction strategy, not only for the SBR but also for statistics based on the SBR. If, for example, there is a revision policy that requires all statistics to be revised after a period of five years, then corrections should be made as far back as the moment of the previous revision. If such a revision procedure that applies only to the national accounts and not to the individual statistical series on which the accounts are based, then the period for which corrections should be carried out depends on the periodicity of the statistical series.

7.6.5 *Error Correction Policy*

A systematic policy for the handling of errors should take into account the following.

- There should be an inventory of SBR users and uses and the consequences of the various types of errors for the various groups of users.
- The SBR should be structured and maintained in such a way that the correction of errors has a minimal impact on statistical surveys.
- SBR inputs, processes and outputs should be systematically monitored to detect potential errors.
- SBR processes should be fully documented so that the handling of changes or errors in the values of characteristics is clear to all concerned. This helps in detecting errors and avoiding discussions on the quality of individual records.
- The policies regarding handling of errors in the SBR should also be fully documented and audited periodically to make sure they are still appropriate.
- Responsibilities regarding handling of errors should be clear and documented. It is advisable to appoint an authority (e.g. the head of the SBR) who has the final say in difficult cases.

- The handling of corrected values in statistics based on the SBR should be fully coordinated and documented.
- The different types of errors detected should be analysed periodically to monitor changes in the pattern of errors over time and thus to inform future development of policies and procedures for handling errors.
- Recording the history of errors facilitates error handling in complex situations.
- If the SBR is used for administrative or commercial purposes, it is advisable to take legal precautions to avoid damage claims arising from errors.

The source from which a potential correction originates is an important factor in deciding whether and how to correct the error to which it refers. Care must be taken to ensure that the SBR remains a known reflection of reality and that corrections do not lead to distortions. As previously noted, if the source is a sample survey based on the SBR, there is a danger that corrections of stratification characteristics may lead to bias in future survey samples. Corrections from sources that are not in any way related to surveys can be processed without such a problem.

8 Survey Frame Methodology

8.1 Introduction

The SBR is the backbone for economic surveys, being a major part of the infrastructure for practical implementation of processes for compiling economic statistics. This chapter presents a framework that describes the underlying concepts in a systematic way so as to provide the context for elaboration of survey frame methodology. The starting point is the Generic Statistical Business Process Model (GSBPM)⁶⁹ which clarifies the needs for inputs, processes and outputs and provides a framework for their documentation in terms of eight different phases of a statistical production process, as indicated in Figure 8.1.

Figure 8.1: Generic Statistical Business Process Model

Specify needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Build collection instrument	4.1 Select frame and sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult and confirm needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify and code	6.2 Variable outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Configure work flows	4.3 Run collection	5.3 Review and validate	6.3 Scrutinise and explain outputs	7.3 Manage release of dissemination products	8.3 Agree on action plan
1.4 Identify concepts	2.4 Design frame and sample	3.4 Test production system	4.4 Finalise collection	5.4 Edit and impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing and analysis	3.5 Test statistical business process		5.5 Derive new variables and statistical units	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare business case	2.6 Design production systems and workflow	3.6 Finalise production system		5.6 Calculate weights			
				5.7 Calculate aggregates			
				5.8 Finalise data files			

⁶⁹ The GSBPM distinguishes eight high level phases of the process and notes that they are guided by quality and metadata management. See <http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0>

Each phase is subdivided into a number of sub-processes as shown in Figure 8.1. The first phase of the GSBPM is to specify the needs and the output objectives and to identify the concepts needed to design a survey. A survey aims to provide information that fulfils the needs and the demands explored in advance with its users. In this phase an inventory of requirements articulated by stakeholders is translated into a framework that enforces coherence in the subsequent statistical processes and provides practical guidance for the implementation phases. The second phase is to design the complete methodology of a survey in terms of the inputs, processing and outputs including the appropriate statistical unit, the variables to be collected, the data collection method, and the indicators to be produced as output.

This chapter describes survey frame methodology in terms of four sub-processes within these first two phases of the model, namely:

- Establish output objectives (1.3)
- Identify concepts (1.4)
- Design outputs (2.1)
- Design frame and sample (2.4)
- Design collection (2.3)

The other phases and sub-processes are not within the scope of the Guidelines, although it is worth noting that data from the SBR should be available during the third phase when the data collection instrument is being built and the production systems are being finalised.

8.2 Establish Output Objectives

Economic surveys provide information to describe different aspects of an economy. But how should an economy be modelled? A simple economy can be seen as a circular flow of money earned and spent in *markets*, for example, labour market, goods and services markets. Economic indicators describe the *transactors* and *transactions* in these markets.

The transactors involved in the income and the expenditure processes are known as the *households* and the transactors involved in the production processes are known as *enterprises*. The concepts of both types of transactors are based upon units that represent organisations that take inputs from, and produce outputs for, markets. They are known as *institutional units* in the SNA2008. An important characteristic of a transactor is that it represents an independently operating decision centre within a national economy, though the global dimension should also be considered as national borders are becoming less visible.

The simplified model in Figure 8.2 refers to an economy and outlines the flow of income in terms of financial transactions between actors. Each actor needs inputs to produce outputs. The production process conducted by an actor comprises a flow of activities needed to produce the output. In other words production factors are directed through actors. In this respect the actor is the actual transactor in the economy. Figure 8.3 provides more detail.

From the *inputs* and the *outputs* of the transactors, enterprises or households in this simple model, the main relations for compiling the Gross Domestic Product (GDP) can be derived.

The starting point for designing an economic survey is to consider its output, namely a set of *economic indicators*, in relation to a population of *actors*. Important from a practical point of view is the ability to collect the information that has been requested when the organisations are surveyed. This leads to the definition of more homogeneous units within a transactor.

Figure 8.2: Circular flow of income⁷⁰

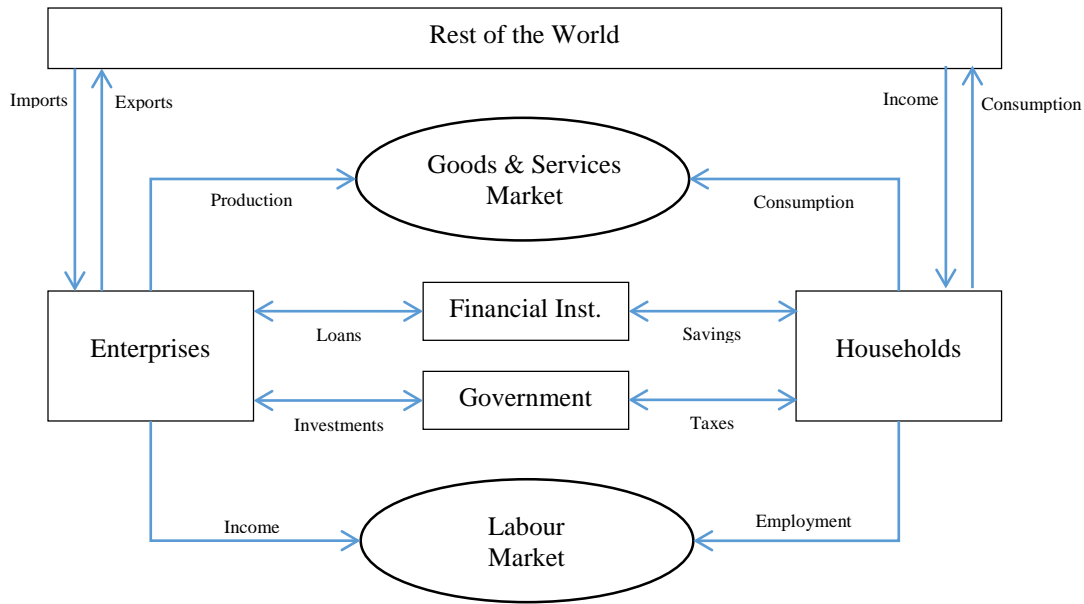


Figure 8.3 Actors, Inputs and Outputs

Actor	Inputs	Outputs
Enterprise	Consumption from the goods and services market Labour from the labour market Imports from the rest of the world Investments or loans from a financial institute Investments or loans from the government	Goods to the market Services to the market Exports to the rest of the world Income to the labour market
Household	Income from the labour market Income from the rest of the world	Savings to a financial institute Taxes to the government Consumption of goods and services on the internal market Consumption of goods and services on the external market Employment to the labour market
Financial institution	Savings from households Savings from Enterprises	Loans/interest to enterprises Loans/interest to households
Government	Taxes from households Taxes from enterprises	Investments to enterprises Services to households
Rest of the world	Export from enterprises Consumption from households	Labour to households Products to households Export to enterprises

⁷⁰ http://en.wikipedia.org/wiki/Circular_flow_of_income

Related to this is the knowledge of the extent to which the available information is consolidated. This is useful to identify the value added by the production process. Also, the geographical dimension should not be forgotten because economic indicators are important for the local and regional economies as well as the national economy. The construction of *local transactors* can account for the consolidation of intermediaries and avoid duplications in the implementation phase. These ideas lead to a *statistical units model* as detailed in Chapters 3 and 4.

8.3 Identify Concepts

It is important to use common definitions in all statistical processes that lead to the planned statistical outputs. For example, it may be desirable to use the changes (by comparison with last year) in the values of a short time indicator (e.g. quarterly production) as a predictor of the overall change in the corresponding annual indicator. Thus, there needs to be coherence between short time and annual indicators and hence between the underlying frames.

A. Let $N(P_1)$, $N(P_2)$ represent a specific frame populations of statistical units extracted from the live register for reference periods P_1 and P_2 respectively, with P_1 being before P_2 . Then the following identities hold:

- $N_{\text{entries}}(P_2) = N(P_2) \setminus N(P_1)$;
- $N_{\text{exits}}(P_1) = N(P_1) \setminus N(P_2)$;
- $N(P_2) = (N(P_1) \setminus N_{\text{exits}}(P_1)) \cup N_{\text{entries}}(P_2)$.

B. Let $Y(P_1)$ represent the sum of a variable (e.g. turnover) based on a frame population $N(P_1)$ for reference period P_1 and $Y(P_2)$ represent the sum of a variable based on a frame population $N(P_2)$ for reference period P_2 , then for the variables based on frame population $N(P)$, where $P = P_1 + P_2$:

- $Y(P) = Y(P_1) + Y(P_2)$
- $Y(P_2) = Y(P_1) + [(Y(P_2) - Y(P_1)) / Y(P_1)] * (Y(P_1))$

The above statistical identities are obvious conceptually, but are not usually satisfied in practice. It is quite likely that the sum of the values of a flow variable such as turnover observed monthly (using possibly a sequence of monthly frames) differs from the annual value of the same variable (observed using an annual frame). There are many reasons for such inconsistency, for example, differences between the monthly data and the annual data as regards concepts, reporting periods, sources and frames.

Differences that can be attributed to the frame populations should be controlled and minimised by the SBR. There should be an established relationship between the monthly survey frames and the corresponding annual frame. This is straightforward for an NSI that renews its frozen frames only annually; more complex where the NSI generates quarterly or monthly frames.

To ensure coherence between the outputs of different surveys (for example, production and employment) for the same reference period, a single frozen frame should be used by all surveys.

8.4 Design Outputs

The outputs of a survey may be specified in terms of a population of interest (e.g. manufacturing establishments, or enterprises with employees) and a set of variables (e.g.

turnover, employment, value added) and a reference period. The survey population is of direct interest to the SBR as it should be derivable from the frozen frames generated from the live register. The choice of variables is also relevant to the SBR insofar as the variables coincide with or relate to characteristics maintained in the SBR.

The SBR itself may be a source as statistics as further discussed in Chapter 9

8.5 Design Frame and Sample

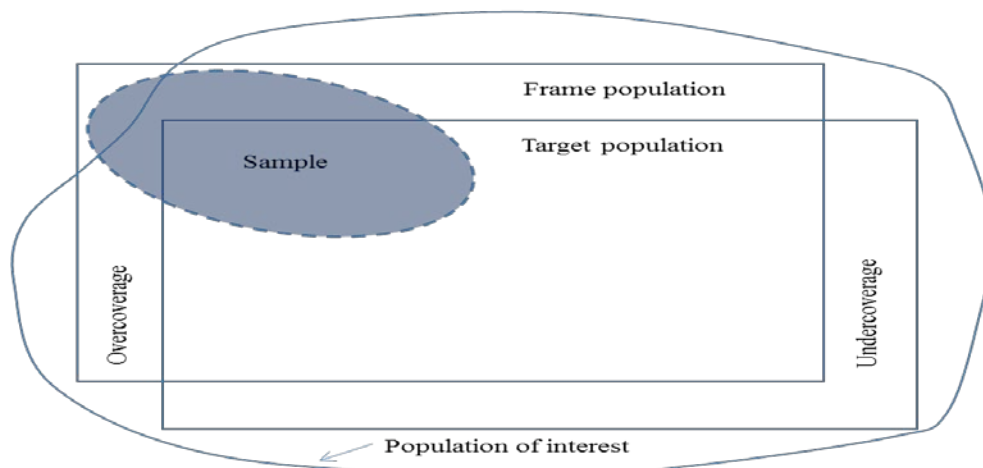
Choice of statistical unit

The target population is the representation of the domain of interest in terms of statistical units in space and time. Statistical units are the building blocks for aggregation of micro-data up to industries, institutional sectors or geographic areas. The type of statistical unit that should be used depends on the economic indicators within the domain of interest. The *enterprise group* is particularly useful for financial analyses and for studying company strategies, but it is too varied in nature and unstable to be adopted as a standard unit for observation and analysis. The *enterprise* is used for financial data and the *kind-of-activity unit (KAU)* is used for production data. *Local unit* or *establishment* is used if indicators are required at regional level. Information on the definitions, derivations and characteristics of statistical units was provided in Chapters 4 and 5.

Target populations, frames and coverage

The population that would exactly meet user needs is termed the (*survey*) *population of interest*. The interpretation of this population in statistical terms is the (*survey*) *target population*. The (*survey*) *frame population* is the best approximation to the target population that can be provided by the NSI. As described in Chapter 2 the survey frame is derived as a subset of a *frozen frame* generated from the *live register*, using unit selection criterion based on characteristics such as economic activity and size class. The *sample* is a subset of the survey frame. Figure 8.4 describes the relationships between these various populations of statistical units, indicating over- and under- coverage.

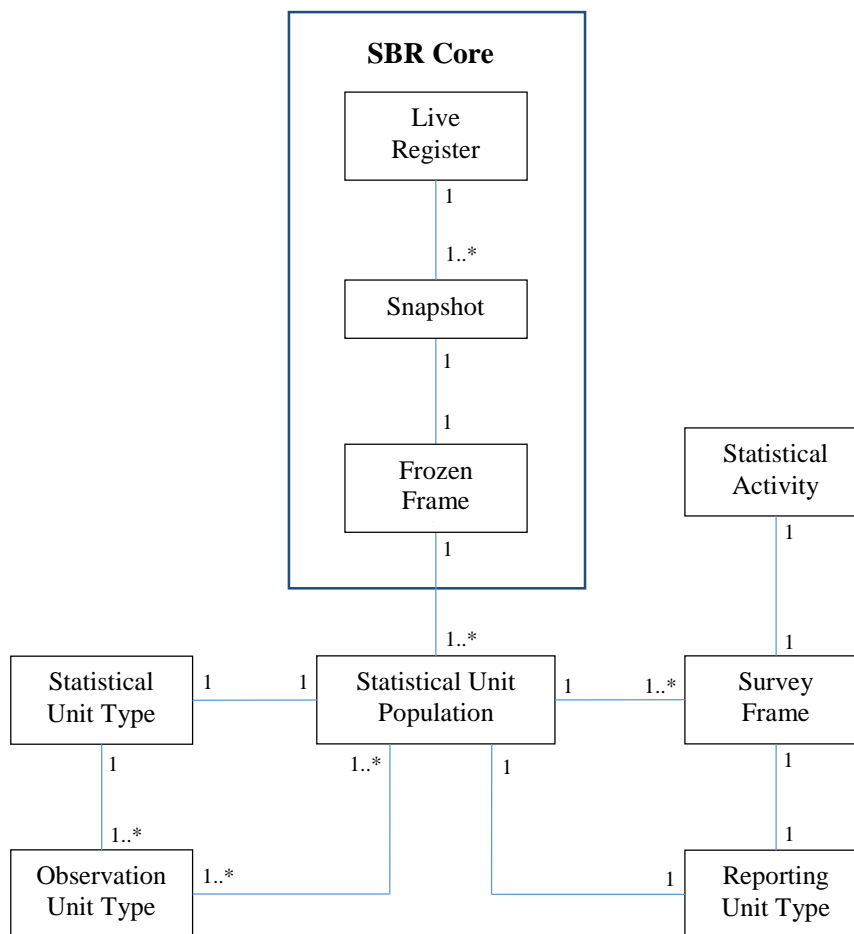
Figure 8.4: Observed, target, and frame populations and sample



Design Frame Methodology

Figure 8.5 indicates the basic steps and data elements involved in creating a survey frame.

Figure 8.5: Creation of a survey frame



The process may be summarised as follows

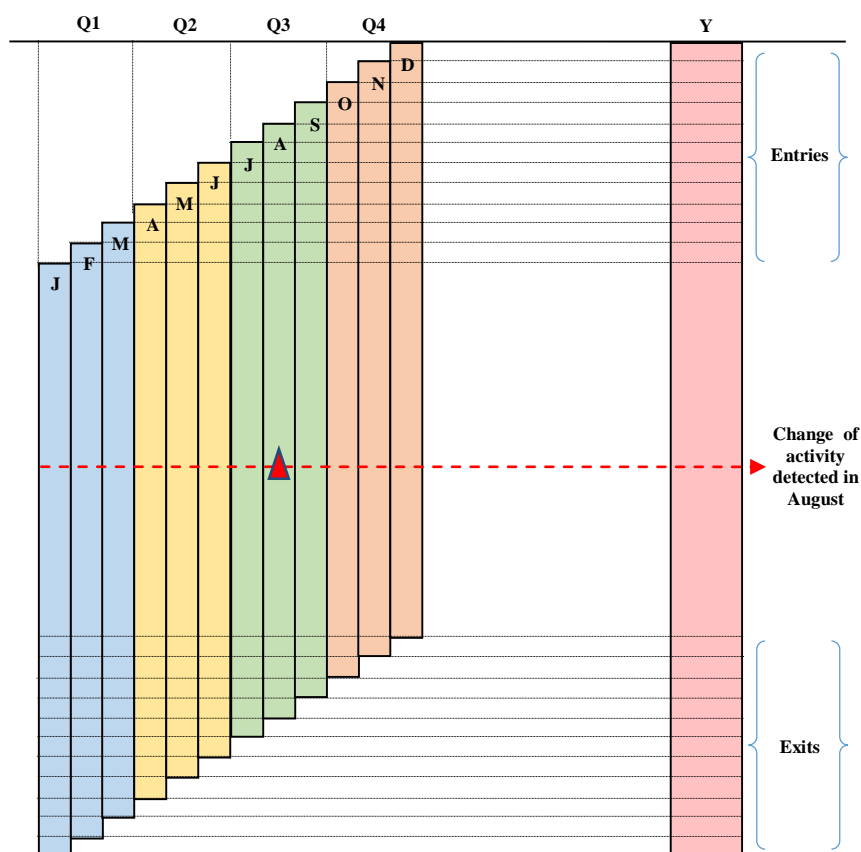
- The live register contains various types of statistical units - enterprise group, enterprise, kind of activity unit, local unit, establishment, etc. The units have been derived and are maintained based on information obtained largely from administrative sources.
- At specified moments in time (month, quarter, or year) a snapshot of the live register is extracted containing all units and their characteristics.
- The frozen frame containing all active statistical units is created by extraction from the snapshot.
- For each type of statistical unit available in the frozen frame, the corresponding set of units may be referred to as the *frame population* for that unit type.
- A survey frame is typically selected from a single frame population and thus consists of a set of one type of statistical unit, e.g. enterprises for a financial survey or establishments for a production survey. (Occasionally a survey frame may be designed using two types of unit.)
- It is possible to derive data for each type of statistical unit using different types of observation units (e.g. by primary observation of the statistical unit, or aggregation of information from the tax office linked to legal units).
- The frames for annual indicators are based on an annual frame population. The frames for collection of short time indicators are ideally based on monthly or quarterly frame populations.

The survey strategy determines at what moment in time a survey frame should be created during the production process. The proposed timeframe for delivery of the output indicators should be specified in the business case. (For example, in Europe the regulations on Short Time Statistics and Structural Business Statistics prescribe the indicators and timeframe for their delivery.) Given the delivery date, the indicators and their quality, the production processes, and the resources available, the moment in time when the survey frame should be available can be computed.

If the frame population for an annual survey about reference period t is also used to observe the short time indicators in year $t+1$, it will not include the information on newly active units for reference year $t+1$. Therefore it may be considered appropriate to create a new version of the frame population at a later point in time for the short term survey. It depends on the availability of sources whether the live register is maintained in such a way that a frozen frame can be extracted more frequently than annually. Some NSIs produce frame populations only annually, some quarterly, and some monthly.

Figure 8.6 illustrates the evolution of monthly frame populations during the course of a year as a result of adding and deactivating units resulting from demographic events.

Figure 8.6: Evolution of frame populations



The annual population may be defined as all units that are or were an element of one of the monthly frame populations during the reference year. In this respect it is important to guarantee consistency among the frame populations by appropriate frame error correction procedures (as further discussed below).

A useful addition to SBR functionality is the possibility of creating improved quality versions of a certain frame population by storing and using a separate list of all frame errors in the live register, including the reference periods to which the errors refer. Based on such a list it is possible to carry corrected values over to the frame population at the time it is needed.

The best way to manage frame errors in a frame population is to avoid their occurrence. This suggests trying to detect anomalies in a frame population in advance before it is used. A good approach is to undertake advance validation based on a set of tables extracted directly from the live register. Units that will appear in the next frame population are compared with the units in the current frame population. Apparent anomalies in the next frame population that could lead to major distortions in the indicators can then be checked (manually) and corrected as appropriate before the new frozen frame is extracted.

Over-coverage and under-coverage errors and their correction

As indicated in Figure 8.4 some statistical units in the target population are not included in the frame population (under-coverage) and some units in the frame population are not in the target population (over-coverage),

Possible reasons for under-coverage are:

- source(s) used to maintain the live register do not include units without employees;
- source(s) used to maintain the live register have size thresholds that exclude smaller unit;
- active units which have been accidentally marked as inactive in the live register or otherwise excluded from the frozen frame;
- recently created units are not yet included in the live register due to the time lag in obtaining and processing information about them;
- recently created units are included in the live register but not in the frozen frame from which the survey frame has been derived.

Possible reasons for over-coverage are:

- non-active units have classified as active in the live register and included in the frozen frame and survey frame.
- some units are duplicates of the same economic entity, which can happen when several administrative sources are used create statistical units in the live register.

Under-coverage is a difficult problem to address because it causes a negative bias in the estimates that cannot be readily estimated from the survey sample. Over-coverage due to the presence of inactive units can be more easily detected during the survey process and allowed for during estimation.

Two ways to deal with over- and under-coverage are:

- to accept their existence in advance (which is reasonable provided they have a minimal effect on the estimation of the indicators); or
- to reduce them by using information from other sources when deriving the survey frame (although ideally this information should be used to update the live register).

If the coverage issues are too large to ignore and cannot be reduced by using other sources of information, then they must be addressed later in the survey process chain by some form of compensation. Known under-coverage should be compensated for by statistical weighting and grossing up procedures or by extrapolation from census results. Over-coverage may be estimated and corrected by conducting a separate survey with the specific aim of estimating the number of units in the population.

Misclassification by principal economic activity

Misclassification of principal activity of units in a frozen frame leads to under- or over-coverage in the subsets of the frame corresponding to specific industries. Special efforts should be made to correct for this kind of error as it leads to increase in variance (due to over-coverage) and bias (due to under-coverage) in the following way. After a frozen frame is produced, a survey frame derived, and a survey sample is drawn, the resulting observation units complete a questionnaire.

- It may turn out that a respondent reports that the registered activity code of the observation unit is not correct. In this case this frame error should be corrected in the survey dataset, because otherwise the observation will be attributed to the wrong industry. This does not cause bias but increases the variance of estimates.
- Alternatively, it may be that units were not included in the survey frame because of being misclassified to out of scope industries. This introduces negative bias in the estimates.

Dealing with frame errors

Frame errors may cause errors in estimates of *levels* of indicators, also, and sometimes more substantially, in estimates of *changes* in indicators over time. When indicator values are compared for different quarters of the same reference year, or for the same quarters of different reference years, it is important to be able to assume that the changes observed correspond to changes in the real economic world and are not just artefacts resulting from the effects of frame errors or corrections of frame errors. Also, if a change (by comparison with last year) in a short time indicator (e.g. quarterly production) is to be used as a predictor for the overall change in the corresponding annual indicator, then it is equally important that the change represents a *real change* in the population and not the result of an error or correction an error.

The way errors in the frame are to be treated should be specified in advance in an official *frame error procedure*. The aim of a frame error procedure is:

- to support industry based statistics in providing frame populations that closely approximate the industries in the real economic world;
- to support coordination of data collection and processing in different industries and in different surveys.

Although it is impossible to monitor all the changes that occur in the economic world in real time, it is possible to correct for those frame errors that are likely to cause major distortions in economic indicators, preferably well before publication. These errors can be categorised as

- Errors that are present for a rather long time in the live register and are detected more or less by accident, for example during data collection and/or when a respondent makes contact. These kinds of errors tend to affect the estimates of indicator levels.
- Errors which distort the relationships between, monthly, quarterly and annual indicators. These kinds of errors affect changes in indicators between reference periods.

An extra complication is that, on occasion, correction of a frame error may result in an apparently anomalous change in the value of an indicator, which itself would then have to be corrected.

Sample Design

The sample is selected in accordance with a predefined sample design, which can be quite different from one survey to another. Appropriate stratification characteristics help in

selecting a sample that minimises the sampling error for a particular indicator. All the stratification characteristics required for sampling should be available in the survey frame.

Inclusion of inactive units in the sample leads to difficulties in obtaining responses during the observation phase and loss of effective sample size. The sample size should be chosen with this in mind.

8.6 Design Collection

The data collection methodology should be designed to optimise survey results in terms of costs, efficiency and accuracy of the indicators. The availability of one or more sources that have the required information for observation units influences the data collection process and the composition of the statistical indicators. It is important to decide in advance whether to collect information directly from statistical units or to use administrative data from linked observation units.

Observation and reporting units

Besides representing the transactors in an economy in terms of statistical units it is useful to introduce the concepts of an *observation unit*, about which data are obtained, and on the basis of which data about a statistical unit can be derived, and a *reporting unit* from which data about the observation unit are obtained. Thus a *survey frame* consists not only of a set of statistical units, but also corresponding *observation units* and *reporting units* together with appropriate contact information.

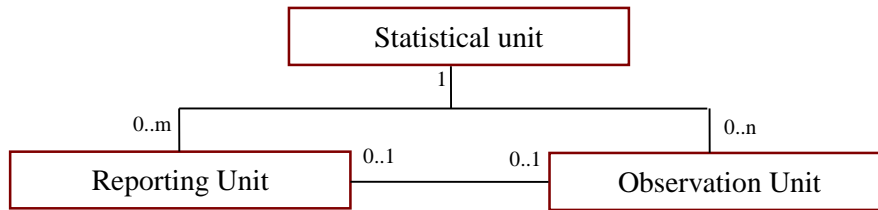
In most cases the statistical unit, the observation unit and the reporting unit are the same, but there are some cases where they are all different:

- because of book-keeping practices no data is available from the target unit, only at another level of the organisation;
- another organisation (e.g. accountancy firms) is responsible for reporting;
- the partial autonomy of the unit;
- data are collected from a group of the units to save costs;
- there is no frame information about the target unit, only the unit to which it belongs.

The relationships between these types of units is indicated in Figure 8.7.

The creation of different observation and reporting units to obtain data is based on discussion with the respondent about how best to obtain the requested information.

Figure 8.7: Statistical, Observation and Reporting Units

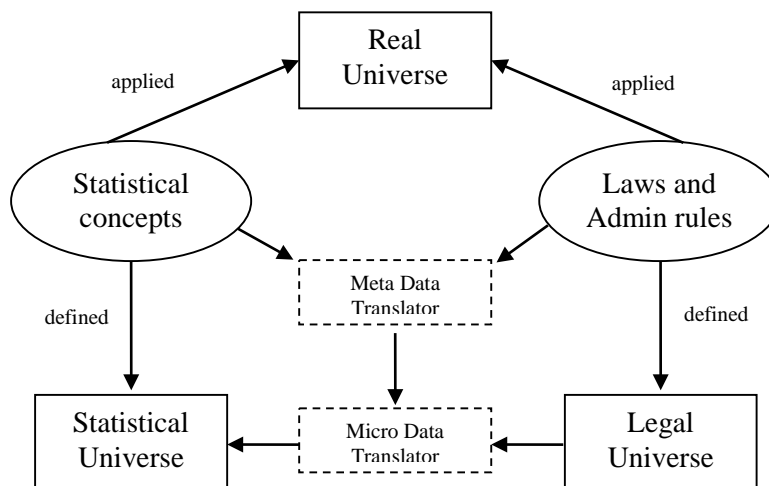


Use of administrative data in place of directly collected data

The main problems that arise in using administrative data in place of direct data collection as illustrated in Figure 8.8 are:

- the differences between statistical concepts and the concepts defined by administrative rules and laws, which results in the need to identify the correspondences in terms of a *metadata translator*; and
- the need to transform administrative data into statistical data using an appropriate statistical methodology, i.e., *micro-data translator*.

Figure 8.8: Relationship between statistical and administrative universes



An example is the desire to disseminate statistics on turnover. The total amount of turnover could be derived from the administration of VAT revenues. However, this derivation might lead to serious inconsistencies for very large enterprises because of:

- their special treatment by the taxation authorities; or
- consolidation effects: or
- global influences that distinguish them from small and medium enterprises in their way to pay taxes.

In this case the metadata-translator should make an appropriate adjustment for large enterprises, or the conclusion drawn that such translation is not feasible and that VAT data cannot be used to derive total turnover for large enterprises. A possible solution might be to use administrative data.

9 Dissemination

9.1 Introduction

This chapter deals with direct dissemination of SBR data to the general public. It covers two roles defined in Chapter 2, namely *Role 5: SBR Statistics* and *Role 6: SBR Information Services*.

In the past SBR data were not usually disseminated directly; rather they contributed indirectly to published statistics that were produced by economic surveys for which the SBR provided the survey frames. However, due to increased use of administrative data and increased computing and storage capacity, the SBR is now often a powerful database from which statistical data can be directly disseminated, or which can be used to supplement data collected by surveys, thereby replacing or reducing the amount of survey data collection.

The SBR may be characterized as comprising a lot of detailed information of high value. In principle, it covers all enterprises, and related units such as establishments, in the formal sector. This includes government units and non-profit organisations. For all these units it contains the values of basic characteristics such as detailed economic activity classification, regional classification and measures of size in terms of turnover and employment. While this range of characteristics is not so extensive as can be obtained by survey, the SBR is nevertheless a unique, rich database as it provides coverage that, in the past, could only be achieved by an economic census. In fact, in several countries the SBR is replacing an economic census, or at least supporting it by providing the frame.

In addition to directly publishing SBR data as economic statistics, the second form of dissemination is of business demographic data showing the creation, continuation (survival) and cessation of enterprises and/or establishments. Such data may be provided by industry, region, legal form, size class and other stratification variables. In this context, the specific unit and continuity definitions used in the SBR are of crucial importance.

A further form of dissemination, corresponding to the role of *SBR as an Information Service* referred to in Chapter 2, is when *individual data* (also termed *unit record level data* or *micro-data*) are made available outside the NSI. This is possible only in countries where it is lawful to disseminate selected data about individual units to other organisations and, possibly, to the general public. Such a release provision may not include all types of units and almost certainly will not extend to all characteristics. Typically it will apply to legal units and/or local units of legal units, and the individual data items that can be released may include name, address, legal form, economic activity code and, possibly, size code.

Dissemination may be in the form of lists or directories of units, either in print or electronically, on-line or off-line, publicly available, or for the organisations specifically requesting them. Such dissemination may be available free of charge or for a fee.

Prior to contemplating this form of dissemination an NSI must establish if there is such an appropriate release provision, and if so, observe it. Otherwise the NSI will be in breach of confidentiality requirements.

9.2 Dissemination of Economic Statistics

9.2.1 *Dissemination Conditions*

Statistical information is considered to be a public good. Thus, direct dissemination of economic statistics from the SBR to users should take place without discrimination regarding

type of user and with simultaneous access to all users. Constant improvements in technology provide increasingly better tools for data communication, and these should be used.

Before SBR data are disseminated their quality needs to be assessed. The data should be sufficiently accurate, timely and coherent that they adequately represent the actual economic structure in the country in those aspects that are disseminated. They should provide good coverage of the economy, industrial structure, regional distribution, size structure, etc.

Given appropriate data quality, one way of disseminating SBR data aggregates is to release regular standard reports, each comprising a set of tables. The reports can be produced annually, quarterly or monthly, depending on data quality, which in turn depends upon SBR creation and maintenance procedures.

9.2.2 Confidentiality Considerations

As in all statistical data dissemination, the need for statistical confidentiality must be taken into account. The broad level requirement is that no data should be published that would allow a user to determine the value of a characteristic for a specific unit. The usual approach in ensuring confidentiality is to publish only data values that refer to more than a specified number of units, typically three or more units. In a table where a certain data cell refers to less than this number of units, the cell value is suppressed and replaced by a symbol indicating that the value is available but *confidential*.

Furthermore, suppression of confidential cells must be accompanied by further cell suppression to prevent the values in the suppressed cells being calculated from the values in other cells by a process referred to as *residual disclosure*. For example, consider a two dimensional table including row and column totals and suppose the one cell value has been suppressed because it is confidential. This value can readily be computed from the other cell values in the same row (or column) and the row (or column) total. Thus the values of other cells that are not themselves confidential have to be suppressed. This is referred to as *complementary suppression*. There are typically many options for the choice of additional cells to suppress and there are algorithms to guide this choice. For example the aim may be to choose the data cells with lowest numbers of units for complementary suppression.

In summary, before publishing a set of tables, confidential cells need to be identified and the values in these cells suppressed. Then the tables need to be reviewed and further cell values suppressed to prevent residual disclosure. Computer applications to do this are available.

9.3 Dissemination of Business Demographics

9.3.1 Background

Business demographic statistics provide data on the numbers of births and deaths of enterprises (and/or establishments) in a specific period, and on the number enterprises (or establishments) that were born in a previous period and continued in (i.e., survived to) the specific period. These data are usually structured according to activity, legal form, size class, region and various other stratification characteristics. The main descriptive characteristics are employment and turnover. Thus it can be shown how many jobs were created by the newly born enterprises as well as how many jobs were lost due to enterprise deaths.

Another important feature of business demography statistics is the analysis of the development of the newly born enterprises over the following years: how many enterprises born in t were still active in $t+1$, $t+2$, ... $t+5$.

All these demographic data about enterprises should also be related to the appropriate total numbers of enterprises to produce birth rates, death rates and survival rates.

Business demographics are sought by analysts and policy makers concerned with business dynamics and entrepreneurship, for example to assess the impact of policy initiatives on enterprise birth and survival rates.

Measuring entrepreneurship with SBR: the OECD-Eurostat Entrepreneurship Indicators Programme

Even though the role of entrepreneurship as a driver of economic growth and job creation entered the policy debate some decades ago, sound international evidence on the entrepreneurial phenomenon and its determinants and impacts there is not much reliable data for performing comparative analysis. To respond to the need of internationally comparable official statistics the OECD-Eurostat Entrepreneurship Indicators Programme (EIP) was launched in 2007.

The EIP is aimed at the development of policy-relevant and internationally-comparable indicators of entrepreneurship and its determinants, in order to support the analysis of entrepreneurship. To that purpose, the programme has developed a conceptual framework and a methodology for the collection of harmonised entrepreneurship statistics. Two elements constitute the characterising features of the EIP. Firstly, the core set of entrepreneurship performance indicators collected by the programme consists of *business demography statistics on the birth, death, survival and growth of enterprises* (as distinct from other approaches to entrepreneurship measurement that focus, instead, on data on individuals). Secondly, *the SBR is used to compute business demography statistics*, according to the methodology presented in the *Eurostat-OECD Manual on Business Demography Statistics* (2007) – a manual developed by the EIP.

The EIP focuses on "*employer business demography*", where the relevant statistical unit is the enterprise with a least one employee. Employer indicators are found to be more relevant for international comparisons than indicators covering all enterprises, as the latter are more sensitive to the coverage of SBR (e.g. when employment (in case of employment thresholds, the basic criteria of employer demography would not be fulfilled) or turnover thresholds are applied by the main sources used to compile SBR).

The development of a database of comparable business demography statistics was much welcomed by policy makers and the research community. The results are also disseminated through the publication *OECD Entrepreneurship at a Glance*.

Using the SBR to produce business demographics requires:

- a typology of demographic events that covers all the main demographic events affecting statistical units and takes account of the links between them;
- definitions of the main demographic events, specifying precisely their characteristics;
- methodology and operational rules for computing the statistics.

Typology of demographic events: Example from the European Union

The Business Registers Recommendations Manual (2010 Edition) of the European Union provides a typology developed specifically for enterprises. It is based primarily on changes in the existence of production factors and their distribution within and among business organisations, while taking into account enterprise structure in terms of organisational units of production, economic activities, locations, and legal/financial links.

As described in Chapter 7, events are separated into two groups:

- those involving existential changes, i.e. the emergence or disappearance of combinations of production factors; and
- those involving distribution changes, i.e. changes in the distribution of production factors between units.

As indicated in the following table (where n implies 2 or more) existential changes (births and deaths) involve a transition from no enterprise to one enterprise, or vice versa, whereas changes in the distribution of production factors require that at least one enterprise is present both before and after the event.

<i>Event</i>	Real, observable world		Business register	
	<i>Number of enterprises before the event</i>	<i>Number of enterprises after the event</i>	<i>Number of creations</i>	<i>Number of deletions</i>
Birth	-	1	1	-
Death	1	-	-	1
Change of ownership	1	1	-	-
Merger	n	1	1	n
Takeover	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

9.3.2 Advantages of Using the SBR

Coverage

In principle, business demographic statistics can be compiled from sources other than the SBR, in particular business censuses and surveys. However, there are limitations that make these other approaches less convenient. With respect to censuses, the coverage of units is typically very comprehensive, but there is a major problem with census data periodicity

(generally every five years or less frequently). Such infrequent data are a source of comparability issues.

While survey based approaches to business demography may be useful, for example to capture the creation of informal enterprises, they suffer from the usual constraints of survey errors and of sample sizes that limit detailed data breakdowns. Also, it is difficult to identify enterprise deaths by means of surveys, as dead units are not available to be surveyed. Moreover, in calculating business demography rates that require an estimate of the total population of enterprises as the denominator, conceptual consistency between the denominator and numerator populations can be realistically achieved only by using the same source of information for both.

- A survey may be used to estimate both numerator and denominator but in compiling the latter, there is a risk of multiple counting of enterprises, and survey respondents may be asked to differentiate between statistical units with which they are not familiar.
- If a survey is used to estimate births and the SBR is used for the total population, there is a risk that inconsistencies between the numerator and denominator arise, for example at the industrial sector level or because the numerator includes births of informal enterprises that not included in the SBR population.

Finally, using data from an SBR is generally quicker and cheaper than conducting a survey and minimises the burden on businesses.

International Comparability

Thanks to the progressive harmonisation in the basic requirements of SBRs across countries, the SBR constitutes the most convenient source of business demographics from the perspective of international comparisons. However, it is important to acknowledge some potential limitations in this respect. Specifically, the various thresholds used in SBRs are perhaps the most important source of incomparability in business demographic statistics. Although SBRs aim for comprehensive coverage, in practice, they each have to use some sort of threshold that excludes small or difficult to find enterprises. Typically, thresholds are based on monetary values, using turnover as the indicator for example, or they are based on employment levels, or on registration with one or more administrative sources. They may be dependent on other criteria, for example they may exclude agriculture and/or unincorporated enterprises. The net result is many variations and corresponding lack of comparability across countries. One of the aims of these Guidelines is to reduce the reasons for lack of comparability.

An additional source of incomparability is that the appearance of a business on administrative register (and hence the appearance of an enterprise in the SBR) does not necessarily coincide with the date at which the business first becomes active. Furthermore, this can vary by country. For example, in some countries, a business may be required to register, or may voluntarily register, with an administrative source before any production occurs or turnover is recorded. In fact, registration does not necessarily mean that the businesses will ever be economically activity. It may instead remain permanently inactive. In other countries, administrative sources register a business only after it is already been active for a while, and perhaps only after its production exceeds some threshold, commonly based on turnover or employment.

9.4 Micro-Data Dissemination

Access to individual unit data in the SBR is restricted by confidentiality requirements. However, in recent years, some countries have been able to give access to such micro-data to

academic institutions, researchers, and students, but always under a number of conditions to ensure its confidentiality.

For some research purposes the SBR data alone may not be sufficient and other individual data may be linked to the SBR data, depending on the research goals. Here again, an adequate legal basis must be assured.

Restricted access to individual unit data is not equal to publication of these data. However, as previously noted, in some countries publication of such data from the SBR is allowed by law. The content of such a publication is usually restricted to identification code, name, address and legal form of the enterprise, economic activity code, and possible other classification codes. Usually, no quantitative economic data, such as employment or turnover are included. The coverage of individual units may also be restricted to companies and government units, or even to bigger units only. The intervals at which such data are updated may vary from daily to annually. Usually, such data are published online.

National Directory of Economic Units (DENUE) in Mexico

Although its main function is to serve as a sampling frame, the National Directory of Economic Units (DENUE) has made individual statistical unit data available to all, as a public good, and not only for decision makers at government or political level. With the benefit of modern technology, DENUE provides readily accessible information on the distribution of the economic activities and establishments by territory throughout the country.

In addition to the data that are openly shared, DENUE contains data for internal work purposes only, for example, to identify records of the establishments in more than one database or to solve errors among the databases used.

See: <http://www3.inegi.org.mx/sistemas/mapa/denue/default.aspx>

There are other countries that provide access to individual unit identification and location data because this information has been disseminated by the units themselves as part of their contact and promoting information. These data can be used for public and private policy planning, as well as for economic research. They include identification characteristics, stratification by size, activity class and geographical location of the establishments in the national territory. Typically they are made available under the general principle that statistical information should be considered a public good and should be made available to everyone in a simultaneous way and without detriment to the protection of personal data.

10 Quality

10.1 Introduction

Quality may be defined as *the degree to which a set of inherent characteristics fulfils requirements*⁷¹, and thought of in general terms as *fitness for purpose*. The SBR purposes are to provide:

- information that enables identification of statistical units;
- populations of statistical units for survey frames and sampling;
- statistics on the structure of the economy – in terms of units and their economic classifications;
- data for business demography analysis; and
- tools for using administrative data for statistical purposes.

It achieves these purposes through the eight roles described in Chapter 2.

The standard *quality dimensions* used in assessing a survey can be used in assessing the SBR. For example, users want SBR data to be relevant, accurate and timely. *Relevant* means that the data meet current and potential needs of the users in terms of units and characteristics to support the production of statistics. *Accurate* means that the information recorded corresponds to the reality. *Timely* implies that the data provide a picture of the real world with the least possible time lag. However, there are significant differences between assessing the quality of a survey and assessing the quality of the SBR due to the wide range of inputs to the SBR, the multiple types of units it contains, the fact that the most important outputs are micro-data not statistical aggregates, and the importance of links to units in other databases. Thus, it is not possible to use exactly the same approach for measuring SBR quality as for measuring survey quality.

SBR data quality is closely related to how the data are used and to whether they satisfy the users' needs. However, the SBR has several different groups of users and each group has its own needs. The requirements with respect to frozen frames as regards timing and accuracy are different depending upon whether the frame is to be used for a short term survey or a structural business survey. For example, if value added is being estimated by a survey based on an SBR frame or using administrative data, the accuracy of the values of key characteristics such as economic activity code and size code for large units will be paramount given the economic significance of these units. On the other hand, if business demographic indicators are based on the SBR, the quality of the smaller units will be very important due to their high rate of involvement in demographic events. In summary, as the SBR is complex and there are numerous and different groups of users of its products, the criteria for evaluating its quality are also complex.

The differing requests from differing groups of users result in conflicting demands regarding the timing with which the SBR records changes and delivers frames. The solution, as already described, is to maintain a *live register* that reflects the latest available information and to produce a set of *frozen frames* each pertaining to a particular reference period. The frozen frames serve for sampling and coordinating the results and for business demographics. They can be extracted as frequently as needed (monthly, quarterly) in order to support surveys

⁷¹ ISO 9000: 2005 Para 3.1.1

according to their timetables, and include corrections of previous mistakes, new units, and recent updates to key characteristics. The live register is continuously updated and can be accessed for the very latest information when required.

10.2 SBR Quality Dimensions

Quality dimensions (sometimes called quality components) are *the concepts used to describe some part or facet of the overall concept of quality, when applied to statistical outputs*⁷².

Although articulated primarily with censuses, surveys and administrative collections in mind this set of dimensions applies equally well to an SBR. The need for an SBR to be relevant, accurate and timely has already been discussed.

As regards *punctuality*, SBR users certainly need to know when to expect the frozen figures and statistics based on the SBR.

As regards *accessibility*, the possibility of internal users obtaining individual data by directly connecting to the database, and external users obtaining aggregate tables directly from the NSI output database, should be considered.

In terms of *comparability*, there are two aspects to consider: region and time. Comparability requires that the concepts and methods underlying the maintenance of the units and their characteristics do not change across regions or over time.

Example from Europe

In Europe comparability over region is promoted at a European level through a regulation, so the quality of the SBR in an EU country in this respect can be measured in terms of level of adherence to the regulation.

Coherence includes both internal coherence of units and characteristics within the SBR and coherence with other registers. While internal coherence involves a consistent treatment of the SBR data, coherence with other registers is promoted by creating and maintaining links. The use of a common identification code across all official business registers (administrative and statistical) is an excellent way to obtain greater coherence.

Finally, even though it is not a quality criterion, consideration of *cost* must accompany measurement of quality. *Cost* is a *quality constraint* and must be considered when allocating resources to improve any quality dimension.

Cost includes burden on enterprises as well as cost to the NSI. It is important to ensure that enterprises are not routinely obliged to provide the same data to the SBR as they have already given to an administrative source. This does not preclude the possibility of conducting an SBR quality improvement survey to check or improve the characteristics of SBR units.

Although described separately, the various quality dimensions are inter-related in the sense that attempts to improve the SBR with respect to one dimension may lead to deterioration with respect to another.

Quality measurements, involve the availability of metadata necessary to correctly understand the information.

⁷² United Nations Statistics Division: National Quality Assurance Frameworks

10.3 Difference between Survey Quality and SBR Quality

In the context of an NSI, *quality assurance* can be defined as *all the planned and systematic activities implemented that can be demonstrated to provide confidence that the processes will fulfil the requirements for the statistical output*⁷³. Quality assurance procedures have been intensely studied and developed in relation to the conduct of surveys. Whilst many of the resulting concepts and methods may be applicable also to the SBR, the SBR has certain specific characteristics that distinguish it from standard surveys and that require different quality assurance procedures. An example is the heterogeneity of SBR users. Another example is that the primary SBR output is individual unit level data not statistical aggregates. The following paragraphs describe some of these specificities and their impact on the quality issues.

10.3.1 Use of Administrative Data

The use of administrative data for statistical purposes has increased in the last decades. The SBR is a statistical process for which massive input of administrative data is the main feature. Thus SBR quality is strongly linked to the quality of the administrative data it uses, over the generation of which it has limited or no control. This is different from a survey process that is under NSI control. For example, the SBR has little or no influence on the definitions of the characteristics that an administrative source provides, whereas it determines the definitions of variables it collects by survey.

10.3.2 Heterogeneity of Inputs

SBR construction and maintenance require data from many sources (administrative and statistical) to be integrated. Each source provides only partial information with regard to the units and their characteristics, so no single source can meet all needs. Sub-populations of units are covered by different sources. For example, data about units in the agricultural sector may be derived from a quite different source than the sources that provide data about units in manufacturing or services. The values of different characteristics may be acquired from different sources, for example, turnover from VAT declarations, employment from the Social Security Register. Information on the large and complex units are typically obtained by surveys or profiling, whereas information on small units are obtained using (often exclusively) administrative data.

Thus, an approach to quality that is appropriate for a survey based essentially on data collection from a sample is not sufficient in assessing SBR quality. Rather the approach must cover the treatment and quality evaluation of data from a variety of sources. In this context, quality evaluation of the SBR as a whole may be difficult. In the first instance, it is necessary to split SBR data into subsets (for example, according to source, unit type, and/or group of characteristics) and to develop a set of quality indicators applicable for each subset.

10.3.3 Variability of Inputs over Time

In the case of a repeating survey, the process and the data items collected do not usually vary much over time, and when there are changes to the process, they are under survey control.

73 United Nations Statistics Division: National Quality Assurance Frameworks

Many of the problems that arise when using administrative data can be related to changes in processing or data made by the administrative source, some of which may not even be known to the SBR. There may be changes in registration and cancellation rules, in the units involved and/or their classifications, in the data collected, and in the control processes. Therefore, an important objective of the SBR is to verify the stability of the administrative process and data, and thus to avoid having changes that are purely administrative causing structural changes in SBR data that do not reflect changes in the real economic world.

10.3.4 Technological Requirements

The processes for construction and maintenance of an SBR are characterized by:

- quite large amounts of data from many different sources;
- changes over time in administrative source contents;
- complex procedures for data integration, processing and production of outputs; and
- changes over time in classifications and output data requirements.

Thus SBR procedures and systems must be adaptable over time. Control and evaluation of the technological aspects of SBR processes is a central task in SBR quality assurance.

Given that all NSIs require much the same SBR fundamental components, the scope for industrialisation of SBRs based on international best practice is an important consideration. In particular development of a generic SBR system that could be readily adopted to the particular circumstances in a country might prove very useful for NSIs with limited resources to undertake SBR development.

10.3.5 Primary Output is Micro-Data

The main objective of the SBR is to produce individual data – data about individual units - for use by surveys in constructing survey frames and samples. Whereas with statistical aggregates random errors tend to average out, with micro-data there is no such notion. Over-coverage and under-coverage errors do not balance out. The effects of both have to be evaluated.

10.3.6 Heterogeneity of Users

The quality policy adopted by the SBR must take into account the wide range of users' needs. Given that the primary SBR role is producing sampling frames for surveys, the SBR update processes have to be as timely as possible, have to guarantee a stable reference population, and have to provide the most complete coverage possible. For surveys producing macro-economic variables such as value added, the coverage of large units and the accuracy of characteristics, such as economic activity code and size, is critical. On the other hand, coverage of smaller units is very important in producing business demographic indicators.

Another important aspect of quality is coherence. Whereas individual surveys might optimise satisfaction of their external user needs by using their own specially defined units and characteristics this is not the best approach. It is far more important to ensure all surveys use the same conceptual and physical set of frozen frames generated by the SBR as this leads to more coherent and cohesive economic statistics. It is much easier for national accounts staff to reconcile statistics coming from the various economic surveys programs when all these surveys start from the same base.

Whereas for internal users accuracy at micro level is important (for example using a wrong address for a unit may result in legal repercussions), for external users accuracy and comparability of SBR derived aggregates are more important.

10.3.7 Continuous Updating of SBR Data

Updating a unit in a SBR requires deciding how to identify and treat the actual changes in the characteristics that have occurred. Such changes can involve the unit's existence, its characteristics and the links between units. Events should be referenced to a specific time point or period. During a given time period, say (t, t+1), the incoming/outgoing flows of units to the SBR, and to the survey frames generated from it, are determined, first, by births and deaths of enterprises, and, second, by changes in the classification characteristics.

The updating of the SBR over the period (t, t+1) is determined not only by actual economic changes occurring in the real world, but also by correction and updating of characteristics of the units following from SBR maintenance processes. Thus, SBR data cannot be considered in the same way as survey data. A survey does not acquire any additional information once it is concluded, whilst the SBR may acquire data during the period (t, t+1) that refer to a time even earlier than (t).

More specifically, in the SBR it is possible to correct retroactively errors in the values at time (t) of characteristics like economic activity code, measure of size, or date at which a cessation or a birth is registered. Changes are virtually always recorded at a later date than they actually occurred. For example, there can be a delay in recording births and deaths or in recording changes in characteristics in the administrative registers used for updating the SBR. Effective and timely updating of a SBR depends to a large extent on how the administrative sources register the information that they receive from the enterprises. There may be errors and delays, especially in connection with small enterprises.

Continuous updating of SBR means there is a possibility of revising any dataset produced by the SBR that is stratified by activity status, or other key characteristics such as economic activity code or number of employees taking into account these two different kinds of changes - actual changes and error corrections. Distinguishing real economic changes from error corrections and delayed updates is a fundamental task in assessing and improving SBR quality.

10.4 Assessing the Quality of Administrative Data for the SBR

SBR quality assessment involves taking into account the entire process of acquisition, loading and processing of administrative data. Administrative sources can be assessed in terms of the usual standard quality dimensions. Alternatively, or as well, the quality framework developed for registers by Daas et al (2009)⁷⁴ may prove useful. The quality dimensions are expressed in terms of *hyper dimensions* (Karr et al., 2006)⁷⁵, which are *source, metadata, and data*. Each hyper dimension is composed of several dimensions of quality and each dimension contains a number of quality indicators.

⁷⁴ Checklist for the Quality Evaluation of Administrative Data Sources (2009), Piet Daas et al, Discussion paper (09042), Statistics Netherlands. More details are provided in Annex F3.

⁷⁵ Data quality: A statistical perspective. Karr, A.F., Sanil, A.P., Banks, D.L. (2006), *Statistical Methodology*, 3, pp. 137-173.

The *source hyper dimension* measures the extent to which information contained in a data source is exploited. The key associated quality dimensions concern the frequency of delivery (yearly, monthly, continuously), the relevance with respect to the information needed, the extent to which the information actually meets SBR needs, the relationship with the administrative authorities, and the procedures taken to mitigate the dependency risk on the source.

The *metadata hyper dimension* focuses on the conceptual and process related quality components of the source metadata. Prior to use, it is essential that the SBR fully understands the metadata and their quality because any misunderstanding is likely to have a serious impact on quality of output based on data from the source. Metadata include the administrative regulations and the clarity with which changes in legal environment are described. The regulations determine the administrative units, the definitions of the data they are to provide, the reference time periods, and the forms used in data acquisition. Comparability is adversely impacted when administrative data for specific reference periods cannot be readily transformed to the reference periods/points required for SBR purposes (for example weekly data, or averages having to be transformed to point in time values). The identification numbering system is also important. If different systems are used by different sources, combining data from the sources is more difficult and error prone.

The *data hyper dimension* focuses on quality indicators that can describe, in a quantitative or qualitative manner, the quality of data that are input to the SBR. They refer to technical checks, accuracy, completeness, timeliness and integrability.

Source and metadata quality assessment is usually done by assigning scores to different key dimensions according to the sources used and related metadata.

10.5 Frame Errors and Their Implications for Surveys⁷⁶

10.5.1 Types of Frame Errors

As noted in Section 7.6, an error may be defined as “*a difference in the information presented in the register and the information as it should be, according to a chosen image of the real world produced and maintained by an accepted instrument and documented procedures*”.

The following types of errors may be identified.

Errors in existence

This type of error is due to false information regarding the demographic characteristics (date of creation and date of cessation) of a particular unit. There are two categories of existence error.

- Units are recorded as economically active, but are not yet active, or no longer active, in the real world. This results in *over-coverage*, which may lead to response problems for statistical surveys based on the SBR.
- Units are economically active but are not present in the SBR. This type of error results in *under-coverage*, which will lead to under-estimation in survey outputs.

⁷⁶ Eurostat – Business Register Recommendations Manual – Chapter 18 – The treatment of errors

Errors in identification characteristics

Errors in names, addresses, telephone numbers, etc., can hamper data collection due to problems locating and contacting the units. Errors in names and addresses also impede the use of the SBR as a tool to link and coordinate data from different sources.

Errors in stratification characteristics

This error type includes errors in legal form, economic activity code, size class (based on number of persons employed, turnover and/or net assets) and geographic location. These errors may affect the inclusion of units in survey samples and certain strata, and in SBR statistics. They may result in inefficient sampling and sample allocation for surveys based on the SBR, and in inaccurate population estimates derived from the SBR.

10.5.2 Impact of Errors on Survey Frames

SBR errors can have a big impact on the processes and results of surveys based on the SBR.

Non-response and coverage problems may result in very significant non-sampling errors. These two problems are inter-related and can be attributable to SBR errors. For example it may be impossible to contact a unit included in the frame because the contact characteristics are wrong. Units wrongly included in the frame may be selected for the sample and then found to be out of scope, thus reducing the effective sample size. The impact of frame errors on survey estimates is a measure of the accuracy of the SBR.

Frame errors and their impact on overall survey error can be classified according to the following types.

- *Under-coverage.* The frame derived from the SBR does not include all units within scope for the survey. Reasons for under-coverage are: omissions (lags and leakage), errors in activity status (falsely inactive units), and errors in stratification characteristics (units miscoded so as to be out of scope). SBR under-coverage results in under-estimation.
- *Over-coverage.* The frame derived from the SBR includes units that are out of scope. Reasons for over-coverage are: duplication, errors in activity status (falsely active units), and errors in stratification characteristics (units miscoded so as to be in scope). If an out of scope unit is not identified as being out of scope, the result is over-estimation. If it is identified, its exclusion results in a reduction in effective sample size and hence an increase in the sampling error.
- *Errors in unit characteristics.* Errors in stratification characteristics such principal economic activity code, size code, and location cause inefficient sampling and sample allocation. Errors in contact data result in increased non-response and non-response follow-up.

10.6 Classification of Metadata

SBR metadata can be classified using the following schema.

Source of data

A source code generally consists of an alphanumeric code that identifies exactly the source used in assigning the value of a characteristic of a unit. For example, it may indicate that the turnover value of a particular enterprise has been taken from VAT data, or that the number of employees' value has been obtained from a particular survey. Every value is assigned a source code.

Procedures used for attribution of unit characteristic values

These metadata include the procedures used in resolving conflicting information from different sources;

Production process

These metadata include descriptions of the processes for:

- acquisition and processing of administrative and statistical data;
- SBR profiling and improvement surveys;
- production of frozen frames and survey frames;
- production of statistics; and
- integration of administrative registers.

Updating and error correction process

These metadata include descriptions of updating processes and determining whether changes occurring in a period reflect real world economic changes or corrections.

Reliability of data

These metadata include quality indicators and references to the production sources and processes.

Characteristic value updating history

These metadata include the date/period to which the value of a characteristic relates and the date on which that the value was last updated in the live register.

Other documentation about sources and methods

These metadata may help users to assess the quality of SBR data. Documentation on sources and methods may be disseminated with SBR statistics or be made available through a database (web application) that can be consulted by users.

10.7 Quality Assessment Methods

10.7.1 Introductory Remark

The following paragraphs outline the methods that can be used to assess SBR quality. Some are generic in the sense that they are applicable to all statistical processes and products; others are specific to the SBR.

10.7.2 User Survey

Given that quality is defined in terms of user needs, collecting users' views of the SBR is a first step in SBR quality assessment. The aim is to obtain the views of major groups of users regarding each of the quality dimensions. The views of survey staff that use SBR data are particularly important as they are the primary users. Often the assessment results differ from one type of survey to another one. For example, managers of structural surveys may have a different view of SBR relevance to those responsible for short term statistics. Complete satisfaction of all users is not a realistic goal. In interpreting the results, the relative importance of each of the types of users has to be taken into account.

10.7.3 SBR Audit

A quality audit is a systematic, independent and documented process for obtaining quality audit evidence (records, statements of fact or other information that are relevant to the quality audit criteria and verifiable) and evaluating it objectively to determine the extent to which the quality audit criteria (set of policies, procedures or requirements) are fulfilled.⁷⁷

Auditing is a “powerful tool.....by providing important information⁷⁸ to improve the quality of the SBR.

From time to time, an overall audit of the SBR processes and outputs should be undertaken. Since use of external audits can be expensive, it may be desirable to conduct an internal audit. Some NSIs have a unit within the organisation whose role is auditing, and in this case the unit should undertake the audit. In NSIs with no such organisational unit an ad hoc team of auditors should be assembled including both SBR users and other statisticians with limited (or no) exposure to the SBR.

10.7.4 SBR Improvement Surveys and Quality Measurement Surveys

The primary purpose of SBR improvement surveys is usually to improve the quality of unit characteristics, by detecting and correcting errors and by filling in missing values. However, such surveys are also a means of measuring SBR accuracy. Indeed, some such surveys may be conducted with the primary goal not of correction but of measurement of the quality of characteristics such as activity status, location, economic activity code and size.

Such surveys must be designed carefully because of the difficulty in determining the “real” values for comparison with the SBR values. Furthermore, to avoid self-referential findings, the survey results should preferably be analysed by statisticians from outside the SBR.

10.7.5 Auditing Clerical Work

Quality audits are a useful tool for monitoring the quality of clerical processing and of automatic updating. They may involve checks on representative samples of clerical update actions (described immediately below) or regular analyses of key characteristics at aggregate level (as described in Section 10.7.6), or, preferably, a combination of the two approaches.

The aim of auditing clerical processing is to monitor the quality of the clerical input to a SBR. It involves checking samples of clerical updates. The checks should be based on regular, random and representative samples.

Clerical audits are normally undertaken by experienced staff, who investigate the work of SBR staff to see if the actions that have been taken comply with the processing guidelines. They may also be done assigning the same update actions to different clerks and then comparing the resulting updates.

Clerical error rates can be monitored over time and reduced, for example by improving procedures and by informing and training staff. Regular summary reports should be produced to inform managers and users.

The quality auditing function should be closely linked to documentation and training functions, perhaps in a form of quality circle, so that issues are identified, resolved,

⁷⁷ United Nations Statistics Division: National Quality Assurance Frameworks

⁷⁸ Eurostat – Data quality assessment and tools

documented and covered in future staff training. By this means it should be possible to ensure sustained improvement in quality over time.

10.7.6 Macro-Editing

Another class of quality audits involves monitoring changes in the SBR at aggregate level. This should be done on a regular basis and linked to the production and dissemination of the frozen frames.

Frequency distributions for key characteristics are compared before and after periods of updating (automatic and/or manual) to assess the impact of the changes on various subsets of units and to ensure that all changes can be adequately explained. In performing the analyses priority is given to large units, and to units that are particularly relevant for specific survey or under particular observation from users.

10.7.7 Defining and Monitoring Quality Indicators

A system of quality indicators in BR is a very important means of assessing quality. The construction and use of quality indicators is discussed in the following section.

10.8 SBR Quality Indicators

10.8.1 Introductory Remarks

The starting point in defining a set of quality indicators for the SBR is an appropriate conceptual framework. This is discussed in Section 10.8.2. As described in the following three sub-sections, SBR quality indicators may be divided into three groups corresponding to processing phases, dealing with:

- quality of *input*;
- quality of *processing* (or *throughput*); and
- quality of *output*.

10.8.2 Conceptual Framework for Quality Indicators

Key factors

Key factors in defining indicators of SBR data quality are the following.

Time. The SBR is constantly evolving, thus quality indicators are characterised by a reference date (t) or period (t-x, t). In presence of a lag between the reference period (t-x) and the assessment period (t), the indicator describes quality at (t-x) measured at (t).

Scope. A quality indicator is applicable to a particular type of unit (say enterprise or local unit) and, within that type, to a subset of all possible units. The subsetting may be formally defined by a filter presented in the form of a logic formula operating on the relevant set of units and their characteristics, for example, *active enterprises born before (t-x)*.

Sub-population. A quality indicator must be defined and measured at the level of the sub-populations of interest. Having only global indicators may mask weaknesses in specific sub-populations. It is common to define sub-populations in terms of size (small/medium/big), region, and economic sector.

Characteristic. A quality indicator applies to a particular characteristic of a set of units.

Criteria. To construct a quality indicator it is necessary to have criteria for assessing the quality of the characteristic, unit by unit. For each unit, and for each characteristic, it must be possible to assess whether the value is right/wrong (true/false), or where it lies on a scale of quality between 0 and 1.

Quality assessment mechanisms and criteria

There are various mechanisms and criteria by which to assess quality as follows.

Use of external information source. A value of a unit in the SBR can be considered as *correct* if it is *sufficiently close* to a reference value from an external source. This is the most commonly used criterion. It focuses on compliance (i.e., whether the value of a characteristic in the SBR complies with the value of the same characteristic in an external source). It is a proxy for accuracy (whether the value of a characteristic is correct or not) but is the only possible measure when the real value is not known. The *compliance rate* (% of units for which the characteristic assumes a sufficiently close value) is a practical substitute for *reliability rate* (% of units for which the characteristic is correct) when the latter cannot actually be measured.

SBR improvement survey (quality survey). Quality is assessed by comparison with a reference value obtained by an SBR survey. This approach is expensive.

Internal consistency. A value is deemed *correct* if it is consistent with the other characteristics of the same unit (for example, turnover/employees, main activity/legal status). Definition of consistency edits is difficult. Often they are based on *plausibility criteria*, meaning most true values satisfy these criteria. Even if a characteristic value passes the edit there is no guarantee that it is correct.

Temporal consistency. The quality of a characteristic is defined on the basis of a comparison with its values in previous time periods. The aim is to identify impossible or implausible changes from one period to another.

Quality without witness. It is possible to make an assessment of quality without having a specific reference value or element of comparison. Examples of the basis for such an assessment are:

- the date on which the value was most recently checked or updated - with the underlying principle that the more recent, the better;
- the name of the information source – the likely accuracy of data from a source may be known;
- the methodology adopted – the appropriateness of the methodology may be known.

10.8.3 Quality of SBR Inputs

The SBR typically has many input sources, mostly administrative. As the qualities of these sources affect SBR output quality they are very useful indicators in themselves. From the SBR perspective it is impossible to control ex-ante the quality of each administrative source. Rather, the quality of the source can be assessed only ex-post by means of suitable analyses to identify errors in data supplied and take these into account.

Quality assessment includes consideration of the accessibility and clarity of administrative data, meaning the ease with which the SBR updating process can access and interpret the data. Good metadata management requires that data format and content are precisely reported and data are accompanied by metadata, examples and advice.

Other simple indicators of the quality of the source are:

- *Time lag*: difference between the date on which the data are supplied and the reference period to which they refer;
- *Completeness of characteristic*: number of missing values; and number of missing values as a proportion of the total number of values.

Many problems that arise using administrative sources are caused by changes in the source that are not known to the SBR. Such changes may include changes in registration or cancellation rules, in classification criteria, or in the administrative control processes. In order to detect such changes, *big changes* - meaning changes that are impossible or implausible during a given period (for example a quarter or a year) – should be identified and reviewed to determine the underlying causes. The main objective is to check the *stability* of the sources and avoid the situation whereby changes that are merely administrative produce structural changes in the SBR that do not reflect economic reality.

Simple indicators based on the comparison of the values provided in two different years may be formulated along the following lines.

Percentage/proportion of variations (per characteristic)

For example a simple indicator is the percentage of units that have changed principal economic activity from year t to year $t+1$. This can be computed for the whole population or for a specific subset. Further analytical investigation can be initiated if the value in the reference year exceeds the average level of the indicator over time.

Comparison of data from a single source over time is also fundamental in analysing the completeness of enterprise births and deaths. Such analysis is essential in helping to detect and minimise under-coverage and over-coverage. Simple counts of the numbers of births or deaths during a reference period do not of themselves provide much information. Viewed as a time series, more complex and useful indicators may be defined.

Indicator for over/under coverage:

The effect of update delay associated with a source can be obtained by comparing, where available, two values of a characteristic at different points in time, for example:

- the number of cessations, $N_{cess}(T)$, occurring in year t as reported in data received during a period ending at time T ; and
- the number of cessations, $N_{cess}(T+1)$, occurring in year t as reported in data received between time T and $T+1$.

In this way the lag in the registration of the cessation dates in the input source can be estimated using a simple indicator such as $1 - N_{cess}(T+1)/N_{cess}(T)$. Analysis of this indicator can give an idea how long it is worth waiting to receive cessations before using the data to generate survey frames. The shorter the time the more cessations will be not have been received. The longer the time the fewer the cessations that will not have been received but the later the frames generated relative to the survey reference year.

10.8.4 Quality of SBR Processes

For the purpose of defining quality indicators, SBR processes can be divided into three phases, namely, integration of input data, assignment of values of characteristics, and editing, as described in the following paragraphs.

Phase 1: Integration of input data from administrative sources

The purpose of this phase is to integrate administrative data from various sources and to create *clusters* referring to the same unit (the enterprise). Two different sub-phases can be distinguished and quality indicators defined for each.

First sub-phase: analyse within a source. For each input source, records that pertain to the same legal unit (as identified, for example, by a common taxation identification number) are integrated. Possible quality indicators are:

- *number of duplicates as a proportion of the total number of supplied records.* A decrease over time indicates an increase of quality of the source from the perspective of the SBR.
- *number of new records during a given reference period.* This provides a measure of coverage in terms of unit creations. Comparisons over time and with a benchmark can be used to identify possible problems in supply.

Second sub-phase: link between sources. Records coming from different sources that pertain to the same legal unit (again, for example, as identified by a common taxation identification number) are integrated in order to build-up a *cluster of records* for the same enterprise. The taxation register is, typically, the base used to define the set of legal units and to integrate all the other sources. This phase has the particular aim of identifying the set of administrative data records relating to each individual legal unit. Errors in this phase, such as missing or wrong links, can greatly affect the data produced in the following phases. Possible quality indicators are:

- (a) number of clusters of records in period (t) linked to legal units;
- (b) number of clusters of records in period (t) not linked to legal units;
- (c) number of clusters of records in year (t-1) not linked to legal units;
- (d) under-coverage indicator: $[(a) \cap (c)]/(a)$.

The indicator (d) measures the under-coverage due to time lag in the registration of units in the taxation register. It represents the percentage of units already provided in (t-1) by other administrative registers, not linked in time (t-1) and then linked in the succeeding time (t). They are units that could have been linked and would have been included in the SBR in time (t-1) if the base source (typically the taxation register) had successfully contained such units at that time.

Phase 2: Assignment of values of characteristics

The second phase of the SBR production process comprises the assignment of the values of the characteristics of each unit and the identification of active units in year t. The quality of the assignment procedure for each characteristic can be evaluated using indicators based on outputs produced at each step in implementation of the procedure.

The characteristic of a unit of which to check the quality first is its *activity status*. This is because the frozen frames and subsequent survey frames contain only active units (more precisely, active, recently active and potentially active units). Thus, other characteristics need only be checked for these units. Quality indicators for each source may be defined along the following lines.

- *Active status concordance rate* = number of units for which unit has active status in both source and SBR as proportion of number of units in common to source and SBR.
- *Inactive status concordance rate* = number of units for which unit has inactive status in both source and SBR as proportion of number of units in common to source and SBR.

- *Activity status discordance rate* = number of units for which unit activity status differs between source and SBR as proportion of number of units in common to source and SBR.

Sources that may be available for determining or checking activity status are:

- small and medium size enterprise survey – checks for active and inactive units;
- foreign trade survey – checks for active units;
- bankruptcy database – checks for inactive units

Phase 3: Editing procedures

The third phase of the SBR production process comprises implementation of the editing and imputation rules used in final identification of the frozen frame for a particular reference period t . In order to measure the quality of the rules, quality indicators can be defined to measure the errors produced by each rule.

The quality checking plan usually consists of a set of separate quality *project modules*, usually executed sequentially. The project modules may be changed in number and composition according to their results over time. Each project module comprises a set of rules having a similar structure and affecting data in a similar way. The main project modules are:

- *cleaning* – using rules determining the exclusion of some units from further checks;
- *deterministic* – using if/then clauses that cause the automatic changes in the values of the relevant characteristics whenever certain conditions occur; and
- *errors* – using rules that generate error warnings whenever certain conditions occur.

Since the number of edits can be very high, deterministic rules usually focus on peculiar subsets of units and on specific characteristics (for example, on the cross combination of economic activity code and size). The editing process produces error warnings for possible clerical follow-up. There is a need to limit the number of warnings in accordance with the resources available so that SBR staff can concentrate their efforts on economically significant units rather than smaller ones. This is sometimes referred to as *significance editing*.

The calculation of quality indicators based simply on the number of warnings or errors in a given time period may be misleading. More useful indicators can be built by looking at trends, i.e., measuring the increase or reduction in the number of units that fail each type of edit over time.

A synthesis of quality indicator values can be achieved along the following lines:

- variation (in absolute terms and percentages) between t and $t+1$ of counts of units by type of error;
- variation (in absolute terms and percentages) between t and $t+1$ of counts of units for which values of characteristics have been automatically changed by deterministic rules;
- variation (in absolute terms and percentages) between t and $t+1$ of counts of units for which a warning is generated and/or that have been manually verified.

10.9 Quality Policy and Improvement

10.9.1 Quality Policy

Elements of a quality policy can be (1) the decision to measure the quality of SBR, (2) the decision to communicate information about the quality of SBR to all users, and, based on the

measurements and the feedback from users, (3) the decision to initiate a program of improving the quality of the SBR.

As discussed above, the measurement of quality is aligned with three principal groups of activities namely the *input*, *processing* and *output* phases. For each of these phases relevant quality dimensions and indicators are identified and defined, based on the specificity of the sources, infrastructure and resources available.

A by-product of measuring SBR quality is the generation of metadata and paradata that allow a better knowledge of the state, content, structure and processes of the SBR. Detailed examples from Italy, Columbia, and the Netherlands are provided in Annex F.

An example of communicating quality policy is an *SBR quality declaration* aimed at all users, internal and external, presenting the quality dimensions and associated indicators. Another example is informing users, in as timely a manner as possible, about significant events related to SBR maintenance, such as the availability of new administrative data, the next frozen frame, changes in important units, and changes in classification.

The decision to implement an SBR quality programme must be based on a careful analysis of costs and benefits. Benefits should be seen primarily in terms of results of improved quality from a perspective of users, in particular statisticians conducting economic surveys and the national accountants that use the resulting outputs. In addition, the SBR quality programme should be viewed within the broader envelope of a NSI wide program to improve economic statistics rather than being seen as an end in itself.

10.9.2 General Approach to Improving Quality

Quality improvement is an iterative process based on (1) construction of a set of *quality and performance indicators* for SBR inputs, processes and outputs, (2) setting *quality and performance targets*, and (3) defining *quality assessment tools*. These indicators, targets and tools enable monitoring of SBR quality as the basis for formulating, reviewing and implementing quality improvements. The results of monitoring quality should be summarised in quality reports that are made available to SBR staff and to users.

The following subsections contain examples of ways in which the quality of the SBR can be improved. The examples often concern more than one quality dimension as improvements cannot be made entirely independently in each quality dimension. There are trade-offs that need to be considered.

10.9.3 Improving Timeliness

In order to have the most representative picture possible of the population of enterprises, the SBR updating process should be as timely as possible whilst at the same time providing the as good coverage as possible. As discussed in Chapters 6 and 7, this is achieved by systematically applying updates available from relevant administrative sources, for example data from corporation registration systems, income tax, VAT and social security systems, chambers of commerce and other trade associations. The aim is to detect creation, structural change and cessation of units as quickly as possible. Administrative data such as sales tax remittances, income tax returns, and payroll deductions provide clear signals that an enterprise is active. Signals indicating when a business becomes inactive are less numerous and timely.

Another aspect of improving timeliness is reducing the time required to apply updates so that these changes can be quickly reflected in the survey frames. Often survey managers complain that changes they have detected via responses to survey questionnaire are not immediately

used to update the SBR because BR staff spend time in processing the changes and deciding on the appropriate updates. One way of reducing processing time is to permit the survey staff themselves (after having received the appropriate training) to perform the updates. Another possibility is to let the businesses themselves to update their own information directly via an appropriate external portal.

10.9.4 Improving Completeness

There should be constant investigation of new ways to extend the SBR by collecting information about enterprises from additional sources. For example, it may be possible to link the SBR to registers like an employment database or industry specific database to obtain additional information on employment size or economic activity.

10.9.5 Improving Coverage

Different strategies must be set up for reducing under-coverage and over-coverage, especially duplication of units. The risk of duplication depends almost entirely on the particular input sources and the matching procedures used in bringing together and unduplicating data from these sources.

- When matching is based on a common identification system there is a risk of misinterpreting the identification codes, for example a fiscal code may be confused with the VAT code.
- In absence of a common identification, matching depends upon record linkage techniques, for which there are numerous high performance software applications but still a significant risk of missed matches and mismatches.

Automated procedures must be adapted to the peculiarities of the situation in any given country.

10.9.6 Improving Quality Reports

Reporting on quality is a crucial aspect of a quality programme. The general aim should be to deliver short sub-annual (preferably monthly or quarterly) quality reports to alert users to significant, recent changes in the SBR - changes that affect the frozen frame and hence survey frames and ultimately published economic statistics. Within the report it should be possible to drill down by industry, region, legal status, and administrative source.

Sub-annual reports should be accompanied by more comprehensive annual quality reports that track smaller changes by month and that show trends over years. They may also report on the results of monitoring particular cases, for example holding companies, foreign controlled transport companies, and discrepancies between employment/turnover for big construction enterprise groups operating abroad.

10.9.7 Providing Survey Support

To assist surveys and to monitor the quality of the SBR as seen from a survey perspective, it is vital to provide functionality that:

- allows a survey to monitor its population, since every survey needs to understand the significant changes that have occurred between survey cycles and wants to track births, deaths, and changes to large units;

- provides a frame on the basis of which efficient sampling schemes can be designed and panels monitored, and the results from sample surveys can be grossed up to population estimates;
- provides the information for assembling mailing lists, dispatch of questionnaires, monitoring responses, and contacting units in the event of non-response.

Such functionality enables SBR and survey staff to monitor their operations and to identify areas or weakness and potential quality improvements.

10.10 Examples of Quality Indicators by Quality Dimension

The particular choice of quality indicators for an SBR depends on an NSI's situation in terms of survey programme, statistical infrastructure and the available resources. Examples of possible quality indicators by quality dimension follow. A considerably more comprehensive set of indicators developed by the Italian NSI (Istat) for its SBR is detailed in Annex F1.

Input Phases

Dimension	Indicators
Timeliness	Time lag for each administrative source (the time lag between registration/updating of data with the administrative source and the delivery of this administrative data to the SBR).
Accuracy	Coverage of each source (measuring under- and over- coverage). Completeness (of characteristics) for each source.

Processing Phases

Dimension	Indicators
Punctuality	Time lag (whether survey frame available according to agreed schedule).
Accuracy	Coverage and completeness (measuring the extent to which administrative data are integrated into the SBR). Extent of missed matches and mismatches in linking administrative units and statistical units.
Accessibility	How readily available relevant data are to SBR staff.
Coherence	Internal coherence of units and characteristics within the SBR Coherence with other registers.

Output Phases

Dimension	Indicators
Timeliness	Time lag for each survey (time lag in delivery of the survey frame relative to the survey reference period). The currency of the values of the characteristics.
Accuracy	Coverage (completeness of the survey frame for given reference period).

	Completeness (of values of characteristics). Error rates (for example, percentage of units with wrong address, percentage of units with wrong activity code). Non response rates (as indicator of errors in contact information).
Accessibility	How readily available relevant data are to the users
Coherence	Coherence with other registers

11 Key Considerations in Establishing an SBR

11.1 Introduction

This chapter provides guidance on the planning, organizational, legal and technical (IT systems) factors that position an SBR for success. It is intended to offer practical suggestions without being overly prescriptive. The environmental circumstances and factors within which countries build their SBRs can be vastly different. The legal frameworks for acquiring data and the access to human, financial and technical resources ultimately greatly influence how an NSI can proceed.

Some key themes recur, including the following.

- Effective partnerships must be built with data suppliers, funding providers and SBR users, first, by ensuring recognition of the critical role of the SBR in delivering a coherent and reliable national economic statistics program, and, second, by setting up robust governance structures and partner engagement mechanisms.
- Implementing and operating the SBR must be managed in a manner that enables it to focus on and achieve its primary purpose, which is to identify the population of economically active units in a country so that it can be surveyed to acquire useful economic data. An SBR can fulfil other secondary yet highly desirable roles, such as acting as a data collection management and tracking tool, as described in Chapters 2 and 8. The SBR design and implementation plan should allow for adding the corresponding components, but only once the SBR has fully matured as a source of quality statistical frames.
- The approach should be to maintain simplicity to the extent possible. Conceptual and technical complexities should be added only when they serve a practical purpose; they should never divert the SBR from meeting its larger goals.

The purpose of Chapter 11 is to provide guidance to SBR developers and managers. Section 11.2 is about the key considerations in establishing an SBR. It recommends a modular approach to the development of the SBR. Section 11.3 is about governance and organizational structure, legislative framework, funding, human resources and relations with other registers. Section 11.4 offers recommendations on the IT infrastructure and programming requirements and, finally, Section 11.5 is about data retention or *safekeeping* of the SBR historical information in order to satisfy operational and analytical purposes.

The primary reference for the chapter is the AfDB's *Guidelines for Building Statistical Business Registers in Africa (2014): Chapters 13-14, 16-17*.

11.2 Planning Considerations

11.2.1 Initial versus Longer Term Scoping: the Importance of a Modular Approach

An SBR's primary function is as a central frame for economic surveys. This enables conceptual coherence and creates the foundation for an integrated economic statistics program. This fundamental purpose should be the focus at the outset of SBR development.

The longer term vision, however, should, from the beginning, also allow for addition of other features and components that will further enhance the SBR's value-added. The secondary outputs to be potentially developed after the SBR has become operational as a survey frame are described in Chapters 2 and 8 and are as follows.

- A *survey feedback mechanism* that facilitates the update of frame information based on the information pertaining to frame-based characteristics, such as economic activity classification.
- A *receptacle for tracking survey collection outcomes, response rates and other metrics*.
- A *module to track respondents and response burden*. A *source of register-based statistics*. (Note: It may be advisable for statistics derived from SBR data to be based on a satellite approach, because then the ‘base population’ is coordinated as noted in section 2.3.5).
- A source for national and *international data exchange*. This requires the usage of common concepts and methods to ensure the coherence in the units and statistics produced by the NSIs involved. This is important as SBR data are used in studies comparing (the structure of) different national economies and other international studies.

In particular the IT professionals who design the SBR data structures and the overall system benefit from having a longer term vision clearly defined in as much detail as possible. This simplifies addition of modules as the SBR evolves. Defining all the different modules right at the beginning enables their development to be prioritized and their implementation to be phased.

An important development task is to determine the characteristics that need to be included in the survey frames, and hence the characteristics that have to be reliably populated and maintained with data from either administrative or statistical sources.

11.2.2 Key Considerations for Establishing a Survey Frame

To reiterate, the SBR must first and foremost provide a reliable listing of enterprises (in the form of a frozen frame) on the basis of which surveys can accurately measure the economic trends of a country. Creating the SBR is challenging, as is keeping it up to date once it is in use. The challenges of creation and maintenance are greatly eased by adhering to the principles outlined below.

Do not over-extend resources in the early stages by trying to cover all types of enterprise

While a highly developed SBR may cover a vast portion of the economic population, a new SBR must focus on covering the part of the population that is most important economically and that can be most reliably captured and reflected. The need to manage with limited human and technological resources, and to use initial funding efficiently, should limit the scope of the initial SBR population.

Reflecting the informal economy, which is highly diversified and for which no administrative data exist, cannot be a focus of the SBR development project. However, in certain countries—particularly developing and emerging economies—the informal economy is important, and collecting economic data about it is a priority. Thus, as was discussed in Chapter 3, although the informal sector is not the main focus in setting up the SBR, in such countries it is vital to consider to what extent, if at all, the SBR will cover the informal sector and how the economic production of those enterprises not covered by the SBR can be measured or estimated.

It is essential to maximise the coverage in every sector of the economy to obtain reliable macroeconomic indicators. Since certain sectors are composed of numerous small enterprises, maximising coverage in these sectors implies the addition of a large volume of records to the SBR that result in very small contribution to overall macro-economic indicators (such as GDP) but significant contributions in certain sectors.

Including these small businesses is also useful as the data they provide support policy analysis pertaining to business formation strategies, small business financing and other micro-economic issues, and greatly assist in the economic analysis of entrepreneurship and the SME sector. Thus, a progressive approach is recommended whereby a country first focuses on developing an SBR system designed to maintain large units that are economically most relevant to its GDP, and subsequently aims to develop automated processes to include the smaller units. Completing the development of the SBR in terms of coverage is one of today's common concerns in developed countries. The initial development plan for the SBR system should incorporate the flexibility to accommodate a substantial increase in coverage due to the later inclusion of a (potentially very large) number of small units.

In determining appropriate coverage, the environment within which the NSI operates should be considered as well as the data that can be reliably acquired from administrative sources and the resources that are required to efficiently obtain, process, and load the administrative records.

Example from Statistics Canada of increasing SBR coverage

Statistics Canada has gradually expanded its SBR population coverage. Over the past five years, it has added businesses that are essentially constituted by self-employed individuals who pay income tax on their business income as persons not as corporations.

Plan for a system that provides both live and frozen versions of the register

As initially discussed in Chapter 2, the SBR environment should include both the live register and the frozen frames:

- the live register receives updates that are instantly recorded and used for contacting the businesses; and
- the sequence of frozen frames is produced from the live register monthly, quarterly or annually and is the source from which survey programs draw their sample files.

A frozen frame contains all the salient characteristics that are stored in the SBR for a particular unit, including its unique identifier and its size, geographic, sector and economic activity characteristics. Survey programs can use a frozen frame to identify and stratify their sub-population of interest and draw samples.

The set of frozen frames also provides a basis for period-to-period comparisons of frame quality. Point-in-time estimates can be calculated and compared with one another to examine the number of units being birthed, died, re-classified, etc. This greatly facilitates the identification of anomalies and problem records. Frozen frames are also used in research and analysis projects, for example, business demographics and longitudinal studies such as entrepreneurship analysis.

11.3 Governance and Organizational Considerations

11.3.1 Introduction

Organizational aspects of an SBR, such as relationships with administrative data providers (for example, the tax authority), stakeholders (for example, the central bank) and SBR users (especially survey statisticians) are discussed in this section.

In almost all NSIs, the SBR resides in an economic division/department or a statistical infrastructure division/department within the NSI.

Example from Statistics Canada

In Canada, the SBR is a critical component of the economic statistics programs of the national statistical agency, Statistics Canada. The SBR is therefore held, managed and maintained within the agency. Statistics Canada is within the Minister of Industry's portfolio. It also has close ties to the Department of Finance, other federal departments, and provincial, territorial and local government organizations to ensure that the economic statistics are relevant.

Strategic direction of Statistics Canada's program' is governed by various committees and working groups across the different levels of government. The program's governance model provides clear direction, enables periodic review of results and enables identification and execution of adjustments to achieve expected outcomes. The governance framework also enables transparent, effective and efficient decision-making, and supports accountability and continuous improvement of the program.

The governance and organizational structure of the SBR within NSI are important — both in developing the SBR and, even more, in maintaining it and providing support for users. The SBR should be, if possible, a separate organisational unit with a dedicated manager within the NSI. The unit should assume the following responsibilities:

- define and document all concepts, in line with international, national and local statistical standards;
- plan and direct the development of SBR system processes and functionalities;
- plan and implement a quality assurance program for the SBR with the goals of :
 - assessing its quality and ensuring its continued integrity;
 - defining and producing SBR quality indicators;
- identify system improvements and recommending adjustments to the training program or procedures where required;
- profile businesses to delineate those that are larger and more complex, and thus to properly represent their production output;
- ensure that businesses are classified according to the standard economic activity classification;
- create statistical units, using data from administrative sources, profiling activities and other surveys, to create a complete and unduplicated SBR aligned with the needs of the system of national accounts (SNA) and other users;
- validate new development strategies, specifications and procedures;
- develop and deliver courses and material to educate the full range of SBR users and stakeholders, including profilers, frame specialists, analysts, coders, survey areas and collection areas;
- develop a certification process so that those wishing to access the SBR must first achieve an appropriate level;
- develop an online system for all internal users, enabling them to browse and update specific enterprise structures and specific units;

- support those who use SBR data, which includes evaluating their needs as related to surveys or analysis;
- provide direction and support on legal aspects related to SBR data, such as access and dissemination; and
- maintain a dedicated group tasked with producing data for users and processing all data related to SBR maintenance.

Relationships with users

To understand and gather users' needs the SBR should have a consultation mechanism. The most important users are the survey areas and ultimately, the SNA. The SBR staff should meet survey statisticians regularly to understand the changes in the economic world and corresponding requirements for SBR data. Every request from a user should be reviewed and prioritized based on its potential impact (benefits) and the availability of resources.

As the SBR is a single source serving multiple users, balancing requests within the limits of the resources available is challenging. The SBR manager must always be aware of the role of the SBR within the context of the larger statistical program.

11.3.2 Legislative Framework

As stated earlier, access to administrative records, such as corporate tax returns, business registrations, payroll deductions and value added tax remittances, is fundamental in building a centralized SBR. Many countries legally require the provision of such administrative data to the NSI for the purpose of compiling official statistics, and this authority, where it exists, must be fully leveraged to build an SBR.

In implementing the legislation, or even in the absence of it, detailed agreements, usually formal *memoranda of understanding (MOU)* or *service level agreement (SLA)*, should be signed with the administrative agencies as discussed in Section 6.2.1. MOUs are crucial because they establish a framework of general rules and procedures for interdepartmental data exchanges, and they specify data protection measures. More specifically, an MOU should clearly spell out the terms for acquiring the administrative data, it should stipulate the types of data required, the likely treatment of the data and the pre-established acquisition schedule, and it should also specify the data security measures and the means of transmission

Example: Portion of the Statistics Act allowing Statistics Canada to access Administrative data:

Statistics Act (Section 13): A person having the custody or charge of any documents or records that are maintained in any department or in any municipal office, corporation, business or organization, from which information sought in respect of the objects of this Act can be obtained or that would aid in the completion or correction of that information, shall grant access thereto for those purposes to a person authorized by the Chief Statistician to obtain that information or aid in the completion or correction of that information.

to be used.

11.3.3 Funding and Development Phases

The funding model for the SBR is itself an important aspect of its governance and decision-making process, and should be carefully considered at the outset.

The costs of creating and maintaining the SBR are not likely to be recovered on a fee-for-service basis as the SBR is a *public-good* type of resource used by multiple programs for diverse purposes. Some individualized activities, such as preparation of special data requests for individual external users, may yield revenues to offset the incremental costs of serving those users. However, developing and maintaining the SBR are generalized expenses, the majority of which are likely funded from the public purse, e.g., as part of the NSI's budget.

The key questions are thus: who should make the budget requests? and who should control the budgeted funds once they are received?

The precise answers depend on the organizational governance structures in place, but the overall objectives should be the same, namely to ensure that funds are allocated (and guarantee an ongoing functional budget) so as to fulfil the larger objectives of the economic statistics program in an optimal fashion. The manager of the SBR should be able to make specific budget decisions within the framework of this larger governance structure.

Developing and implementing an SBR consists of three distinct phases from a funding perspective: *pre-build*, *development*, and *post-build operations*. It is important that the last mentioned is included the development funding plans.

Conceptual development (pre-build) phase

This phase requires initial funding and resources for the following pre-build tasks:

- defining the purposes, uses and roles of the SBR, i.e., the data and services that should be produced to meet users' needs: to this end, consultations with users and stakeholders should occur from the start through all phases;
- outlining a clear overall picture of the future SBR;
- defining the initial scope of an SBR development project;
- defining the initial coverage of the population of enterprises;
- determining method(s) to derive the statistical units;
- analysing and determining all the data inputs needed to populate the live register and to create a coherent, accurate and timely set of frozen frames;
- obtaining access to the input data sources; this may require drafting MOUs between the NSI and other departments; and negotiating legal, ongoing data transfer agreements with organisations that will supply data, e.g., the authorities responsible for tax data;
- determining the availability of, and analysing, the necessary stratification characteristics;
- defining activation rules - what signal(s) or specific characteristics determine when an enterprise is to be considered active;
- defining the most viable means of receiving business inactivation and cessation information;
- identifying and determining the availability of administrative data from various sources; initial and ongoing cost is greatly influenced by the accessibility and usability of data input;
- determining if there will (or should be) be an economic census as a source of input which itself depends on the availability and accessibility of administrative data;

- outlining the processes that will assess the quality, validate and treat the data before they are loaded into the SBR;
- defining components for future developments, using the modular approach as discussed earlier in the chapter, like a module to track respondents and response burden.

Development phase

Development phase funding depends on the decisions made during the pre-build phase. A modular development approach is advantageous because development can be phased in and compartmentalized, module by module. For a fully operational SBR, the following modules should be funded and fully functional for the initial SBR:

- live register, covering targeted population;
- batch load and update processes;
- online tools supporting SBR maintenance;
- frozen frames, outputs for surveys; and
- quality assurance.

Post-build operations phase

The SBR should be seen as a continuously evolving entity within the NSI. A long-term vision supported by senior management is needed. A senior steering committee should oversee the needs and usage of the SBR. This ensures that the SBR will evolve in accordance with the agency's general data requirements and functionalities/roles. A ten year investment plan should be outlined as input to preparation of the funding schedule. In addition to the funds reserved to support the ongoing production and system maintenance of the SBR, this long term plan should be reviewed annually and adjusted to new business requirements and the current NSI budget reality.

11.3.4 Human Resources

The human resources allocated to the SBR depend on the financial support that the NSI is prepared to dedicate to developing and maintaining an SBR. Investing in a high quality SBR that is timely, accurate, coherent and user-friendly results in lower costs and/or better quality in other parts of the economic statistics program, for example by facilitating more precise sampling methodology, lowering collection costs and respondent burden, improving response rates, and providing higher quality estimates. To the fullest extent possible, the human resources allocated to the SBR should be *dedicated to the SBR*, i.e., not have any other function.

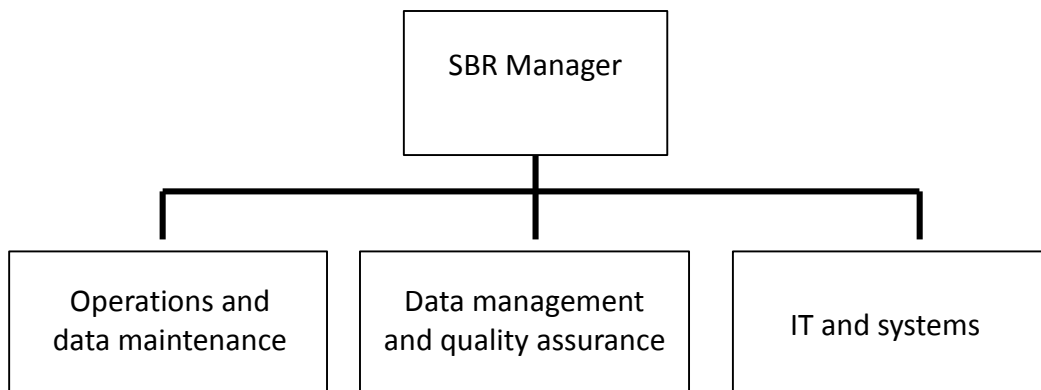
Regarding the organizational structure of the SBR, there is no international standard, or even commonly used practice, amongst NSIs. Therefore, each NSI has to determine its own particular SBR structure and the SBR's place within the economic statistics infrastructure and program as a whole. Factors to consider in this determination are:

- numbers of staff in the NSI and economic statistics program – evidently the SBR cannot expect to have more than its fair share of the human resources;
- number of surveys that the SBR is servicing, or will service;
- degree of centralization/regionalization of the NSI and scope/need for regional business registers; and
- sophistication of computer systems supporting SBR operations.

Ongoing operational requirements and organization

The operational structure of a completed SBR, as well as the corresponding funding structure, should be clearly set out in advance to prevent the build project from creating an ineffective or underutilized SBR. Although many alternative groupings of the functions are possible, Figure 11.1 and the following paragraphs present a feasible structure involving three SBR sections, i.e., organisational work units, with distinct roles: *operations and data maintenance*, *data management and quality assurance*, and *IT and systems*.

Figure 11.1: Operational structure of an SBR



Each section can be expanded, contracted, or modified as need be to take account of the particular circumstances of the NSI and the functions that the SBR performs. The sectional structuring has been devised to delineate clearly the essential tasks, responsibilities and funding commitments. Opportunities for efficiency and collaboration between the sections can be pursued.

Operations and data maintenance section

The staff of this section undertake all the basic tasks associated with operating an SBR, including processing administrative data, processing frame data feedback from surveys, performing profiling activities, producing frozen frames and training staff and users. Securing and committing full-time personnel to these tasks is essential to ensure that the SBR data remain relevant and accurate. The size of the section depends on the coverage of the SBR at its outset. Its structure can be expected to change as SBR coverage and content improve.

A realistic, well-funded and evolving maintenance strategy is a crucial aspect of the SBR. For example, in the case of profiling, the exact nature of the signals and processes that prompt a profiler to review a particular enterprise structure should be reassessed over time. When an SBR is new, the section might simply schedule regular profiling of the largest so many enterprises. Subsequently it might target profiling operations more specifically to those large enterprises for which data collection is proving most difficult.

The section should comprise:

- specialists in SBR concepts and business structures, to perform profiling;
- specialists in administrative data sources and processes;
- industry classifications specialists; and,
- operational staff.

The section should be responsible for:

- acquiring and processing administrative data;
- processing frame data feedback from surveys;
- performing profiling and classification activities;
- investigating and correcting errors and inconsistencies that are identified by users or the quality assurance section;
- producing frozen frames; and
- training SBR staff and users

Data management and quality assurance section

While the SBR is being established, and after it is established, the scope of the staff that are initially responsible for analysing the incoming data should be extended to assessing quality requirements and developing a quality assurance strategy. A person, or a team, ideally with a good understanding of SBR concepts and of enterprise structures and strong technical skills in validating data and data processes, should analyse frame data and investigate potential quality issues.

This *data management and quality assurance section* should be maintained as part of the SBR's ongoing operation. The team should comprise:

- data source specialists, to create specifications and/or use cases for the development of business rules and automated processes for acquiring and processing administrative and statistical data;
- specialists in SBR concepts and business structures (having experience in economic statistics and survey cycle is relevant);
- survey interface specialists; and
- data analysts.

This team should be responsible for:

- providing specifications of the characteristics to be populated in the database tables, maintaining the content of the tables and submitting new or modified specifications as needs arise and shift;
- perform acceptance testing of changes made to the SBR systems, interfaces and processes;
- assessing data from administrative sources;
- conducting internal analysis and coherence checks on frame information;
- reviewing requests from survey statisticians and other SBR users (including SBR staff themselves) for improvements or modifications to SBR metadata, characteristics and outputs, prioritizing these requests and creating specifications for future development and improvements;
- liaising with survey staff to ensure that they understand SBR processes and outputs, and that they are submitting and receiving files from the SBR as needed;
- ensuring the quality of frozen frames before they are released to users;
- identifying significant errors;
- ensuring that updates are timely, coherent and accurate; and
- documenting concepts, processes or coverage changes over time.

IT and systems section

Maintaining the SBR requires dedicated IT staff. The section ensures that the systems and software needed for extracting data are properly programmed and optimized. The section also maintains, and possibly advances or further develops, the graphical software that enables SBR staff to easily access and update the SBR content. This may initially be restricted to packaged server software (e.g., Microsoft SQL Server Management Studio). In addition the section ensures that certain derivation tasks are programmed and run correctly.

At least one database administrator (DBA) is required. The DBA interacts with the database management system processes and tables that make up the SBR, and ensures that the tables are accessible and available for production of the outputs. The DBA is responsible for ensuring that the database is functional. The DBA also ensures that the database can be updated and queried as required.

The number of dedicated IT personnel required for the implementation and maintenance of the SBR depends on the volume and nature of the functional and non-functional requirements for the specific SBR implementation. They must also be competent to handle specific IT requirements, to develop and maintain the SBR application, and to test newly developed features. The number required could range from one or two dedicated persons for a simple SBR to a team of experts during the course of a complex implementation.

The types of IT specialists required include:

- architect;
- database administrator;
- developer, with knowledge (as needed) of .NET, Java, SAS, T-SQL, PL/SQL;
- information analyst;
- IT project manager;
- systems analyst;
- tester.

In smaller teams one person may perform more than one type of task.

11.4 IT Considerations

11.4.1 Introductory Remarks

This section provides notes and recommendations on the IT infrastructure and programming requirements for the build phase of an SBR system. Cost and resource requirements can be inferred.

When establishing an SBR, there are many possible technologies. The choice should take into account scalability, cost and maintenance. The technology should be flexible enough to evolve with new requirements.

As there is no international standard or even commonly used practice amongst NSIs regarding the design of an SBR system per se, general international IT standards and guidelines, such as the Open Group Architecture Framework (TOGAF) and Solutions Integrated Development (SID) should be considered. The main consideration is to develop an SBR system that fits within the NSI IT architecture and that is as compatible as possible with other systems like the administrative data acquisition systems and the business survey collection systems.

11.4.2 Project Management Methodology

The development and implementation of a new SBR is a significant undertaking and should be managed as a project. If the NSI has a defined project management methodology, then it should be applied. If there is no existing organizational standard, then well accepted international standards such as PRINCE2 and PMI/PMBOK should be evaluated with the goal of selecting one of them.

As a general rule, the level of rigour in implementing the project management methodology should be scaled to the size of, complexity of, and risk associated with, the project.

11.4.3 Software Development Methodology

The development and implementation of the system required to support the establishment of the SBR is always a substantial part of the overall effort required. Many organizations have IT standards and, of course, these should be followed. Where this is not the case, an appropriate methodology should be adopted. Modern methodologies place a high focus on agility, demonstrating value to stakeholders quickly and regularly, managing change, proving the technical architecture early, managing risk continuously, and establishing frequent checkpoints (quality gates) for evaluation and realignment of scope, cost, schedule and quality.

IT development can be implemented in two phases:

- *acquiring an initial database infrastructure* - keeping in mind the modular approach to growing the SBR, the infrastructure should be expandable and flexible;
- *programming and process development* required in order to create:
 - the tables that make up the database;
 - the main statistical outputs of the frame (sampling frames, sample files, etc.);
 - the applications, programs and tools required to query and update the database to support business workflows and interface with other systems, for example, the administrative data, survey, and collection systems;
 - data transfer processes in a standard data format (for example SDMX).

As with project management methodology, the software methodology should be scaled to the size, complexity and risk of the project.

11.4.4 Solution Architecture

The requirements for an IT solution are commonly split into two groups; *functional requirements* and *non-functional requirements*.

Functional requirements indicate what the system is expected to do, i.e., the set of functions and user tasks that the system must be able to implement. A function can be described in terms of inputs, behaviour and outputs. The scope of the SBR coverage, the number of characteristics (real, estimated and derived), and the required interfaces and interactions with other systems are, along with other items, reflected in the functional requirements of the system.

Non-functional requirements indicate overall requirements that do not pertain specifically to functions of the system. Some examples are compliance with organisation's technical standards, performance, scalability, security, availability, accessibility, quality, and usability.

Functional and non-functional requirements are necessary inputs to decisions that need to be made regarding the solution architecture. The architect considers patterns, practices, tools

and technologies to arrive at a design and system architecture that optimally satisfies the requirements. Most modern, complex solutions are implemented with a layered architecture, of which a simple generic example is illustrated in Figure 11.2.

Figure 11.2: Layered architecture

Presentation layer	Implements the user interface and manages user interaction with the system
Service layer	Exposes interfaces and system functionality to other systems, and may also be the boundary between the presentation and business layers.
Business layer	Implements the core functionality and business logic.
Data access layer	Implements access to and interaction with data stores.
Data sources	Typically DBMS.

11.4.5 Database

There are a number of options, including a relational database management system (RDBMS), an object-oriented database management system (OODBMS), a key-value store, hierarchical database system, flat files, or even a spreadsheet based solution. An RDBMS is the predominant choice for a core piece of infrastructure such as an SBR. It is recommended on the basis of the following strengths:

- extensive capacity to implement security;
- extensive capacity for scalability;
- extensive capacity for support of concurrent users;
- capacity to enforce referential integrity and improved data integrity;
- efficient storage facilitated by normalization and lack of data duplication;
- efficient application code;
- flexible and standardized query language; and
- ability to be extended for future requirements.

A relational database enables proper segmenting of the areas that make up the whole, which in turn enables a proper maintenance strategy. The database should be scalable as there may be an increasingly large number of observations to process.

SBR data should be stored in a purpose built database that can support all the types of units in the economic units model, and their inter-relationships and characteristics, and that has appropriate confidentiality and access provisions.

There should be provision in the database, or within an associated, controlled environment, for storing and accessing data generated by all SBR functions, including snapshots, frozen frames, survey control files, respondent reporting obligations and statuses, and respondent burden.

Frame, collection and respondent burden modules

In its primary capacity of providing frames for surveys, the SBR should be seen as a database that stores both the frame data (including unique identifiers and stratification characteristics) as well as data about the collection processes. The required collection data include the units sampled, information on how and from where data about these units are to be collected, and the outcomes of the collection efforts. These data are input to the SBR respondent burden

module that provides an agency-wide view (i.e., across economic statistical programs) of the reporting burden that the NSI places on enterprises, and a record of any efforts to mitigate this burden, for example by using tax data in place of survey data.

A key benefit of using a relational database as the foundation of the SBR is that tables can be created, developed and maintained apart from one another. This provides great flexibility regarding resource requirements and allocation, the sophistication of the various tables, and the capacity to build on different modules at different rates. Thus, initial resources should be devoted to developing the core SBR tables and its main outputs, with simple collection tables to track which units in the SBR are sampled most heavily. Future resources can be allocated more efficiently as usage is tracked. The respondent burden module may be developed at a later stage.

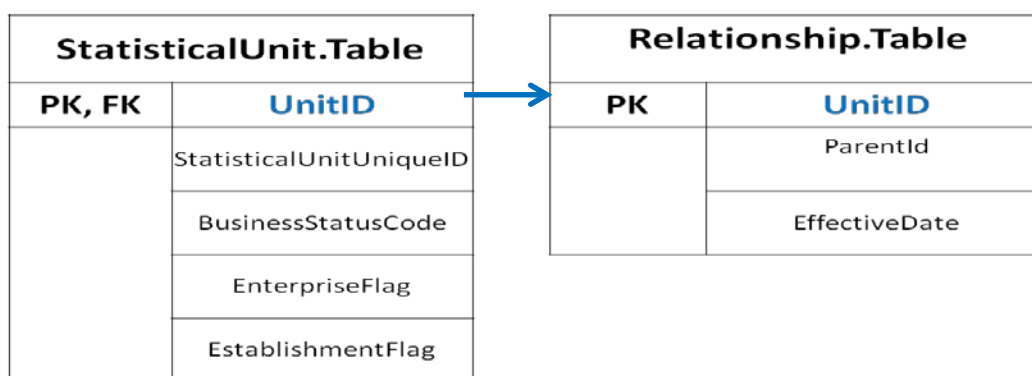
11.4.6 Database Management System (DBMS) Options

Relational databases and SQL

The modular approach to building the SBR is aided by using a relational database structure that enables creation of different tables and modules in isolation from one another. Using a RDBMS solution such as Microsoft SQL Server, or Oracle, offers other benefits too, including predefined software server applications, like SQL Server Management Studio, and international standards for querying the database using Structured Query Language (SQL).

Figure 11.3 below, represents two tables from a relational database model. The first table (the Statistical Unit table), contains the statistical unit's unique identifier with a few basic characteristics of the units, and the second table (the Relationship table), contain the parent's identifiers of each unit.

Figure 11.3: Simplified data model example of 2 tables



Database cost considerations

Functional and non-functional requirements for the SBR can have significant cost implications at the database layer in terms of licensing, amount and type of storage required, quantity and capacity requirements of the database servers, and levels of effort required by the database administrator (DBA) and other IT support staff. Here are some items that need to be considered.

- *Functional requirements:* coverage levels, number of characteristics, scope of functionality implemented (i.e., number of subsystems, screens and business processing logic), and interfaces with other systems; these requirements have a direct effect on the number and size of database tables required, and therefore on the storage required.

- *Non-functional requirements:* performance, load, availability, and number of environments; these requirements have an effect on the cost of the IT hardware required to implement the solution.

An enterprise class RDBMS enables addition of servers and databases as resources and growth dictate.

Another benefit of a relational database is that users can link tables across several databases to perform analyses and track data.

Simple database solutions

Compared with enterprise class RDBMS solutions, there are other, lower cost options for simple database solutions. However, the reduced costs are accompanied by loss of functionality and flexibility (for usage and future growth) compared with what a relational database offers. For example, MS Access can be used to run a limited SBR, but with several shortcomings:

- the inability to program sophisticated data constraints;
- limitations on how users can query and update tables simultaneously;
- constrained ability to proceed with a modular approach that enables gradual introduction of databases and tables;
- limited security controls.

Example of database structure from Statistics Canada

In Canada, the SBR is now maintained in about 30 databases on five separate servers. The online transaction processing (OLTP) Production Database is a separate database and server that can only be accessed and updated using the Business Register System (BRS) application. A snapshot of the frame is stored in a separate database on a separate server, so that analysis, quality assurance and reporting activities do not impact the OLTP performance. Likewise, the Survey Interface Repository (SIR), which enables survey areas and systems to access information required for sampling and analysis of survey collection information, is stored on a separate server and database so that its workload does not impact the OLTP performance and scalability.

Database construction and organization

Designing database tables involves deciding which tables will house which characteristics, and the manner in which these tables should interact. The IT team should also determine how data are backed up and secured, and should create a set of database roles (for example, *write access*) to be assigned to users, as needed.

Another important task in table design is creating unique identifiers for the records in the SBR, both to identify specific records and to identify relationships between the records. Creating and maintaining hierarchical relationships is essential as they enable users to denote the parent–subsidiary relationships of enterprises, and the operational links between enterprises and establishments.

Advantages of relational database: quality assurance and data coherence processes

Another great advantage of a relational database system is that data constraints can be placed on tables. This ensures that inputs, either from statistical or from administrative sources, can be required to meet certain criteria in order to be loaded. This in turn enables automated

quality checks on data when they are loaded. Furthermore, dividing different types of information into multiple tables allows for data segregation, and thus analysis can be atomized further.

The IT development team should be contracted to insert appropriate data constraints. As well, they should design *data coherence programs* that can be used to determine outlier data in tables and to validate the internal coherence of certain information, for example, that the revenue of an establishment is not greater than the revenue of the corresponding enterprise. Constraints on data should also be implemented as a control on the quality of the data being received and loaded from external sources.

11.4.7 Programming Requirements

Graphical user interface and database server software

The manner in which users interact with the SBR is a crucial decision in the build process. In deciding whether or not to develop a tailor-made graphical user interface (GUI), two groups of factors should be considered: first, the *types of users and their requirements*; and, second, *the relationship between the SBR itself and its outputs*.

Users of the SBR include:

- SBR staff whose task is to ensure that SBR information is accurate and up to date;
- survey statisticians and analysts, who want to review the enterprises and corresponding structures in the SBR that affect their own statistical programs with a view to updating their particular populations;
- sampling statisticians and other statistical staff, who analyse their frames to select samples.

Each group of users has different requirements. Sampling statisticians are most interested in analysing the outputs from the frame at aggregate level to help in optimising sample design, and thus, for them, a GUI is not particularly important. On the other hand, profilers and survey statisticians greatly benefit from a GUI that enables them to easily search, browse and update individual units in the SBR. A GUI also minimises the need for all staff to understand SQL and database querying.

On the other hand, a GUI can involve large development and maintenance costs, including those for design, coding and additional server infrastructure. Without a GUI, there is still plenty of capacity to group and analyse enterprises based on the database tables. Initially running database server software like SQL Server Management Studio, or SAS, to access and query SBR data may be economical. The SBR can be made operational and can start producing the needed outputs more quickly and cheaply than if a GUI has to be developed.

In summary, given that a GUI is not an essential SBR tool, in the initial stages of SBR development the cost-benefit is tilted against introducing a GUI. However, a GUI may be integrated into some standard database software. Also, keeping in mind that survey frames are ultimately the main output of the SBR, a GUI is a tool that may be worth introducing at a later stage to enable statistical staff to more easily browse, profile and update the units in their frames.

Production of survey frames and sample files

The extraction process by means of which survey frames are produced requires relatively little programming, as it is a derivative product of the database tables themselves. The exact design and types of SBR database tables, and the unique identifiers to be used, should have already been determined before the design of the extraction processes. The main effort

required from survey and sampling statisticians is in determining which characteristics should be available in the frames, and which in the sample files.

Creating a data repository to archive frames and sample files is recommended as these are the primary outputs of an SBR. Maintaining these files is important from a data management perspective. They also offer a key input for future analysis and development work.

11.4.8 IT Environments

The IT solution should properly manage the deployment of new features. A full scale solution may involve five distinct environments.

- 1) *The production environment* should be a dedicated version of the system, including the active data that are being updated.
- 2) *The practice environment* has the same code as production, but the data are not being continuously updated. This environment is typically used to perform updates that mimic what would happen in production. It enables users to check and see how their changes would affect the overall process.
- 3) *The user acceptance environment* enables testing of new programming functionality before moving the code into the production and practice environments. It typically has fictitious data that are developed solely to test various scenarios.
- 4) *The development environment* is dedicated to systems programmers, who use it to test their own programming. Once code is system-tested, it can be moved into the user acceptance environment to be tested by users.
- 5) *The analysis environment* is dedicated to analysts performing quality evaluation and simulation testing. This environment offers analysts access to a mirror image of the data available in the production environment without disturbing the production processes.

The first environment is, of course, essential, and the third and fifth environments are highly desirable. At a bare minimum, there should be a production environment and one other environment in which all forms of testing and analysis can take place.

It is important to consider how all other systems link to the SBR when establishing these environments. For example, if end-to-end testing across systems is required, the SBR user acceptance environment needs to be linked to the user acceptance versions of the linked systems.

11.4.9 Programming Language

The choice of programming language is crucial, and the options should be researched thoroughly. A poor choice could lead to premature redesign and costly maintenance processes. The chosen language should be sufficiently common that knowledgeable staff and training are readily available, and it should have a demonstrated long shelf life.

11.4.10 Relationships with Other Systems and Registers

The SBR should be established knowing that it may be linked to other registers or lists. Thus, from the start, standard definitions and concepts should be used, and all the most commonly used unique identifiers for each type of businesses should be included.

If a commonly known identification numbering system for legal units does not exist, the SBR should create one itself.

Developing a flexible SBR system that readily enables adding to and maintaining the various concordances between identifiers is important. It facilitates linking the SBR to satellite systems that create and maintain their own identifiers.

Example from Canada

In Canada every business that registers with the Canada Revenue Agency is assigned a *Single Business Registration Number (SBRN)*, commonly known as the BN, or Business Number. Thus, it is essential, first, for SBR staff to have a clear understanding on how to relate and/or link the BN to statistical units and, second, to include the BN within the SBR as a key identifier.

11.4.11 Establishing a Unique Identifier for Statistical Units

The establishment of unique identifiers is essential for accurate maintenance of the SBR. There is a technical need to have a unique identifier to load and maintain each unit within the SBR database. The sequential assignment of unique identifiers is effective. It is even more effective when the unique numbers are created and managed centrally, and are used by all statistical systems and processes throughout the entire statistical organisation. Having SBR assigned identifiers is also a good means of reducing the risk of inadvertently disclosing confidential micro-data by use of easily identifiable and recognizable information such as the tax number or the business name.

Here are some of the key elements to consider when creating a unique identifier for the SBR.

- Create an identification numbering system *for each statistical unit, no matter what type*. Multiple business numbers lead to added complexity in SBR systems, and introduce risks such as duplication and omission of statistical units.
- Use a *non-confidential identifier* in order to facilitate the statistical processing. It should not contain any information about the business. All the statistical processing such as edit, imputation and estimation can make use of this non-confidential identifier in their processes.
- Ensure these unique identifiers have *no meaning* (other than, possibly, indicating the type of unit) and are *only be used for statistical purposes*.
- Ensure that the unique identifiers *cannot be reused*. Ensure that the length of the identifier is sufficient to account for the total number of statistical units the SBR may ever contain over time, including dead units and units of all types.

Technical considerations in the generation of a unique identifier are as follows.

- In order to minimise errors, consider including a check digit function, especially if the unique identifier will need to be manually captured (e.g., through *heads down data entry*) by SBR users when they enter information about specific units. (An example of how to calculate a check digit according to Modulo 11 is included in Annex G3.)
- If the identifier is always generated centrally by the SBR, use can be made of a *key generator function* that guarantees to generate a unique sequential number. If need be a check digit can be added to the number.
- Alpha characters combined with numbers can be used in order to avoid confusion with any other numeric data value on the database.

Example from Canada

The unique identifier of the Canadian SBR is called the *Statistical Number*. The Statistical Number is composed of a letter and followed by an 8-digit number. The letter “S” is used in order to easily identify the Statistical Number. The alpha character combined with numbers avoids confusion with other numeric data value in the database.

In the database, there is a function called the *Key Generator* that generates a unique sequential 8-digit number, so no digit check is required. This function also ensures the uniqueness of the generated number. Here is an example of a Statistical Number structure generated in the SBR to identify uniquely a new created business entity: *S12345678*.

The *Key Generator* function is used when a new statistical unit is added on the SBR. A new unit can be added online by an employee or from a batch process. The *Key Generator* is invoked and uses the Statistical Number Table that contains all potential 8-digit numbers from 00000001 to 99999999 to determine the next available number. When the function locates the next available number, then it automatically assigns an “S” in front of the number and assigns it to the new statistical unit. As it performs this assignment, the *Key Generator* automatically removes this number from the Statistical Number Table, leaving in the Table, only numbers that have not been assigned. This technique permits multiple processes to call the *Key Generator* function simultaneously and to have access to a new 8-digit number.

11.4.12 Tools/Software for Record Linkage

The results quoted for the performance of automated data matching tools and software tend to be overly optimistic. Experience has shown that, when bringing together data from two sources without a known and unique correspondence between the sets of identifiers, deterministic matching or probabilistic record linkage often yields mismatches when the matching rules are too loose and a high percentage of missed matches when the rules are too rigid. Nevertheless, circumstances may dictate the need for record linkage. No specific tools or software are recommended in this chapter. Chapter 6.8 provides some guidance.

11.4.13 Job Scheduling Software

The administrative data used to update the SBR usually requires a great deal of time to process. Thus, to reduce the impact on other forms of processing, administrative data are usually processed when no one is accessing the database — typically at night or over the weekend. Job-scheduling software enables efficient management of the process of running computer jobs automatically. It monitors processing and should remotely notify technicians of any problems to be resolved before users next access the database.

11.4.14 Documentation

Documentation related to the system is needed to ensure the long-term functioning of the SBR. It helps staff identify and understand the changes they need to make from time to time.

The IT team should invest significant time in fully documenting each module and process, both during initial development and as the SBR evolves. The documentation should be detailed enough for a new programmer to continue where a previous one left off. Wikis enable free-form information entry over time.

11.5 Data Retention

SBR data retention strategy should be articulated in accordance with operational and analytical needs and should begin with the determination on how changes made to the SBR will be tracked and what historical information will need to be kept.

Tracking changes

The following paragraphs summarise briefly the options to consider for the safekeeping of historical information, like the usage of logs, the creation of snapshots and frozen frames on a regular basis, and the recording of the *effective date* (i.e. the time that an event really occurred) as well as the *update date* (i.e. the date when the event was recorded on the SBR).

In determining the strategy, a key factor is how changes to SBR data are handled. One possible approach is *to always add data, never to replace data*. With this approach, the (new) value of characteristic is recorded together with the date and time it was recorded. The old value (if any) of the characteristic, and the date and time it was recorded are retained. This approach enables the creation of a view of the SBR as of any past date and time. In a more sophisticated approach the date and time from which a new value is *effective* is recorded in addition to the date and time of the update.

Another, much simpler approach is to take periodic snapshots of the database and keep these for as long as seems necessary. To satisfy all operational and analytical needs, using both approaches is suggested.

Frequency and content of snapshots

In so far as snapshots are used, data retention should be frequent enough to offer users a reasonably comprehensive view of the past. Monthly snapshots of all statistical units and their characteristics going back two years would likely offer users sufficiently frequent data for the majority of their immediate needs. As regards data retention for analytical purposes a set of frozen frames may suffice for most users. Some NSIs retain these files indefinitely, as they are a great source of information for longitudinal studies.

Example from Statistics Canada

At Statistics Canada a complete copy — called a snapshot — of the SBR database (live register) is taken just prior to the first day of every month. A generalised survey universe file (GSUF), i.e. frozen frame, containing every statistical unit, is created from the snapshot every month. Although frozen frames are primarily used for sampling, normally soon after their creation, they are retained for an extended period for analysis purposes. The table below shows the current retention period for each frozen frames.

Monthly Frozen Frames	Retention period
January	Indefinite
February to December	24 months

Administrative updates

Keeping track of updates from administrative sources is also recommended. An automated log should include the update date and source. This information may prove useful in

troubleshooting issues with a specific maintenance or update processes. It may also serve to document and explain changes to SBR users.

Documentation

As the SBR undergoes constant evolution in concepts, coverage and methods, the data retention strategy must also contain documentation explaining these changes. Each file stored should have a full account of major changes attached to it. This enables future analysts to understand the data. The quality assurance team may also use the log information to assess the quality of certain maintenance processes or sources of SBR updates.

Once it is determined *how* changes made to the SBR are tracked and what historical information the SBR stores, then an *information management directive (IMD)* should formalize *what* should be kept on the long term.

Information Management Directive

Assuming an IMD exists, the SBR should follow it in determining its data retention strategy. If no IMD exists, a data retention strategy should be articulated in accordance with operational and analytical needs, taking into account future cost and space requirements and general NSI policies and legislative restrictions.

To determine appropriate retention periods for specific sets of information, it is suggested to develop:

- groups for the various types of statistical micro-data files; and
- information management categories, i.e. categories of retention and documentation requirements.

For example, four primary groups (each group relating to one or more phases of the generic statistical business process model) could be created.

1. Statistical microdata – design and collection.
2. Statistical microdata – processing.
3. Statistical microdata – analysis and dissemination.
4. Statistical microdata – other.

Five information management (IM) Categories could be created.

1. Indeterminate retention and “high” level documentation.
2. Maximum 20 year retention and “medium” level documentation.
3. Maximum 10 year retention and “medium” level documentation.
4. Maximum 5 year retention and “minimum” level documentation.
5. Other (for a few file types that require special attention).

Archiving and deleting

As the SBR content continuously grows, the data retention strategy must also contain documentation and rules on data and information that does not need to be retained (i.e. data and information that can be deleted/disposed of) and data and information that do not need the same ease of access (i.e., data and information that can and should be archived).

Below is a list of business rules that should be considered when developing any archiving and deleting processes.

- *Automation*: The process needs to be automated with minimal manual intervention.
- *Scheduling*: The process needs to be able to schedule using a standard scheduler.

- *Logging*: A log of the tables (files) deleted (or flagged for deletion) needs to be recorded, centralized and accessible to all users having access to the process.
- *Reporting capability*: The process should create a report.
- *Possibility of review*: The process must allow the user to review the actions to be performed before being applied.
- *Possibility of executing report only*: The process should allow the user to execute the report portion of the process only (i.e. not apply the updates).
- *Possibility of resuming execution*: The process should allow the user to resume execution of the update portion of process when the reporting portion has been executed beforehand. The amount of time spent between the execution of the report portion and the update portion should be flexible.
- *Possibility of retrieval*: The process must allow tables that have been deleted to be retrieved if deleted by mistake (at least the last deletion batch)
- *Validation of existing cycle*: The process must ensure that a previous cycle exists for all tables (files) that have been identified to be deleted.
- *Flexible parameter modification*: The process should use the already existing files, tables and information so that modifications to concepts are automatically applied to the job. Parameters should not be hard coded in the process.

Annex A: References

UN and Other International Agencies

Generic Statistical Business Process Model

UN Economic Commission for Europe

<http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0>

Guidelines for Building Statistical Business Register in Africa

African Development Bank (latest version, 2014)

<http://www.afdb.org/en/knowledge/publications/guidelines-for-building-statistical-business-registers-in-africa/>

International Standard Industrial Classification of All Economic Activities (ISIC) Revision 4

Statistical Paper Series M No.4/Rev4, United Nations (2008)

http://unstats.un.org/unsd/publication/seriesM/seriesm_4rev4e.pdf

ISO 9000:2005 Quality management systems -- Fundamentals and vocabulary

International Standards Organisation (2005)

http://www.iso.org/iso/catalogue_detail?csnumber=42180

Measuring the Non-Observed Economy: A Handbook

Organisation for Economic Cooperation and Development (2002)

<http://www.oecd.org/std/na/1963116.pdf>

National Quality Assurance Frameworks

United Nations Statistics Division

<http://unstats.un.org/unsd/dnss/QualityNQAF/nqaf.aspx>

Resolution concerning statistics of employment in the informal sector,

adopted by the Fifteenth International Conference of Labour Statisticians

(January 1993)

http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087484.pdf

System of National Accounts 2008

United Nations et al, New York, 2009

<http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Working Paper on Statistical Business Registers

United Nations Statistical Division and Eurostat (2011)

European Commission and EUROSTAT

Business Registers Recommendations Manual

Eurostat, Methodologies and Working Papers, Publications Office of the European Union, Luxembourg (2010)

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

Business Register Regulation 2008

Regulation (EC) No. 177/2008 of the European Parliament and of the Council of 20 February 2008

<http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32008R0177>

European System of National Accounts

EU Regulation (EC° N° 549/2013) of 21 May 2013 on the European system of national and regional accounts in the European Union

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:174:0001:0727:EN:PDF>

Framework Regulation for Integrated Business Statistics (FRIBS)

Eurostat (Draft, 2014)

Handbook on Data Quality Assessment Methods and Tools

EUROSTAT (2007)

http://ec.europa.eu/eurostat/ramon/statmanuals/files/Handbook_on_data_qual_assess_tools.pdf

Manual on Business Demography Statistics

EUROSTAT, OECD

ISBN 978-92-79-04726-8 (2007)

<http://www.oecd.org/std/39974460.pdf>

Seventh Council Directive 83/349/EEC of 13 June 1983 based on Article 54(3)(g) of the Treaty on consolidated accounts

http://europa.eu/legislation_summaries/internal_market/businesses/company_law/126010_en.htm

Report of the Work Package B of the ESSnet on profiling large and complex MNEs: Conceptual framework, methodology, rules and standards

Eurostat (2014)

Statistical Units Regulation

Council regulation (EEC) No 696/93 of 15 March 1993 ()

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1993R0696:20081211:EN:PDF>

Task Force on Statistical Units: Conclusions of meeting 21-22 January 2014 ESTAT/G1/TF SU/21-22

Eurostat (January 2014)

National Statistical Institutes and Statisticians

A theory for record linkage

Fellegi, I.P. & A.B. Sunter, *Journal of the American Statistical Association*, 64, pp. 1183-1210, 1969

Beurteilung ausgewählter Stringvergleichsalgorithmen zur Eignung für Record Linkage an Hand einer empirischen Datenbank.

Weghofer E.: Diplomarbeit an der WU Wien. 2004

Checklist for the Quality Evaluation of Administrative Data Sources

Piet Daas, Saskia Ossen, Rachel Vis-Visschers and Judit Arends-Tóth Piet Daas, Saskia Ossen, Rachel Vis-Visschers and Judit Arends-Tóth,

Discussion paper 09042, Statistics Netherlands (2009)

<http://www.cbs.nl/nr/rdonlyres/Odbc2574-cdae-4a6d-a68a-88458cf05fb2/0/200942x10pub.pdf>

DANE, Metodología Directorio Estadístico de Empresas –DEST–, Bogotá, D. C, Colombia. Versión 5. 2013.

http://www.dane.gov.co/files/investigaciones/boletines/ipp/bol_ipp_ago13.pdf

Data Quality and Record Linkage Techniques,

Thomas N. Herzog, Fritz J. Scheuren, William E. Winkler

Springer 2007,

Data quality: A statistical perspective

Karr, A.F., Sanil, A.P., Banks, D.L.

Statistical Methodology, 3, pp. 137-173 (2006)

Re-thinking Statistics Canada's Business Register

Bérard, Hélène, Pursey, Stuart & Rancourt, Eric., *Statistics Canada.*

https://fcsmsites.usa.gov/files/2014/05/2005FCSM_Rancourt_Berard_Pursey_IVB.pdf

Seeing the whole elephant: A proposed experiment on measuring the activities of multinational enterprises

Richard Barnabé, Statistics Canada (2003)

<http://www.unece.org/fileadmin/DAM/stats/documents/ces/2003/13.e.pdf>

Statistical Business Register: Content, place and role in Economic Statistics.

Ritzen, Jean. Statistics Netherlands

http://ec.europa.eu/eurostat/ramon/coded_files/RITZEN_Jean_SBR.pdf

String comparator metrics and enhanced decision rules in the Fellegi-Sunter model of record linkage“

Winkler, W:E., Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 354-359, 1990

The French Business Register: From a quality approach to a statistical register.

Projects on the Improvement of Business Register, Session 5.

<http://www.oecd.org/industry/business-stats/41691270.pdf>

UNIMATCH –a computer system for generalized record linkage under conditions of uncertainty

Jaro, M.A., Spring Joint Computer Conference, 1972, AFIPSL-Conference Proceedings, 40, pp.523-530, 1972

Other Sources

The barcode solution

Active Barcodes

<http://www.activebarcode.com/codes/checkdigit/modulo11.html>

Circular Flow of Income

http://en.wikipedia.org/wiki/Circular_flow_of_income

Annex B: Glossary⁷⁹

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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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 identified by the topdown method as the activity which contributes most to the total value added of the unit under consideration. The principal activity so identified does not necessarily account for 50% or more of the entity's total value added.
- 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 an unit by providing non-durable goods or services for the use of that entity.

Source: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

⁷⁹ Definitions may slightly differ from the wordings in the original sources as the meanings of the terms are expressed in the way that best fits the purpose of a statistical business register.

Related terms: Active unit, economic activity, classification of activities, ancillary activity, secondary activity, principal 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 purposes (e.g. recording and maintaining certain details of businesses or taxation)

Source: Eurostat, Manfred EHLING, Thomas KÖRNER and others, "Handbook on Data Quality Assessment Methods and Tools", Wiesbaden, 2007

Link:

http://ec.europa.eu/eurostat/ramon/statmanuals/files/Handbook_on_data_qual_assess_tools.pdf

Related terms: Administrative register, statistical business register

Administrative data

Data originally collected for non-statistical purpose. Complete coverage is the aim. 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: Eurostat, Manfred EHLING, Thomas KÖRNER and others, "Handbook on Data Quality Assessment Methods and Tools", Wiesbaden, 2007

Link:

http://ec.europa.eu/eurostat/ramon/statmanuals/files/Handbook_on_data_qual_assess_tools.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 particular set of objects. Typically a register is some sort of structured list of units, containing a number of attributes for each of those units, and having some sort of regular updating mechanism. Register in this way and maintained by administrative authorities for administrative purposes can be considered to be administrative registers.

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000.

Link: <http://www.unece.org/stats/publications/53metadaterminology.pdf>

Related terms: Administrative business register

Administrative source

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

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Administrative register, administrative data

Administrative unit

An administrative unit is defined by a legal unit for the purposes of conforming to an administrative regulation, for example VAT.

Source: International Guidelines for Business Registers

Related terms: Statistical unit

All-resident enterprise group

Enterprise group composed only of enterprises that are all resident in the same country.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Enterprise group, multinational enterprise group, truncated enterprise group

Ancillary activity

Ancillary activities are undertaken to support principal and secondary productive activities of an unit by providing goods or services entirely or primarily for the use of that entity. 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: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Activity, economic activity, principal activity, secondary activity

Attribute

See: Characteristic

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-off 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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)

Branch

A branch is an unincorporated enterprise that wholly belongs to a non-resident unit, known as the parent. It is resident and treated as a quasi-corporation. For the purpose of the business register a branch shall be treated as an enterprise.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Foreign affiliate, subsidiary

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 split-off and can be seen as the opposite of a merger.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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

Term is used as a type of enterprise, namely a "commercial enterprise" or legal unit with commercial economic activity

Source: International Guidelines for Business Registers

Related terms: Enterprise

Business closures

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 time dimension, thus gaining information on their survival or discontinuity. It also covers development in 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 by any statistical unit; however, a clear political interest in Europe is on enterprise demography. 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: European Commission (Eurostat), Organisation for Economic Cooperation and Development (OECD), "Eurostat – OECD Manual on Business Demography Statistics", Methodologies and Working papers, Publication Office of the European Communities, Luxembourg, 2007.

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

Related terms: Continuity, survival

Business start-ups

See: Creation of business

Business statistics

See: Economic statistics

Captive financial institution

Activities of holding companies, i.e. units that hold the assets (owning controlling-levels of equity) of a group of subsidiary corporations and whose Institutional units and sectors principal activity is owning the group are treated as captive financial institutions. The holding companies in this class do not provide any other service to the enterprises in which the equity is held, i.e. they do not administer or manage other units. Other units that are also treated as captive financial institutions are units with the characteristics of SPEs including investment and pension funds and units used for holding and managing wealth for individuals or families, holding assets for securitization, issuing debt securities on behalf of related companies (such a company may be called a conduit), securitization vehicles and carry out other financial functions. The degree of independence from its parent may be demonstrated by exercising some substantive control over its assets and liabilities to the extent of carrying the risks and reaping the rewards associated with the assets and liabilities.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Control, ownership, SPE

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.

Source: International Guidelines for Business Registers

Related terms: Variable

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: European Commission (Eurostat), Organisation for Economic Cooperation and Development (OECD), "Eurostat – OECD Manual on Business Demography Statistics", Methodologies and Working papers, Publication Office of the European Communities, Luxembourg, 2007.

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

Related terms: Business closures, death (of business)

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 a wide range of purposes, whereby each has its own specific area of application depending on the subject of classification. Economic classifications can be broadly divided into two categories: Classifications of economic activities cover all economic activities - from agriculture to services - and are used to classify economic entities (enterprises, local units and similar statistical units). Such classifications therefore form the basis for compiling statistics on output, the production factors entering into the production process (input: labour, raw materials and supplies, energy etc.), capital formation or financial transactions. The outputs of the economic entities

are termed products and are generally divided into goods and services. They are classified in product classifications. The international classification for activity is ISIC (International Standard Industrial Classification of All Economic Activities, maintained by the United Nations, used at world level).

Source: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Activity

Common frame

See: Frozen frame

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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Mergers, 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: Eurostat, "Business registers. Recommendations manual", Methodologies and Working Papers, Publications Office of the European Union, Luxembourg, 2010

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

Related terms: Business demography, Survival

Control

A single institutional unit owning more than a half of the shares, or equity, of a corporation is able to control its policy and operations by outvoting all other shareholders, if necessary. Similarly, a small, organized group of shareholders whose combined ownership of shares exceeds 50 per cent of the total is able to control the corporation by acting in concert.

Corporation B is said to be a subsidiary of corporation A when: Either corporation A controls more than half of the shareholders' voting power in corporation B; or corporation A is a shareholder in corporation B with the right to appoint or remove a majority of the directors of corporation B. Corporation A may be described as the parent corporation in this situation. As the relationship of a parent corporation to a subsidiary is defined in terms of control rather than ownership, the relationship must be transitive: that is, if C is a subsidiary of B and B is a subsidiary of A, then C must also be a subsidiary of A. If A has a majority shareholding in B while B has a majority shareholding in C, A cannot also have a majority shareholding in C. Nevertheless, A must be able to control C if it controls B.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Ownership, subsidiary

Control survey

See: SBR improvement survey

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: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 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: European Commission (Eurostat), Organisation for Economic Cooperation and Development (OECD), "Eurostat – OECD Manual on Business Demography Statistics", Methodologies and Working papers, Publication Office of the European Communities, Luxembourg, 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-ups, new enterprise, take-over

Data collection units

Sometimes data cannot be obtained about and from a statistical unit. To handle this situation, two units are defined specifically with collection of data in mind

Source: International Guidelines for Business Registers

Related terms: Observation unit, reporting unit

Date of birth (of enterprise)

See: Date of creation (of enterprise)

Date of birth (of enterprise group)

See: Date of creation (of enterprise)

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. This variable refers to the death or other deletion date of the enterprise (when it becomes historical) and is interpreted in a way similar 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. It does 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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 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 of its having taken 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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 legal unit)

Date of cessation (of local unit)

Date of final cessation of activities. This variable 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. The disappearance of a local unit, which existed before. 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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 local unit)

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: European Commission (Eurostat), Organisation for Economic Cooperation and Development (OECD), "Eurostat – OECD Manual on Business Demography Statistics", Methodologies and Working papers, Publication Office of the European Communities, Luxembourg, 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 to the Member State 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: European Commission (Eurostat), Organisation for Economic Cooperation and Development (OECD), "Eurostat – OECD Manual on Business Demography Statistics", Methodologies and Working papers, Publication Office of the European Communities, Luxembourg, 2007.

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 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: European Commission (Eurostat), Organisation for Economic Cooperation and Development (OECD), "Eurostat – OECD Manual on Business Demography Statistics", Methodologies and Working papers, Publication Office of the European Communities, Luxembourg, 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: European Commission (Eurostat), Organisation for Economic Cooperation and Development (OECD), "Eurostat – OECD Manual on Business Demography Statistics", Methodologies and Working papers, Publication Office of the European Communities, Luxembourg, 2007.

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

Related terms: Creation (of business)

Death (of business)

See: Date of cessation (of business)

Death (of enterprise)

See: Date of cessation (of enterprise)

Death (of enterprise group)

See: Date of cessation (of enterprise group)

Death (of legal unit)

See: Date of cessation (of legal unit)

Death (of local unit)

See: Date of cessation (of local unit)

De-concentration

De-concentration refers to demographic events involving one enterprise before and more than one enterprise after the events by 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Break-up, split-off

Delineation

The various needs of the users of the business registers require providing different units correctly delineated with respect to structure, and characteristics of the unit. The delineation of statistical units is done by grouping or dividing administrative or other relevant units according harmonised rules and also by using classifications to delineate the unit according its activity, location or any other characteristics.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Statistical unit, profiling

Domestically controlled enterprise group

From the point of view of the country where the group head of an enterprise group is located, the enterprise group is domestically controlled irrespective if the enterprise group is multinational or all-resident.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Enterprise group, foreign controlled enterprise group

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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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 are economic activities. 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 business registers.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Active unit, activity

Economic census

A survey conducted on the full set of observation objects belonging to a given business population.

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000.

Link: <http://www.unece.org/stats/publications/53metadaterminology.pdf>

Related terms: Economic survey

Economic operator

See: Economic unit

Economic organisation

See: 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: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

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: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000.

Link: <http://www.unece.org/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 SNA2008.

Source: International Guidelines for Business Registers

Related terms: Legal unit, economic production, statistical 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: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Activity, economic 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 into 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: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms Employment, number of employees, number of persons employed

Employment

Employment includes all persons, both employees and self-employed persons, engaged in some productive activity that is undertaken by a resident enterprise.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Employees

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 all 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: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Multinational enterprise, standard statistical unit

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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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

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.

Source: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Standard statistical unit

EuroGroups Register (EGR)

The EuroGroups Register (EGR) is a network of registers, consisting of a central register kept at Eurostat and registers in each EU Member State and in EFTA countries. The central register contains information about multinational enterprises groups (MNEs), which have statistically relevant financial and non-financial transnational operations in at least one of the European countries. Registers in the EU Member States and in EFTA countries contain information regarding MNEs active in the respective countries and are fully consistent with the central register. In practice both, Eurostat and European countries exchange confidential and non-confidential data on MNEs by EDI exclusively for statistical purposes. The EGR integrates data coming from commercial providers and from the National Statistical Authorities. The different sources are compiled to unique MNEs in the central register at Eurostat in co-operation with the concerned statistical authorities of the EU Member States and EFTA countries. The aim of the EGR network is to hold a complete, accurate, consistent and up-to-date set of linked and coordinated statistical registers, which offer compilers a common frame of multinational enterprise groups, global as well as truncated national groups, operating in the economy of the EU and EFTA countries, together with their constituent legal units and enterprises and the ownership and control relationships between legal units.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: European System of interoperable statistical Business Registers, multinational enterprise group

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). The ESA framework consists of two main sets of tables:

- Sector accounts: provide, by institutional sector, a systematic description of the different stages of the economic process: production, generation of income, distribution of income, redistribution of income, use of income and financial and non-financial accumulation. The sector accounts also include balance sheets to describe the stocks of assets, liabilities and net worth at the beginning and the end of the accounting period.

- Input-output framework and the accounts by industry: describe in more detail the production process (cost structure, income generated and employment) and the flows of goods and services (output, imports, exports, final consumption, intermediate consumption and capital formation by product group).

The ESA encompasses concepts of population and employment. These concepts are relevant for both the sector accounts and the input-output framework. The ESA is not restricted to annual national accounting, but applies also to quarterly accounts and regional accounts."

Source: European Commission (Eurostat), "European System of National Accounts 2010", Publication Office of the European Union, Luxembourg, 2013.

Link: <http://ec.europa.eu/eurostat/documents/3859598/5925693/KS-02-13-269-EN.PDF/>

Related terms: System of National Accounts

European System of interoperable statistical Business Registers (ESBRs)

Within a project of the European Statistical System (2013 - 2017) the way will be paved for a net of consistent and interoperable national statistical business registers in the EU and the statistical EuroGroups Register (EGR), which will operate as backbone for the European business statistics.

Source: European Commission (Eurostat), "The European System of interoperable Business Registers (ESBRs)", Compact guides, Publication Office of the European Union, Luxembourg, 2013.

Link: <http://ec.europa.eu/eurostat/documents/4031688/5931580/KS-03-13-411-EN.PDF>

Related terms: EuroGroups Register

Foreign affiliate

Foreign affiliate shall mean 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: Eurostat, "Recommendations Manual on the Production of Foreign Affiliates Statistics (FATS)", Methodologies and Working papers, 2007 Edition

Link: http://ec.europa.eu/eurostat/ramon/statmanuals/files/FATS_2007_EN.pdf

Related terms: Foreign Affiliates Trade in Services, foreign controlled enterprise group

Foreign Affiliates Statistics (FATS)

FATS mean statistics describing the overall activity of foreign affiliates. As there are two point of views 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: European Commission, Eurostat, "Foreign Affiliates Statistics (FATS) Recommendations Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2012.

Link: <http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-12-016>

Related terms: Foreign affiliate, foreign controlled enterprise group

Foreign controlled enterprise group

In general, a non-resident unit controls a resident corporation if the non-resident unit owns more than 50 per cent of the equity of the corporation. Branches of non-resident corporations are by their nature always under foreign control. Therefore an enterprise group controlled by a group head that has its headquarters located outside the country where part(s) of the enterprise group is (are) located is a foreign controlled.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Enterprise group, foreign controlled enterprise group

Foreign direct investment (FDI) statistics

FDI statistics embody four distinct statistical accounts: Investment positions, financial transactions, associated income flows between enterprises that are related through a direct investment relationship, and other changes in the value of assets, especially revaluation terms. Direct investment is a category of cross-border investment associated with a resident in one economy (the direct investor) having control or a significant degree of influence on the management of an enterprise (the direct investment enterprise) that is resident in another economy.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Foreign Affiliate Trade in Services, inward foreign affiliate statistics (I-FATS), outward foreign affiliate statistics (O-FATS)

Franchise

The operation of a franchise network is a method of doing business that is popular in a number of service activities, especially hotels, restaurants, and retail sales. Franchisees are independent legal units which sign a contract with another legal unit, the franchiser, to engage in an activity making use of trademarks, trading styles and marketing support provided by the franchiser, usually in return for a fee or a share of the profits. A franchise contract typically includes a number of restrictive clauses limiting the franchisee's freedom of choice, for instance imposing standards as to the goods and services to be produced, their quality and their price. The franchisee may be compelled to obtain supplies from the franchiser and must pay possibly access rights. The franchisee remains entirely responsible for his investment. Contribution towards certain services organised by the franchiser that is common to the entire network. The franchiser, in turn, offers scale economies without completely taking away the autonomy of the franchisee, for example by taking care of collective marketing. Franchise operators may or may not belong to the same enterprise group. Franchisees are deemed to be separate enterprises because they consist of a complete combination of factors of production, and they run the full entrepreneurial risk. Moreover, the definition of the enterprise requires autonomy but allows for this autonomy to be somewhat restricted ("a certain degree of autonomy" is required), and full accounts tend to be available only at the level of the separate franchisees. The franchiser is also regarded as a separate enterprise.

Source: Eurostat and Organization for Economic Cooperation and Development (OECD), "Eurostat - OECD Manual on Business Demography Statistics (Edition 2007)", Methodologies and Working Papers, Office for Official Publications of the European Communities, Luxembourg, 2007

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

Related terms: Enterprise

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 are extracted from the same common set of units, i.e. the frozen frame.

Source: International Guidelines for Business Registers

Related terms: Frozen frame

Frame population

See: 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 and nothing else. 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. However, as an SBR improvement survey may be one of the surveys, this may not be much of a restriction as the survey may be conducted precisely for the purpose of determining missing values of characteristics and hence require only minimal data about the unit.

Source: International Guidelines for Business Registers

Related terms: Frame, live register

Frozen register

See: Register snapshot

Global decision centre

Unit where the strategic decisions referring to an enterprise group are taken. The term UCI, ultimate controlling institutional unit, is equivalent to this term.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Global group head, multinational enterprise group, ultimate controlling unit

Global enterprise

See: Multinational enterprise

Global group head

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. The group head is a parent legal unit which is not controlled either directly or indirectly by any other

legal unit. In case of multinational enterprise groups global and domestic group heads can be identified. The global group head is the group head of the multinational enterprise group, the domestic group head is on the top of the truncated national part of the multinational enterprise group.

Source: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Global decision centre, multinational enterprise group

Government unit

Government units are unique kinds of legal entities established by political processes that have legislative, judicial or executive authority over other institutional units within a given area.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Institutional unit

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 historical and not to be deleted physically.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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

Horizontal integration

Horizontal integration occurs when an activity results in end-products with different characteristics. This could theoretically be interpreted as activities carried out simultaneously using the same factors of production, in which case it would often be impossible to separate such activities statistically into different processes, assign them to different units or generally provide separate data for them, nor would rules relying on allocation of value added or similar measures be applicable. Alternative indicators, such as gross output, might sometimes be applicable, but there is no general rule for identifying the single activity that best represents the mix included in this horizontal

integration. Since patterns of horizontal integration have been considered in the preparation of the classification, in many cases commonly integrated activities are included in the same class of activity even though their outputs have quite different characters.

Source: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Activity, vertical integration

Historical register

Historical register is the version of the register with the capacity to view the content of a live register at some time point in the past, or more generally a set of such versions for a set of register snapshots. A historical register has three possible manifestations:

- First as stated in the text, the set of frozen frames may be said to constitute a historical register.
- Second, the SBR may take measures other than production of frozen frames to generate a historical register
- Third, if all updating transactions in the SBR are time stamped then, in principle, the state of the SBR as of any point in time up to the present can be reconstructed. Thus the live register itself has the capacity to generate an historical register as of any given time point.

Source: International Guidelines for Business Registers

Related terms: Historical data, live register

Indirect control

Indirect control means that a parent unit has control over a sub unit (sub-subsidiary) through one or many other subsidiaries. Indirect control does not require the parent unit to own a majority of an integrated shareholding in the capital share of the sub-subsidiaries.

Source: European Commission, Eurostat, "Foreign Affiliates Statistics (FATS) Recommendations Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2012.

Link: <http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-12->

016

Related terms: Control, ownership, parent corporation

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 International Standard Industrial Classification (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 on similar types of activities.

Source: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Establishment, kind-of-activity

Informal sector

The informal sector is broadly characterised as consisting of units engaged in the production of goods or services with the primary objective of generating employment and incomes to the persons concerned. These units typically operate at a low level of organisation, with little or no division between labour and capital as factors of production and on a small scale. Labour relations – where they exist – are based mostly on casual employment, kinship or personal and social relations rather than contractual arrangements with formal guarantees. The informal sector thus defined excludes households producing exclusively for own final use.

Source: Measuring the Non-Observed Economy: A Handbook, OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, 2002.

Link: <http://www.oecd.org/dataoecd/9/20/1963116.pdf>

Related terms: Institutional sector

Institutional sector

The institutional units are grouped together to form five institutional sectors, on the basis of their principal functions, behaviour and objectives:

- a. 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.
- b. Financial corporations are institutional units which are independent legal entities and market producers that are principally engaged in financial services including financial intermediation.
- c. 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.
- d. 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.

e. 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.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.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: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Institutional sector, standard statistical unit

Inward statistics on foreign affiliates (I-FATS)

Inward statistics on foreign affiliates describe the activity of foreign affiliates resident in the compiling economy.

Source: European Commission, Eurostat, "Foreign Affiliates Statistics (FATS) Recommendations Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2012.

Link: <http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-12->

016

Related terms: Foreign Affiliate Trade in Services, foreign direct investment (FDI) statistics, outward foreign affiliate statistics (O-FATS)

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: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

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 making bodies.
- Enterprises with other forms of legal constitution: This group includes nationalised industries, publicly-owned enterprises and state or local authority monopolies.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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", which is used in the legislation of a number of countries, though not all, 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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 to keep track of changes in, statistical units and their characteristics that occur in the economy. 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 to be 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 communications with the owners of the sources 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: International Guidelines for Business Registers

Related terms: Administrative business register, EuroGroups 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: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Enterprise, statistical unit

Local unit of a legal unit

A local unit of a legal unit is an administrative or operational unit that engages in productive activity at or from one location. Legal, administrative and operational units are the basis for identifying statistical units.

Source: International Guidelines for Business Registers

Related terms: Administrative unit, legal unit, operational unit

Majority ownership

A single investor controls an enterprise by holding a majority (more than 50 %) of the voting power or of the shares, directly or indirectly. While majority ownership is the major criterion for determining control, it is not indispensable for exercising control. A government can exert control through a legislative decree or regulation, empowering it (the government) to determine the enterprise's policy or to appoint (a majority of) directors.

Source: European Commission, Eurostat, "Foreign Affiliates Statistics (FATS) Recommendations Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2012.

Link: <http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-12-016>

Related terms: Control, ownership

Market producer

A corporation that is created for the purpose of producing goods or services for sale on the market at prices that are economically significant is defined as a market producer.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Non-market output, production unit

Master frame

See: Frozen frame

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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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)

Multinational enterprise

A multinational enterprise is an enterprise with cross border links to foreign affiliates.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Enterprise, global enterprise

Multinational enterprise group

A multinational enterprise group is an enterprise group which has at least two enterprises or legal units located in different countries. The reason why both enterprises and legal units are mentioned is that branches which do not constitute separate legal units and are dependent on foreign enterprises are deemed to be (quasi-)enterprises for business register purposes.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: All-resident enterprise group, enterprise group, truncated enterprise group

National Statistical Authority (NSA)

See: National Statistical Institute

National Statistical Institute (NSI)

The national statistical institute is the leading statistical body within a national statistical system that is in charge of producing and disseminating official statistics.

Source: Measuring the Non-Observed Economy: A Handbook, OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, 2002, Annex 2, Glossary

Link: <http://www.oecd.org/dataoecd/9/20/1963116.pdf>

Related terms: National Statistical Authority, National Statistical Institute

National Statistical Office (NSO)

See: National Statistical Institute

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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Legal person

Nature of business survey

See: SBR improvement survey

New enterprise

See: Birth of enterprise, creation (of business)

Non-market output

Non-market output consists of goods and individual or collective services produced by non-profit institutions serving households (NPISHs) or government that are supplied free, or at prices that are not economically significant, to other institutional units or the community as a whole. A price is said to be not economically significant when it has little or no influence on how much the producer is prepared to supply and is expected to have only a marginal influence on the quantities demanded.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Market producer

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;
- 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: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Non-market output

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: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.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: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms Employees, employment, number of employees

Observation unit

A unit about which data are obtained during the course of a survey. Usually a statistical unit, or, if data cannot be obtained about a statistical unit, then some other unit about which data can be obtained and from which data for a statistical unit can be compiled.

Source: International Guidelines for Business Registers

Related terms: Data collection unit, reporting unit, statistical unit

Offshoring

Offshoring is generally defined as companies' purchases of intermediate goods and services from foreign providers at arm's length or the transfer of particular tasks within the firm to a foreign location, i.e. to foreign affiliates.

Source: Organisation for Economic Cooperation and Development (OECD), "Measuring Globalisation: OECD Economic Globalisation Indicators 2010", OECD, 2010.

Link: <http://www.oecd-ilibrary.org/docserver/download/9210031e.pdf?expires=1421077118&id=id&accname=ocid194935&checksum=9AC2705CA7526D789461FA65C002DF58>

Related terms: Outsourcing

Operational unit

A unit defined by a legal unit for the purposes of organising itself, for example a division, branch, workshop, warehouse, or outlet.

Source: International Guidelines for Business Registers

Related terms: Legal unit

Outsourcing

Outsourcing refers to the purchasing of intermediate goods and services from outside specialist providers at arm's length either nationally or internationally.

Source: Organisation for Economic Cooperation and Development (OECD), "Measuring Globalisation: OECD Economic Globalisation Indicators 2010", OECD, 2010.

Link: <http://www.oecd-ilibrary.org/docserver/download/9210031e.pdf?expires=1421077118&id=id&accname=ocid194935&checksum=9AC2705CA7526D789461FA65C002DF58>

Related terms: Offshoring

Outward statistics on foreign affiliates (O-FATS)

Outward statistics on foreign affiliates describe the activity of foreign affiliates abroad controlled by the compiling economy.

Source: European Commission, Eurostat, "Foreign Affiliates Statistics (FATS) Recommendations Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2012.

Link: <http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-12-016>

Related terms: Foreign Affiliate Trade in Services, foreign direct investment (FDI) statistics, inward foreign affiliate statistics (I-FATS)

Ownership

The ownership of a unit or a group of units is related to the property of its assets and determines the distribution of financial flows and income. If a unit or group of units is owned by shareholders, its ownership is vested in the shareholders collectively and can be seen as diffused among the legal units that own its shares in proportion to their shareholdings, and independently of voting rights.

Source: Eurostat, "Business registers. Recommendations manual", Methodologies and Working Papers, Publications Office of the European Union, Luxembourg, 2010

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

Related terms: Control, parent corporation, subsidiary

Parent corporation

See: Control

Principal activity

Principal and secondary activities cannot be carried out without the support of a number of ancillary activities, such as bookkeeping, transportation, storage, purchasing, sales promotion, cleaning, repair and maintenance, security etc. At least some of these activities are found in every economic entity. Thus, ancillary activities are those that are undertaken to support the main productive activities of an unit by providing goods or services entirely or primarily for the use of the respective entity.

Source: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Activity, economic activity, ancillary activity, secondary activity

Producer

See: Production unit

Producing unit

See: Production unit

Product

Products are goods and services (including knowledge-capturing products) that result from a process of production. They are exchanged and used for various purposes: as inputs in the production of other goods and services, as final consumption or for investment. The SNA makes a conceptual distinction between market, own final use and non-market goods and services, allowing in principle any kind of good or service to be any of these three types.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009"

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Non-market outcome, production unit

Production unit

Enterprises are likely to be very heterogeneous as they may 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. 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.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009"

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Economic unit, institutional units, market producer, product, statistical unit

Profiling

Method to analyse the legal, operational and accounting structure of an enterprise group at national and world level, in order to establish the statistical units within that group, their links, and the most efficient structures for the collection of statistical data.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Delineation

Property

See: Characteristic

Quality improvement survey

See: SBR improvement survey

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. The enterprise is considered to have survived as long as the period of inactivity does not encompass a whole calendar year.

Source: Eurostat and Organization for Economic Cooperation and Development (OECD), "Eurostat - OECD Manual on Business Demography Statistics (Edition 2007)", Methodologies and Working Papers, Office for Official Publications of the European Communities, Luxembourg, 2007

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

Related terms: Dormant unit

Real world units

Units that exist in the "real world" "i.e., the world outside the NSI (as distinct from standard statistical units that are NSI constructs used to model real world economic units in a standard way across within countries and across countries).

Source: International Guidelines for Business Registers

Related terms: Data collection unit, reporting unit, standard statistical unit

Register picture

See: Register snapshot

Register snapshot

A copy of a live register (or the statistical units in the live register) as of a given point in time

Source: International Guidelines for Business Registers

Related terms: Live register

Reporting unit

Unit from which data about an observation unit are obtained during the course of a survey. It may, or may not, be the same as the observation unit. An example where it is not the same is where an accounting business reports data on behalf of a client business that is the actual subject of the survey.

Source: International Guidelines for Business Registers

Related terms: Data collection unit, observation unit

Restructuring

Restructuring within an enterprise does not affect the continuity of the enterprise, but changes its structure in the process. 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 could affect the way the enterprise is included in demographic statistics. Restructuring will be reflected through changes to relationships or characteristics recorded in the register. 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, involving many enterprises and possibly, but not necessarily, entailing a change in the number of enterprises of the group. Complex restructuring is a similar event, but this is not constrained to one enterprise group. An example is the transfer of a number of enterprises or parts of enterprises between groups. Restructuring within an enterprise group, or complex restructuring, may entail any number of register creations and deletions.

Source: Eurostat and Organization for Economic Cooperation and Development (OECD), "Eurostat - OECD Manual on Business Demography Statistics (Edition 2007)", Methodologies and Working Papers, Office for Official Publications of the European Communities, Luxembourg, 2007

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

Related terms: Continuity, survival

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: International Guidelines for Business Registers

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: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Activity, economic activity, ancillary activity, principal activity

Small and medium-sized enterprises

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 new definition provides for an increase in 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

Snapshot

The snapshot is a copy of all the statistical units in the live register, and all administrative units or at least all links from statistical units to administrative units. The snapshot is an intermediary step between live register and 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: International Guidelines for Business Registers

Related terms: Frame, frozen frame

Special purpose entity (SPE)

There is no common definition of an SPE but some of the following characteristics may apply. Such units 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 SPEs in particular are often resident in a territory other than the territory of residence of the related corporations. In the absence of any physical dimension to an enterprise, 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 any other institutional unit 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: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 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 and Organization for Economic Cooperation and Development (OECD), "Eurostat - OECD Manual on Business Demography Statistics (Edition 2007)", Methodologies and Working Papers, Office for Official Publications of the European Communities, Luxembourg, 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 units

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 units are enterprise, enterprise group, kind-of-activity unit (KAU), local unit, establishment and homogeneous unit of production that are delineated according their legal, accounting or organisational structure, their location and their activity.

Source: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Legal unit, institutional unit

Statistical Business Register (SBR)

The statistical business register is a fully 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: Administrative business register

SBR survey

See: SBR improvement survey

SBR coverage survey

See: SBR improvement survey

SBR improvement survey

A survey conducted by statisticians to improve SBR quality.

Source: International Guidelines for Business Registers

Related terms: Statistical business register

SBR quality improvement survey

See: SBR improvement survey

Statistical compiler

A statistician involved in producing statistics (aggregates) from individual data or other aggregate

Source: International Guidelines for Business Registers

Related terms: Statistician, survey statistician

Statistical unit

Unit defined by an NSI for statistical purposes; unit for which information is sought and for which statistics are ultimately compiled.

Source: International Guidelines for Business Registers

Related terms: Administrative unit, observation unit

Statistician

A statistician involved in the conduct of a survey

Source: International Guidelines for Business Registers

Related terms: Statistical compiler, survey statistician

Subcontracting

See: Outsourcing

Subsidiary

A subsidiary is a company controlled by another company. A subsidiary is a corporation in which over 50 per cent of the voting power is held by their respective parents.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009.

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Branch, control, ownership

Survey statistician

Anyone involved in statistical activities

Source: International Guidelines for Business Registers

Related terms: Statistical compiler, survey statistician

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 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). 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: European Commission (Eurostat), Organisation for Economic Cooperation and Development (OECD), "Eurostat – OECD Manual on Business Demography Statistics", Methodologies and Working papers, Publication Office of the European Communities, Luxembourg, 2007.

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

Related terms: Business demography, continuity

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 strict 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 such items as gross domestic product (GDP), the most frequently quoted indicator of economic performance. The accounting framework of the SNA allows economic data to be compiled and presented in a format that is designed for purposes of economic analysis, decision-taking and policymaking. The accounts themselves present in a condensed way a great 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, that takes place on markets or elsewhere. The framework of the SNA provides accounts that are:

- comprehensive, in that all designated activities and the consequences for all agents in an economy are covered;
- consistent, because identical values are used to establish the consequences of a single action on all parties concerned using the same accounting rules;
- integrated, in that all the consequences of a single action by one agent are necessarily reflected in the resulting accounts, including the impact on measurement of wealth captured in balance sheets.

The accounts of the SNA provide more than a snapshot of the economy at a point in time, since in practice the accounts are compiled for a succession of time periods, thus providing a continuing flow of information that is indispensable for the monitoring, analysis and evaluation of the performance of an economy over time. The SNA provides information not only about economic activities taking place within a period but also about the levels of an economy's assets and liabilities, and thus the wealth of its inhabitants, at particular points of time. In addition, the SNA includes an external account that displays the links between an economy and the rest of the world. The SNA is designed for economic analysis, decision-taking and peacemaking, whatever the industrial structure or stage of economic development reached by a country. The basic concepts and definitions of the SNA depend upon economic reasoning and principles which should be universally valid and invariant to the particular economic circumstances in which they are applied. Similarly, the classifications and accounting rules are meant to be universally applicable. The System of National Accounts (SNA) has been prepared under the joint responsibility of the United Nations, the International Monetary Fund, the Commission of the European Communities, the OECD and the World Bank. The SNA is designed to give a realistic and compact view of the economy that is suitable for policy and analytical use.

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 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: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 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: European Commission (Eurostat), Organisation for Economic Cooperation and Development (OECD), "Eurostat – OECD Manual on Business Demography Statistics", Methodologies and Working papers, Publication Office of the European Communities, Luxembourg, 2007.

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

Related terms: Cessation (of business), death (of enterprise group), merger,

Truncated enterprise group

A truncated enterprise group consists of the enterprises and the legal units of a multinational enterprise group, which are resident in the same country. It may comprise only one unit, if the other units are non-resident. An enterprise may be the truncated group or part thereof. A truncated group may consist of several units and subgroups, which can appear seemingly unlinked if their parent is non-resident, but actually belongs to the same multinational enterprise.

Source: European Commission, Eurostat, "Business Registers and Recommendation Manual", Methodologies and Working papers, Publication Office of the European Union, Luxembourg, 2010.

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

Related terms: All-resident enterprise group, enterprise group, multinational enterprise group

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: European Union, Commission Regulation (EC) No 2700/98 of 17 December 1998 concerning the definitions of characteristics for structural business statistics Indices of Distributive Trade: Handbook on Good Practices, United Nations, 2010, pag.34

Link: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1998:344:0049:0080:EN:PDF>

Related terms: Administrative business register

Ultimate controlling institutional unit (UCI)

See: Global decision centre

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 (SNA 1993, § 4.140 and 4.141).

Source: European Commission (Eurostat), International Monetary Fund (IMF), Organisation for Economic Cooperation and Development (OECD), United Nations (Statistics Division), World Bank, "System of National Accounts 2008", United Nations, New York, 2009

Link: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Related terms: Enterprise

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.

Source: International Guidelines for Business Registers

Related terms: Characteristic

Vertical integration

The vertical integration of activities occurs wherever the different stages of production are carried out in succession by the same unit and the output of one process serves as input to the next, any activities that cannot be separated at the level of the statistical unit.

Source: United Nations, Statistics Division, "International Standard Industrial Classification of all Economic Activities (ISIC)", Statistical Papers Series M No. 4, Rev. 4, New York, 2008.

Link: <http://unstats.un.org/unsd/cr/registry/isic-4.asp>

Related terms: Activity, horizontal integration

Annex C: Characteristics of Units by Unit Type

C1: Enterprise Group (including Truncated Enterprise Group) - Characteristics

The characteristics listed in the following paragraphs apply equally to a truncated enterprise group that is the national part of multi-national enterprise group as to an enterprise group that is entirely within national boundaries.

C1.1: Enterprise Group - Identification characteristics

Enterprise Group - Identity number

- Purpose** To identify all enterprise groups (all-resident or truncated parts of international enterprise groups) and to follow their continuity.
- Definition** The identity number of an enterprise group is (usually) given nationally in the SBR. It may be common with other national institutions.
- As the enterprise groups are permanently buying, selling or restructuring their affiliates, the question of whether the enterprise group is the same after restructuring than before is a prominent one. General continuity rules apply for all types of enterprise groups, as discussed in Section 7.3.4.
- The continuity of a truncated group is based on the continuity of the global group of which it is part. The question of whether a truncated group can be continuous, even if the global group is not and even if it is split from the global group or sold to another one, remains an issue.
- Sources** The identity number of an enterprise group can be given internally in the SBR or be derived from an external (administrative) source.
- Comments** As the continuity of an enterprise group may remain while the global group head changes, it is not advisable to use the identity number of the global group head as the identity number of the enterprise group.

Enterprise Group Name

Enterprise Group - Address

Enterprise Group - Telephone and fax numbers, electronic mail address and information to permit electronic collection of data (optional)

- Purpose** The storing of the name and address information is important for survey and profiling purposes and may have other national uses.
- The enterprise group (through its head office services) is often the reporting unit for itself and even for data about other statistical units belonging to the group (e.g., enterprises).
- Definition** National standards for the name and address are used for head office, whether it be an all-resident or truncated enterprise group.
- Sources** The name may be given nationally in the SBR or it may be common with other institutions (for example, from the EU perspective). Administrative

sources, surveys, profiling, internet, and SBR information can be used.

Comments The name of the truncated group should not be exactly the same as the name of the global group. It is up to the country to decide on the information it wishes to record for each unit and address. In some countries, at legal level the consolidated accounts may use the legal name of the group head when they refer to the whole group.

Enterprise Group - Identity number of the group head

Purpose To identify the global group head (all-resident or truncated group head).

Definition The identity number of the legal unit which is the resident group head.

Proxies For a natural person who is not registered in the SBR as economic operator or for non-resident group head, a quasi-unit with an identity number can be created when necessary. In this case, it should be possible to identify natural persons separately.

Sources Administrative sources and control links in the SBR.

Comments If the controlling unit is a natural person who is not an economic operator, recording is subject to the availability of this information from an administrative source. In some countries recording is subject to legal limitations or banned; in other countries there is no problem.

Nevertheless, the *OECD Handbook on Economic Globalisation Indicators* recommends the inclusion of natural persons in these cases

The number and importance of natural persons as controlling units may depend on national legislation and thus vary considerably between countries. It would be useful to study the effect of their inclusion/exclusion on the consistency of data between countries.

This characteristic is relevant only if the global group head (GGH) is a resident unit and if the unit is not a natural person. If not, a link should be made with to an international register, if one exists.

The dates of commencement and cessation as resident group head must also be stored, either when the change is recorded in the SBR or preferably as separate variables when the real change takes place (if this is known).

Enterprise Group - Type of enterprise group

Purpose The type is an important stratification characteristic from an economic point of view: The importance of many other characteristics depends on the type.

It allows calculation of many economic indices at national level by type, such as proportion of different group types (and independent enterprises) as regards employment, turnover per person employed, etc. The units belonging to foreign controlled truncated groups define the foreign controlled inward FATS population.

Definition The difference between an all-resident and a multinational group depends purely on whether controlled affiliates exist in only one country or in more than one country.

The difference between domestically and foreign controlled multinational groups is determined by whether the global decision-centre managing unit is resident in the country being considered or in another country.

Sources Administrative and commercial sources, statistical or administrative surveys, the Euro Groups register.

Comments Types of enterprise group:
1. All-resident group;
2. Domestically controlled truncated group;
3. Foreign controlled truncated group.

Although enterprise group type may be inferred from other characteristics it is advisable to record it separately. This is especially the case if much of the national enterprise group information is in a satellite register where the coverage may vary according to the group type.

C1.2 Enterprise Group - Demographic characteristics

Enterprise Group - Date of commencement in the country

Purpose The date is needed for the demography of enterprise groups.

Definition The date refers either to a date when a new enterprise group is born or to other creation date of a new group (by merger, break-up, split-off or restructuring).

The date refers to the beginning of operation for a truncated group: this date can be the same as the creation date of the first affiliate in the country; but more often it refers to the date when the first affiliate is controlled (bought).

Proxies The birth of a new group may be difficult to define in practice, if the smallest groups of no statistical importance to the country are not monitored. The date from which the group is being monitored is then used as a proxy. However, the approximate dates are important in order to determine from which year a certain multinational group is monitored in different countries.

Sources Administrative sources, surveys.

Comments The creation may result from a restructuring process.

Enterprise Group - Date of cessation in the country

Purpose The date is required for the demography of enterprise groups.

Definition Cessation of a group means either death of the group in the country (dissolution of the links of control between the units belonging to the group), or (more commonly) merger with or takeover by another group, break-up, split-off or restructuring into two or more groups.

The cessation of a truncated group is registered separately from that of the global group it is part of, because the previous demographic events can have different consequences for the global group and for its truncated part(s).

Thus the date of cessation of any kind of group is very closely dependent on

the continuity rules used.

- Proxies If no exact date is known for truncated groups, the approximate date is important in order to know the situation of multinational groups in different countries.
- Sources Administrative sources, surveys.
- Comments This characteristic is recorded when the group ceases to exist or is not monitored anymore. In such cases, all the characteristics become historical.

C1.3: Enterprise Group – Economic/stratification characteristics

Enterprise Group - Principal activity code at ISIC 2-digit level

Enterprise Group – Secondary activities at ISIC 2-digit level (non-core)

- Purpose These activity codes can be used for stratification, demographic and economic analysis. Secondary activities are important for observing the homogeneity of the group and, in cases where the secondary activity rather than the primary activity is the subject of a survey.
- Definition The activities are defined according to the ISIC rules. The principal activity is identified by the top-down method as the activity which contributes most to the total value added.
The same definition is applied at the global level as at the national truncated level.
- Proxies If value added data are not available, it is recommended that employment be used as the criterion. The activities performed can be based on the activity codes and number of persons employed in the units belonging to the group at national level. Turnover is less suitable because it is more sensitive to intra-group transactions.
- Sources SBR derivation from its legal units, administrative sources, and surveys.
- Comments The main activity of any of the truncated parts may be different from the main activity of the global group.
Principal and secondary economic activities in the country can be inferred from the economic activities of the enterprises composing the group.
It is recommended that the principal activity be recorded more precisely, if possible at ISIC 3-digit level.

Enterprise Group - Number of persons employed in the all-resident/truncated groups

- Purpose Employment can be used for stratification, demographic and economic analysis. Together with activity code, it provides information on the role of the group in the global framework.
- Definition The persons employed in the group are added up from the units which belong to the enterprise group.
This procedure is valid at truncated level and at global level (provided that a global register of affiliates exist).

Proxies	The number of employees.
Sources	The SBR.
Comments	For a large multinational operating in a large number of countries, a better source for employment in foreign countries might be a specific “note” included in the Annual Report.

Enterprise Group - Consolidated turnover (non-core)

Purpose	Turnover is an important size indicator for stratification. The availability of consolidated turnover at global group, at truncated or at or subgroup level may be useful in profiling.
Definition	Turnover is defined as the revenues gained by the sales of goods and services. For consolidation rules (full consolidation) the International Accounting/ Financial Reporting Standards (GAAPS) should be followed.
Proxies	There can be national specifications of the consolidation rules.
Sources	Administrative sources, surveys, profiling.
Comments	<p>It is difficult to consolidate turnover of different enterprises in different activities and the consolidation rules are not the same everywhere. Also turnover from different sources may not be consistent.</p> <p>It may also be possible to aggregate the turnover of legal units belonging to the enterprise group, but this would include internal turnover which does not correspond to market prices when based on internal prices. Thus storing of any figure other than the consolidated turnover figure may be confusing and is not recommended. The best source for consolidated turnover of an enterprise group is their Annual Report.</p>

Enterprise Group - Country of global decision centre (non-core)

Purpose	To allocate the enterprise group to the country where decisions on its global strategy are actually taken. Foreign affiliate statistics use the resident country of the global group head.
Definition	Country where the strategic decisions referring to an enterprise group are taken.
Proxies	In most of the cases the global decision centre equals to the global group head. In these cases the country of global decision centre is the country where the global group head is located. In special cases the global group head is not able to act as a decision centre for the whole group. Units without substantial physical presence, significant economic activities, or employment are not able to be engaged in decisions. In these cases the global decision centre should be appointed at the next level downwards in the enterprise group tree structure, where the global decisions of the enterprise group are taken.
Sources	Administrative and commercial sources; statistical surveys. (In the case of the EU the Eurogroup Register provides standardised information on the decision centre of the multi-national enterprises covered.)
Comments	The country where the decision-making centre is located is often referred to

as the *nationality of the group*, although nationality can be understood in other ways, for instance to refer to the nationality of the owner. The roles of the group head and decision centre may vary a great deal.

This cannot be taken into account in an operational rule, but can be established during profiling. To gain further knowledge about how the group operates, the group can be further analysed. Factors to be considered include where the main production activity is located, where value added is created, and where research and development (which influences the future strategy of the group) takes place.

These characteristics are relevant when allocating the enterprise group to the country where decisions on its global strategy are actually taken. Foreign affiliate statistics use the resident country of the global decision centre (GDC) for defining the outward FATS population (the resident GDC is the reporting unit on its foreign affiliates) and the inward FATS population (all enterprises which have a non-resident GDC are surveyed).

Coordination with the NSIs of the other countries where the multinational enterprise group is acting is recommended in order to agree on the GDC and the global economic characteristics of the group for producing international consistent statistics.

Enterprise Group - Countries where enterprise group has enterprise(s) or local unit(s) non-core)

Purpose	The characteristic is essential information for outward FATS and for studying intragroup trade, which also needs the names, addresses and identity numbers of legal units belonging to the group.
Definition	Country codes (according to the §4.13 of the Balance of Payments Vademecum).
Sources	Administrative (consolidated balance sheets) and commercial sources, the EuroGroups register.

C1.4 Enterprise group - Relationships between units

(No characteristics.)

C1.5 Enterprise group - Relationship with other registers/update sources

Purpose There can be two kinds of links:

- links with sources of information and data on enterprise groups, mostly either private providers or specific administrative registers or enquiries; several examples show that the concerned administrative enquiries might be managed by statistical offices under specific non-statistical regulations;
- links with other countries relating to multinational enterprise group (as exemplified in the following box).

EU Example of relationship

The example concerns the EU (including EFTA) member states.

- The identity number of a multinational enterprise group comes from the EuroGroups Register (EGR), a supra-national register.
- The ID number of a truncated group is derived from the common ID of the multinational enterprise group and thus given centrally by an EGR procedure.
- The identity number should remain the same so long as the group is considered to be continuing. (Continuity is discussed in Chapter 4 of these Guidelines and in Chapter 22 of the BR Recommendation manual.)
- The global group head and/or the main decision centre may be moved from one country to another with the enterprise group. Thus it is not advisable to give an ID that includes a country code.

Proxies: National numbers can still be used until a system of unique Europe-wide identity numbers is established. From then on, a second identity number may be useful in order to be able to trace the continuity of all-resident groups if they become multinational.

Remark: In principle this approach could be used by any collection of countries that agreed to co-operate on handling multinational enterprises. It would need to be supported by legal rights to exchange information, which would likely require the existence of some kind of integrated economic union. In practice the only known collection of countries where the approach is being applied is the EU.

C2 Enterprise (and Truncated Enterprise) - Characteristics

The characteristics listed in the following paragraphs apply equally to a truncated enterprise that is the national part of multi-national (global) enterprise as to any other kind of an enterprise.

C2.1 Enterprise - Identification characteristics

Enterprise - Identity number

Purpose To identify the unit and to follow its continuity and the demographic events leading to discontinuity.

Definition Given nationally in the SBR.

As the continuity rules for enterprises should be applied, the identity number should remain the same while the enterprise is considered as continuing.

Sources Assigned by SBR procedures.

Comments In case of a 1:1 relation between the legal unit and enterprise, the same identity number is often used but this is not recommended as, conceptually, they are different units.

Enterprise - Name

Enterprise - Address

Enterprise - Non-core: Postal, electronic mail and (Non-core) website addresses

Purpose To provide contact information. The website can also be used for statistical purposes.

Definition Countries can decide what information is required to meet their needs.

Proxies The enterprise name may be the same as the legal unit name, but it may also differ (i.e., be a trade name) even the enterprise comprises only one legal unit. An enterprise often uses the name of the main legal unit that operates it.

Sources Administrative sources, websites, SBR procedures, surveys.

Enterprise - Identification number(s) of legal unit(s) of which the enterprise consists

Purpose To enable the enterprise to be linked to its constitutive legal unit, or legal units if more than one.

Definition This is simple for enterprises consisting of one legal unit only. In the case of complex enterprises, the identity numbers of all legal units must be recorded.

Proxies None.

Sources Created by SBR procedures.

Comments It is recommended that the dates when the links are created and (possibly) dissolved in the SBR are recorded as well as the reference dates in the real world.

C2.2 Enterprise - Demographic characteristics

Enterprise - Date of commencement of economic activities in the country

Purpose The date is required for monitoring the demography of enterprises.

Definition The date refers to the economic birth of the enterprise.

Proxies As explained in Section 5.3, the date of birth is, in principle, the date on which the first financial commitments are made. In practice it may refer to the registration date in the administrative source if the unit starts its economic activities immediately after that.

Sources Administrative sources, surveys.

Comments It is necessary to register the relation with legal units.

The legal unit(s) associated with an enterprise may change and be re-registered, for example after a change of legal form, while the enterprise continues, i.e., remains the same.

Enterprise - Date of permanent cessation of activities in the country

- Purpose** The date is required for monitoring the demography of enterprises.
- Definition** The date refers to the date when an enterprise becomes permanently inactive or dead.
- Proxies** As explained for the corresponding characteristic of legal units, this date may not be available with any precision. Only the fact that the enterprise has ceased to exist or has ceased its activities during the reference year may be known, in which case the date should be estimated.
- Sources** Surveys, administrative sources or defined by SBR procedures.
- Comments** The relationship to legal units needs to be registered.

C2.3 Enterprise- Stratification characteristics⁸⁰

Enterprise - Principal activity code at ISIC 4-digit level

- Purpose** The principal activity code is a key stratification characteristic. The code is also very important in judging the role of certain units in the enterprise group structure, mainly the group head, the decision centre and special purpose entities, and in general in enterprise group structuring and demography. The enterprise group's worldwide dispersion of employment by activity can also be studied through its constituent units.
- Definition** The activity code is derived according to the rules in the ISIC Rev 4.
- Proxies** Other criteria can be used to define the principal activity code, if value added is not available. Employment (preferably) and turnover (with the drawback that it depends on from where it is measured) broken down by the different activities can be used, also activity descriptions.
- Sources** Surveys and administrative sources.
- Comments** If the enterprise has a legal capacity to operate in different trades, data from administrative sources may indicate a range of activities, but quantitative information on their relative importance may not be available. In such a case the data has to be obtained by survey.
- If the enterprise is linked to just one local unit, the principal activity for both types of unit should be the same.

⁸⁰ Enterprises are defined globally (which means at world level when adequate). For establishing national level, they may be truncated to the national territory of each relevant country. In the case of truncated enterprise, double coding of some characteristics (in particular principal activity code) at national and supra-national level is recommended.

Enterprise - Conditional: Secondary activities, if any, at ISIC 4-digit level

Purpose	Knowledge of secondary activities is very important for large enterprises, for example for short-term business surveys and for national account purposes.
Definition	The activity codes are derived according to the rules in the ISIC Rev. 4.
Proxies	Several criteria can be used for defining the activity code(s), if value added is not available. Employment and turnover by different activities can be used (when available), also activity descriptions.
Sources	Surveys (more accurately) or administrative sources.
Comments	The values of this characteristic may lead to delineation of establishments belonging to the enterprises.

Enterprise - Number of persons employed

Enterprise - Number of employees (non-core)

Enterprise - Number of employees in FTE (non-core)

(These characteristics are calculated in accordance with the same rules as for local units, establishments or KAUs.)

Purpose	<p>Many countries are interested in breakdowns by size class, as small and medium size enterprises (SMEs) are often considered to be a major source of new employment. The precise definition of SME has to be decided by each country⁸¹. The share of SMEs and within SMEs, the small and micro enterprises can be calculated on the basis of this characteristic.</p> <p>Employment within the enterprises that comprise an enterprise group may be used to determining the continuity or otherwise of the enterprise group.</p> <p>An enterprise group's worldwide dispersion of employment by activity, and the impact of offshoring, can also be studied by through the employment of its constituent enterprises.</p>
Definition	<p><i>Number of persons employed</i> 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.</p> <p><i>Number of employees</i> 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</p>

⁸¹ For example, in the EU, according to Commission Recommendation 2003/361/EC they consist of independent enterprises below 250 persons employed full-time

of that employment agency and not of the units in which he/she (temporarily) works.

Sources Surveys or administrative sources. Social security sources may be used as sources of information.

Comments As an enterprise consists of one or more local units, the employment of an enterprise as of a given date should equal the sum of the employment for the local units. The same holds true for establishments and KAUs belonging to the enterprise.

SBR employment values should be compared with those of other sources (structural business statistics, business demography, labour force survey) for consistency. Of course, the values may differ if different concepts, definitions or methodologies are used.

Enterprise - Turnover

Purpose For some surveys, mainly repeated ones, and for some spheres of activity, it may not be appropriate to stratify according to employment. Moreover, for accurate calculation purposes, the size of enterprises should also be measured in terms of their turnover.

Turnover is not an additive measure which restricts its direct (non-consolidated) use for the enterprises in the enterprise group framework. However, it can be used for group-related research and development intensity calculations.

Definition Turnover consists of all revenues arising during the specified reference period, in the course of ordinary activities of the statistical unit and is presented net of all price reductions, discounts and rebates granted by it.

The revenue referred to is arising from contracts with customers and are realized through the satisfaction by the statistical unit of the contractual performance obligations usually represented by the sale (transfer) of goods or the rendering of services, however, it can also contain revenues obtained as a yield on the use by others of the statistical unit's assets.

Excluded from net turnover are:

- all taxes, duties or levies linked directly to the output;
- any amounts collected on behalf of any principal, if the statistical unit is acting as an agent in its relationship with said principal; turnover includes only the service charge for acting as an agent;
- all income not arising in the course of ordinary activities of the statistical unit. Usually, these types of income are classified as "Other income", "Financial income", "Extra-ordinary income" or under a similar heading, depending on the respective set of generally accepted accounting standards used to prepare the financial statements.

Proxies Fiscal sources may be used as sources of information, for example, corporate income tax or VAT returns. There may be cases where such information is unavailable, for example where units are not subject to VAT. In such a situation one possibility is to estimate turnover based on employment using a standard turnover per head ratio calculated by the ISIC class.

Turnover is relevant for market units; it should not be estimated for non-market units. However, non-market units can have secondary production.

Sources Surveys and administrative sources.

Comments This turnover is "net" in the sense that it excludes the price reductions, discounts and rebates.

The figure used should be the actual turnover. Turnover may be available only with some delay, especially for newly created enterprises. This may cause problems, mainly for business demography, and it may be necessary to estimate the turnover in such cases.

As turnover from different sources is not always the same, using different sources may lead to inconsistencies.

Turnover is usually measured excluding VAT.

Turnover is an essential piece of accounting data. Even if comparing turnover between enterprises across different ISIC sections may not make much sense, it may be the only parameter that permits a breakdown by product or allows exports to be ascertained.

Enterprise - Institutional sector and sub-sector according to SNA2008

Purpose The institutional sector classification in an SBR has several important functions, as noted in Chapters 3 and 4.

Definition The classification of institutional sector applies to the institutional unit which includes also the enterprise.

The classification is detailed in Annex D.

Proxies From the information available, it may not always be possible to define the sector or sub-sector accurately. In such cases, large and important units could be checked manually. A proxy can be derived from ISIC code, legal form and country of global decision centre.

Sources Defined by SBR procedures in cooperation with national accounts, according to established rules, based on SBR information and using administrative sources when necessary.

Comments In the vast majority of cases, it is possible to calculate the institutional sector code from other explicitly recorded characteristics, especially legal form, ISIC code and controlling country. Algorithms can be used for this purpose. In certain cases the institutional sector must not conflict with the principal activity and whether the enterprise is foreign controlled must be consistent with the respective characteristics of the enterprise group to which it belongs.

The principal activity of the enterprise separates the financial and non-financial corporate enterprises. Two criteria can be applied to separate corporate and quasi-corporate enterprises from enterprises attached to the household sector: The legal form of the legal unit operating the enterprise and, if the legal unit is a natural person, whether or not it keeps annual accounts for the assessment of its tax liability. It may therefore be appropriate to provide an intermediate code indicating, in the case of enterprises operated by a natural person, if they keep accounts or not. That code might perhaps

record whether the enterprise is taxed on profits on the basis of its accounts (actual profits) or on the basis of an assessment.

Business statistics cover only market activities and a code indicating whether the activity of an enterprise is market or non-market is essential for them. The distinction is sometimes difficult to make and can be definitively determined only by looking at the accounts. Market enterprises cover institutional sectors S.11, S.12, S.141 and S.142.

Public undertaking is defined as any undertaking over which the public authorities may exercise directly or indirectly a dominant influence by virtue of their ownership of it, their financial participation therein, or the rules that govern it. A dominant influence on the part of the public authorities shall be presumed when these authorities, directly or indirectly:

- hold the major part of the undertaking's subscribed capital; or
- control the majority of the votes attaching to shares issued by the undertakings; or
- can appoint more than half of the members of the undertaking's administrative, managerial or supervisory body.

This could be one legal form or it could be obtained from an administrative source. It may be interpreted in different ways in different administrations. The key issue is that public authorities control public undertakings. The definition does not clearly specify whether public undertaking should be market-oriented and distinguishing between public undertaking and government unit is sometimes difficult and may require the information from its accounts (whether more or less than 50 % of production costs are met by sales). The precise method to be used for separating public enterprises is to define them by the sum of institutional sub-sectors: S.11001+S.12201+S.12301+S.12401+S.12501.

C2.4 Enterprise - Relationships between units

Enterprise - Identification number of the all-resident/truncated enterprise group to which the enterprise belongs

Purpose	To identify the enterprises belonging to an enterprise group, for example when looking at size classes, in order to avoid confusion between independent enterprises and enterprises belonging to enterprise groups. It is important also for the enterprise confidentiality treatment, for example when all enterprises in a cell belong to the same enterprise group.
Sources	SBR procedures
Comments	The date when enterprise becomes part of the group should also be recorded.

Enterprise- Relationship with other registers/update sources

Purpose	(See Annex C1.5.)
---------	-------------------

C3 Establishment - Characteristics

An *establishment* is also known as a *local kind-of-activity unit (LKAU)*.

C3.1 Establishment - Identification characteristics

Establishment - Identification number

- Purpose** To identify a unit and to be able to record its continuity.
- Definition** It is recommended that the SBR registers the identification numbers of establishments in accordance with the advice given in Section 11.4.11. The identification number of an establishment should remain the same while the unit continues (according to the continuity rules defined in Chapter 6) even when the enterprise to which it belongs changes.
- Proxies** Using a unique administrative identity number (if it exists) is also possible, but may cause problems with continuity and thus is not recommended.
- Sources** SBR procedure, unless an administrative number is used.

Establishment - Name

Establishment - Address

Establishment - Telephone and fax numbers, electronic mail address and information to permit electronic collection of data

- Purpose** To contact the unit.
- Definition** The official name of an establishment is generally the same as the legal unit that controls it, with some additional part (usually) specifying location or activity. If there is only one establishment belonging to the legal unit, a separate name may not exist.
- The actual address of the location of the unit and the contact information should always be recorded.
- Proxies** Different establishments within a legal unit may use different *trading names* (also known as *signboard names* or *commercial names*) and provision should be made for recording these different trading names where they exist.
- Sources** Administrative sources and surveys.
- Comments** The address could be given a code referring to a national territorial nomenclature. It would be preferable if that nomenclature or national geographical code enables the unit to be pinpointed as accurately as possible, at the level of the street, section of street and building.
- In addition to the physical address of the establishment it is useful to provide for the possibility of recording a correspondence address where different.

Establishment - Identity number of the enterprise to which it belongs

- Purpose** The establishment must be linked with the enterprise to which it belongs.

Definition	This link can be included in the register by adding the identity number of the enterprise to the Establishment file (and vice versa).
Proxies	Other ways are conceivable, for example when the enterprise consists of one establishment only, another (simple) arrangement of the business register is possible.
Sources	SBR procedures.

C3.2 Establishment - Demographic characteristics

Establishment - Date of commencement of activities (economic birth)

Purpose	The date is needed for monitoring the demography of establishment.
Definition	This date should refer to the birth or other creation date of the unit according to the continuity rules.

Establishment - Date of final cessation of activities (economic death)

Purpose	The date is needed for monitoring the demography of establishments.
Definition	The date refers to the death of a unit.
Proxies	As explained for the corresponding characteristic for legal unit, this date may not be available with any precision. Only the fact that the establishment has ceased to exist or has ceased its activities during the reference year may be known, in which case the date should be estimated.
Sources	Survey, administrative source and SBR procedures.

C3.3 Establishment – Economic/stratification characteristics

Establishment - Principal activity code at ISIC 4-digit level

Purpose	The principal activity code is a stratification characteristic and important in the compilation of regional and small area statistics.
Definition	The activity code is determined according to the rules in the ISIC Rev. 4 handbook. The activities actually pursued in the unit are taken into account. They may not coincide the principle activity of the enterprise.
Proxies	Several criteria can be used for defining the principal activity code. Employment by different activities (if available) can be used, also activity descriptions.
Sources	The code can be obtained from surveys or administrative sources.
Comments	If the activities are ancillary in the context of the enterprise, the value recorded should indicate this.

Establishment - Secondary activities, if any, at ISIC 4-digit level

Purpose	Secondary activities may be used in determining establishments that are
---------	---

	of key importance for national/regional accounts.
Definition	Activity codes are allocated according to ISIC Rev 4.
Proxies	Several criteria can be used for defining secondary activity code(s). Employment by different activities can be used (if available), but also activity descriptions, etc.
Sources	Surveys and administrative sources.

Establishment - Activity carried out in the unit constituting an ancillary activity of the enterprise to which it belongs (yes/no) (non-core)

Purpose	To distinguish ancillary units. This characteristic enables statistical analyses to reallocate the cost of ancillary activities to the activities for the benefit of which they are undertaken.
Definition	An ancillary activity is incidental to the main activity of an enterprise. It facilitates the efficient running of the enterprise but does not normally result in goods and services that can be marketed. For enterprises that are relatively small and have only a single location, ancillary activities are not separately identified. For larger enterprises with multiple locations, it may be useful to treat ancillary activities in the same way as a secondary or even a principal product ⁸² .
Proxies	Activities in certain ISIC classes often constitute ancillary activities.
Sources	Administrative sources and surveys.
Comments	If a unit has been identified as an ancillary unit, this fact should be positively indicated in the SBR.

Establishment - Number of persons employed

Establishment - Number of employees (non-core)

Purpose	<p>The SBR should record the actual numbers of persons employed and employees, the latter both as a head count and as a FTE.</p> <p>The main aim is to provide a stratification characteristic. Persons employed are preferable for stratifying survey samples for very small units.</p> <p>In addition, SBR employment figures are sometimes disseminated directly, especially as small area statistics where the SBR may be the only comprehensive source.</p>
Definition	<p>For stratification purposes, the SBR aim is to provide end of year values (including seasonally active units). As the end date approach is not harmonised across countries, the annual average can also be used.</p> <p>The number of employees as a FTE may be calculated for a full year as well as for the period in which an enterprise is active. If the number of employees</p>

⁸² SNA2008: 5.10

in FTE is used as stratification characteristic, it should be calculated as an average over the active period, whilst in compiling annual statistics it should be calculated as the average over the full 12 months.

0 means less than half a person, whether calculated as a head count or FTE

Proxies Number employed can be obtained directly in some countries in administrative sources, while in other countries only an administrative source providing number of *paid employees* is available. However, number employed can be calculated from data on paid employees by adding working proprietors and unpaid employment calculated according to legal form and activity. In particular:

- for sole proprietors, total employment = paid employees + 1;
- for partnerships, total employment = paid employees + number of partners.

Depending on the availability of administrative sources, more sophisticated methods have been devised in some countries.

Sources Administrative sources, surveys, calculations.

Comments The reference period used for the measurement of employment in business demography is a year, i.e. the labour force should be an annual average, though this can be approximated by using the number of persons employed at any given moment during the year if this is the only information available.

Both head counts and FTE have certain advantages and the latter should be recorded if possible. Head count is the number of physical persons, full-time and part-time, employed by a unit. FTE is the number of full-time equivalent jobs, defined as total hours worked divided by average annual hours worked in full-time jobs. FTEs are a more accurate measure of labour input but they are more difficult to measure as additional data would be required which is difficult to survey. As the concept of 'full-time' may vary over the economic sectors and countries, the definition does not really make the FTE data comparable. Given the administrative origin of the data, it may not be possible to calculate FTEs in some countries. Another possibility would be to use 'hours worked' directly. This is gaining favour in employment statistics, but the comment on data availability is also likely to apply to this characteristic.

Establishment - Geographical location code

Purpose The geographical location code complements the address and postal codes and can be used to derive classifications relating to the geographical location of units at the most detailed level, also other national classifications such as administrative regions, travel-to-work areas, health regions, and education regions.

Definition Countries can decide which code is most useful for their own purpose.

Proxies The geographical location code can refer to a classification at the most detailed level used in the country. It can refer to geocoding according to latitude and longitude points

Sources Administrative sources.

C3.4 Establishment - Relationship between units

Establishment - Identification number of the enterprise to which the establishment belongs

Purpose To identify the establishments belonging to an enterprise: for example to link geographical details to the economic enterprise data.

Sources SBR procedures

C4 Local Unit of Enterprise (or Truncated Enterprise) Characteristics

The characteristics listed in the following paragraphs apply equally to a local unit of a truncated enterprise as well as to any other kind of enterprises.

C4.1 Local Unit of Enterprise - Identification characteristics

Local unit of Enterprise - Identification number

Purpose To identify the unit and to be able to record its continuity.

Definition The identity number of a local unit should remain the same while the unit continues according to the continuity rules defined in Chapter 6 (even when the legal unit to which it belongs changes).

Proxies Using a unique administrative identity number (if it exists) is also possible, but it may cause problems with continuity and thus is not recommended.

Sources SBR procedure, unless the administrative number is used.

Local unit of Enterprise - Name

Local unit of Enterprise - Address

Local unit of Enterprise - Telephone and fax numbers, electronic mail address and information to permit electronic collection of data

Purpose To contact the unit.

Definition The official name of the local unit of an enterprise (and truncated enterprise) is generally the same as the legal unit that controls it, with some additional part usually specifying location or activity. If there is only one local unit in the legal unit, a separate name may not exist.

The actual address of the location of the unit and the contact information should always be recorded.

Proxies Different local units within a legal unit may use different trading names, also known as *signboard names* or *commercial names*, and provision should be made for recording these different trading names where they exist.

- Sources Administrative sources and surveys.
- Comments The address could be given a code referring to a national territorial nomenclature. It would be preferable if that nomenclature or national geographical code enables the unit to be pinpointed as accurately as possible, at the level of the street, section of street and building.
- It is always useful to provide for the possibility of recording a correspondence address in addition to the physical address of the local unit if the unit wishes correspondence relating to statistical surveys to be sent elsewhere.

Local unit of Enterprise - Identity number of the enterprise to which it belongs

- Purpose The local unit of enterprise (and truncated enterprise) must be linked with the enterprise to which it belongs.
- Definition This link can be included in the SBR by adding the identification number of the enterprise to the local unit of enterprise (and truncated enterprise) (and vice versa).
- Sources SBR procedures.

C4.2 Local Unit of Enterprise- Demographic characteristics

Local Unit of Enterprise- Date of commencement of activities (economic birth)

- Purpose The date is needed for monitoring the demography of local units.
- Definition This date should refer to the birth or other creation date of the unit according to the continuity rules.
- Sources SBR procedures.

Local Unit of Enterprise- Date of permanent cessation of activities (economic death)

- Purpose The date is needed for monitoring the demography of local units.
- Definition The date refers to the death of the unit or the date on which it became permanently inactive.
- Proxies As explained in connection with the corresponding characteristic of legal units, this date may not be available with any precision. Only the fact that the local unit has ceased to exist or has ceased its activities during the reference year may be known, in which case the date should be estimated.
- Sources Surveys, administrative sources, SBR procedures or estimated.

C4.3 Local Unit of Enterprise- Economic/stratification characteristics

Local Unit of Enterprise- Principal activity code at ISIC 4-digit level

- Purpose The principal activity code is a stratification characteristic and is important in the compilation of regional and small area statistics
- Definition The activity code is determined according to ISIC Rev 4. The activities actually conducted by the unit are what matters, even if they are ancillary in the context of the enterprise.

Proxies	Several criteria can be used in defining principal activity code. Employment by each different activity can be used (if available), also activity descriptions.
Sources	Surveys and administrative sources.
Comments	If the local unit activity is ancillary in the context of the enterprise, this should be indicated.

Local Unit of Enterprise- Secondary activities, if any, at ISIC 4-digit level

This concerns only local units which are the subject of surveys

Purpose	Secondary activities may be used for helping determining establishments, which are of key importance for national/regional accounts. Establishments can be recorded in the SBR as separate units (this practice is commonly applied by many smaller countries). If this is not the case, this characteristic offers the possibility of identifying them for analytical purposes.
Definition	The activity codes are allocated in accordance with ISIC Rev 4.
Proxies	Several criteria can be used for defining the secondary activity code(s). Employment by each different activity can be used (if available), also activity descriptions.
Sources	Surveys or administrative sources.

Local Unit of Enterprise - Activity carried out in the unit constituting an ancillary activity of the enterprise to which it belongs (yes/no) (non-core)

Purpose	To distinguish ancillary units. This characteristic enables statistical analyses to reallocate the cost of ancillary activities to the activities for the benefit of which they are pursued.
Definition	An ancillary activity is incidental to the main activity of an enterprise. It facilitates the efficient running of the enterprise but does not normally result in goods and services that can be marketed. For enterprises that are relatively small and have only a single location, ancillary activities are not separately identified. For larger enterprises with multiple locations, it may be useful to treat ancillary activities in the same way as a secondary or even a principal product ⁸³ .
Proxies	Certain ISIC classes often constitute ancillary activities.
Sources	Administrative sources, surveys.
Comments	Whether or not a unit is an ancillary unit should be explicitly recorded in the SBR. The absence of a value should be interpreted as <i>having no knowledge about this characteristic</i> .

⁸³ SNA2008: 5.10

Local Unit of Enterprise - Number of persons employed

Local Unit of Enterprise - Number of employees (non-core)

Local Unit of Enterprise - Number of employees in full-time equivalent (non-core)

Purpose The SBR should record the actual numbers of persons employed and employees, both as head counts and the latter also in FTEs. The main aim is to obtain stratification characteristics as well as statistical information. Persons employed are preferable for stratifying survey samples for very small units. The business register employment figures are used especially for small area statistics, where the business register is the only comprehensive source. In addition, employment figures can be used by employment statistics where needed.

Definition The structural business statistics definitions should be used, with the exception that the requirement to measure the numbers as annual averages does not apply. For stratification purposes and according to the BR intention is to use the situation at the end of year (including seasonally active units). As the end date approach is not harmonised the annual average can also be used as reference calculated for a certain period. The number of employees in full-time equivalents might be calculated for a full year as well as for the period in which the enterprise is active. If number of employees in fulltime equivalents is used as stratification characteristic, the calculation for the active period should be used, while for statistics covering a year the calculation should cover the whole year.

Proxies These figures can be obtained directly in some countries, while other countries may have an administrative source available only for the number of paid employees. However, the latter countries can obtain total employment by making a statistical adjustment to their figures on paid employees by adding a constant representing working proprietors and any other form of unpaid employment calculated according to legal form and activity, for example:

- For sole proprietors, total employment = paid employees + 1;
- For partnerships, total employment = paid employees + number of partners.

Depending on the availability of administrative sources, more sophisticated methods have been devised in some countries.

Sources Administrative sources, surveys, calculations.

Comments Note that the reference period used for the measurement of employment in business demography is a year, i.e. the labour force should be an annual average, though this can be approximated by using the number of persons employed at any given moment during the year if this is the only information available. How the annual average is calculated depends on the updating frequency of the register. If the unit operates during only part of the year (seasonal, new enterprises), the average should be calculated for that period.

Both head counts and FTE have certain advantages and the latter should be recorded if possible. Head count is the number of physical persons, full-time and part-time, employed by a unit. FTE is the number of full-time equivalent jobs, defined as total hours worked divided by average annual hours worked

in full-time jobs. FTEs are a more accurate measure of labour input but they are available in fewer countries. As the concept of ‘full-time’ may vary, the definition does not really make the FTE data comparable. Given the administrative origin of the data, it may not be possible to calculate FTEs in some countries. Another possibility would be to use ‘hours worked’ directly. This is gaining favour in employment statistics, but the comment on data availability is also likely to apply to this characteristic.

Local Unit of Enterprise- Geographical location code

Purpose	The geographical location code complements the address and postal codes and can be used to derive classifications relating to the geographical location of units at the most detailed level. Other national classifications such as administrative regions, travel-to-work areas, health or education regions etc. can also be derived from it.
Definition	Member States can decide which code is most useful for their own purpose.
Proxies	The geographical location code may refer to classification at the most detailed level used in the country. It may refer to geocoding to latitude and longitude points recorded by GPS in countries where the exact site of local unit can be recorded.
Sources	Administrative sources.

C4.4 Local Unit of Enterprise - Relationship between units

Local Unit of Enterprise - Identification number of the enterprise to which the local unit belongs

Purpose	To identify the local units belonging to an enterprise: for example to link geographical details to the economic enterprise data.
Sources	SBR procedures.

C5 Legal Unit - Characteristics

C5.1 Legal unit - identification characteristics

Legal unit - Identity number

Purpose	To identify the unit and to link it with other units in the SBR and with administrative and statistical sources.
Definition	The identity number of the legal unit can be either specific to the SBR or set by an administrative source and used by the SBR, or shared by several administrative sources and used by the SBR.
Sources	If the identity number of the legal unit assigned by an administrative source is used by the SBR, then its updating has to follow the changes taking place in the administrative source, even if it is an issue for economic purposes. For

instance, a change of legal form from natural to legal person may result in the fiscal administration assigning a different identifier to the same economic unit.

If the identity number of the legal unit in the SBR is not the administrative identifier, the administrative identifier should be handled as a characteristic and a record of its changes should be kept⁸⁴.

Comments Common business identifiers, shared with fiscal and other government departments greatly facilitate the connection of the SBR with other administrative sources.

The continuity rules for legal units depend on national legislation. There are no general recommendations.

Legal unit - Name

Legal unit - Address

Legal unit - Telephone and fax numbers, electronic mail address and information to permit electronic collection of data (Non-core)

Purpose To contact the unit.

In the event that a common identifier is missing, names and addresses can also be used for linking units.

Definition The addresses should be recorded at the most detailed level possible and respect international standards.

The information may refer to a legal or a natural person.

In the case of a legal person, the official business name and address must be recorded.

In the case of a natural person, it may be useful to maintain both a business and a personal address. Also, the following information may be recorded

- family name(s)
- names normally used and possible pseudonyms
- forenames
- gender (for address purposes — Dear Mr/Ms ...)

This information is often insufficient to identify a natural person with certainty. It may therefore be useful also to record either the date of birth or the personal identity number.

Proxies In some countries and for some legal forms, the business name may be very long and have to be abbreviated. In this case, strict rules on abbreviation must be established and applied.

Where applicable, territorial classifications and nomenclatures may be used

⁸⁴ This could be the situation if an administrative legal unit is given a unique identifier related to the region in which it is registered. If the unit moves to another region within the country and then is given a new identifier, there is not actually a unique administrative identifier for the unit for that country.

for coding of addresses. It is even better if addresses respect the ISO standard.

- Sources Administrative sources, mainly trade/company registers and surveys.
- Comments Legal persons and also sole proprietors often use initials, an acronym or a trading name instead of their official name in their business or administrative relations. There must be provision for recording this information separately.

C5.2 Legal unit - Demographic characteristics

Legal unit - Date of incorporation for legal person or date of official recognition as an economic operator for natural person

- Purpose The characteristic is needed for the inclusion of new units
- Definition 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 via a company/trade register, a VAT register or other register.
- Proxies If the exact date is not available, the year from which the unit has been monitored can be used as proxy, including a certain date (like 1 January) that is indicative of the situation, as defined in SBR procedures.
- Sources Administrative sources: Trade/company register, tax administration, social security.
- Comments In general, the date is the prerequisite for a unit to engage in legal economic transactions. Given that the SBR is usually supplied from administrative sources, a date for official recognition should always exist and be stored in the SBR. The legal unit may start its economic activity (and only then be regarded as an enterprise or part thereof) with some delay after its recognition, or it may remain economically inactive. If a legal unit remains inactive, it can either be omitted from the register, or kept in the SBR but marked as inactive. The latter approach is preferable. Keeping legally alive but economically inactive legal units in the SBR facilitates the use of data from administrative sources.

Legal unit - Date of commencement of economic activity

- Purpose To ensure that a legal unit that has started economic production is recognised as an active enterprise.
- Definition The 'date of effective economic activity' should be the date on which the unit starts its economic activity. This date may be declared by the unit when it registers officially and receives an identification number, or when its legal existence was approved, be it via a company/trade register, a VAT register or another register.
- Proxies If the exact date is not available, the year from which the unit has begun its activities can be used as proxy, including a certain date (like 1 January) defined by the SBR procedures.
- Sources Administrative sources: Trade/company register, tax administration, social security.

Legal unit - Date of commencement of economic activity

Comments In general, the date is the prerequisite for a legal unit to engage in legal economic transactions. Given that the SBR is usually supplied from administrative sources, a date for an effective economic activity should always exist and be stored in the SBR.

Legal unit - Date of permanent cessation of activities (economic death)

Legal unit - Date on which the legal unit ceased (administrative death)

Purpose These demographic characteristics are needed for monitoring the death or permanent inactivity of the unit from both economic and administrative perspectives.

Definition A legal unit ceases to be part of an enterprise (economic death) when:

- the unit ceases to be economically active and it is not part of the control chain within the enterprise group;
- the unit ceases to exist.

The record of a legal unit that has ceased (administrative death) should be kept in the SBR marked dead.

Proxies Registration of the year of economic death is important, even if the precise day and month are not known, as is often the case.

Sources Administrative sources (from which the date often comes with considerable delay), surveys.

Comments A legal unit usually has no particular interest in officially announcing cessation of its activities. Cessation may be the end result of a slow process of diminishing activity. Even if the activity completely ceases, the owner may still want to retain the option of resuming in the future and hence may retain a legal name, fiscal number and other legal attributes.

Between economic death and administrative death there is often a period of inactivity during which the unit may be regarded as economically *inactive* (also called *dormant*). Signs of this situation are lack of employees, the cessation of tax compliance and inability to contact the unit after repeated efforts.

C5.3 Legal unit - Economic/stratification characteristics

Legal unit - legal form

Purpose Legal form (also known as *legal status*) is useful in helping to eliminate ambiguity in identification searches. It is also a possible criterion for selection of units into survey frames and for stratification for survey sampling. It is also useful in defining the institutional sector (see Annex D). In monitoring the internal market there is an interest in being able to distinguish publicly traded and incorporated companies.

Definition Defined according to national legislation. Section 4.6.4 gives an example for EU countries.

- Sources Administrative sources, surveys. The legal form of a legal or natural person determines the tax regime applicable to the unit. Thus, if the SBR is supplied with taxation records it has access to the values of this characteristic.
- Comments Adjustments to questionnaires and collection processes are often made in accordance with legal form of the legal unit operating an enterprise.

Legal unit - Principal activity

- Purpose As a stratification characteristic and for the compilation of regional and small area statistics.
- Definition The activity code is determined according to ISIC Rev 4. The code is based on the activities actually conducted by a legal unit even when they are ancillary in the context of the enterprise.
- Proxies Employment or value added for each of the different activities can be used (if available), also activity descriptions, etc.
- Sources Surveys and administrative sources.
- Comments This characteristic indicates whether the activities are ancillary in the context of the enterprise.

Legal unit - Secondary activities

- Purpose For indirect use by national/regional accounts to estimate enterprise or establishment figures.
- Definition The activity codes are determined according to ISIC Rev 4. The code is based on the activities actually conducted by a legal unit even when they are ancillary in the context of the enterprise.
- Proxies Employment or value added for each of the different activities can be used (if available), also activity descriptions, etc.
- Sources Surveys and administrative sources.

Legal unit - Ancillary unit (yes/no) (non-core)

- Purpose To distinguish ancillary legal units. This characteristic enables statistical analyses to reallocate the cost of ancillary activities to the activities for the benefit of which they are conducted.
- Definition An ancillary activity is incidental to the main activity of a unit. It facilitates the efficient running of the unit but does not normally result in goods and services that can be marketed. For units that are relatively small and have only a single location, ancillary activities are not separately identified. For larger units with multiple locations, it may be useful to treat ancillary

activities in the same way as a secondary or even a principal product⁸⁵.

Proxies	Ancillary activities tend to relate to certain ISIC classes, for example book-keeping. However, while book-keeping is a potential ancillary activity, it is predominantly done by the businesses that offer book keeping services on the market and constitute the corresponding ISIC class.
Sources	Surveys and administrative sources.
Comments	Whether or not a unit is an ancillary unit should be explicitly recorded in the SBR. The absence of a value should be interpreted as <i>having no knowledge about this characteristic</i> .

Legal unit - Turnover (with flag for consolidated turnover)

Purpose As a size indicator for stratification, and for profiling an enterprise group using a bottom-up approach.

Definition Turnover consists of all income arising during the reference period, in the course of ordinary activities of the unit and is presented net of all price reductions, discounts and rebates granted by it.

The income referred to is arising from contracts with customers and are realized through the satisfaction by the unit of the contractual performance obligations usually represented by the sale (transfer) of goods or the rendering of services, however, the it can also contain revenues obtained as a yield on the use by others of the statistical unit's assets.

Excluded from net turnover are:

- all taxes, duties or levies linked directly to revenue;
- any amounts collected on behalf of any principal, if the statistical unit is acting as an agent in its relationship with said principal;
- all income not arising in the course of ordinary activities of the statistical unit. Usually, these types of income are classified as "Other income", "Financial income", "Extra-ordinary income" or under a similar heading, depending on the respective set of generally accepted accounting standards used to prepare the financial statements.

Proxies Fiscal sources may be used as sources of information, for example, corporate income tax or VAT returns. There may be cases where such information is unavailable, for example where units are not subject to VAT. In such a situation one possibility is to estimate turnover based on employment using a standard turnover per head ratio calculated by the ISIC class.

Turnover is relevant for market units; it should not be estimated for non-market units. However, non-market units can have secondary production.

Sources Administrative sources and surveys.

⁸⁵ SNA2008: 5.10

Legal unit - Number of employees

Legal unit - Number of persons employed (non-core)

Purpose The SBR should record the actual numbers of persons employed and employees, both as head counts and the latter also in FTEs. The main aim is to obtain stratification characteristics as well as statistical information. Persons employed are preferable for stratifying survey samples for very small units. The SBR employment figures are used especially for small area statistics, where the SBR is the only comprehensive source. In addition, employment figures can be used by employment statistics where needed.

Definition *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.

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.

For stratification purposes the intention is to use the situation at the end of year (including seasonally active units). As the end date approach is not harmonised the annual average can also be used as reference calculated for a certain period. The number of employees in full-time equivalents might be calculated for a full year as well as for the period in which the legal unit is active. If number of employees in fulltime equivalents is used as stratification characteristic, the calculation for the active period should be used, while for statistics covering a year the calculation should cover the whole year.

Proxies These figures can be obtained directly in some countries, while other countries may have an administrative source available only for the number of paid employees. However, the latter countries can obtain total employment by making a statistical adjustment to their figures on paid employees by adding a constant representing unpaid employment (including working proprietors), calculated e.g. according to legal form and activity:

For sole proprietors, total employment = paid employees + 1;

For partnerships, total employment = paid employees + number of partners.

Depending on the availability of administrative sources, more sophisticated methods have been devised in some countries. Note also that 0 means less than half a person, whether calculated as head counts or FTEs

Sources Administrative sources, surveys, calculations.

Comments Note that the reference period used for the measurement of employment in

business demography is a year, i.e. the labour force should be an annual average, though this can be approximated by using the number of persons employed at any given moment during the year if this is the only information available. How the annual average is calculated depends on the updating frequency of the register. If the unit operates during only part of the year (seasonal, new enterprises) the average should be calculated for that period.

Both head counts and FTE have certain advantages and the latter should be recorded if possible. Head count is the number of physical persons, full-time and part-time, employed by a unit. FTEs are defined in SBS (variable 16 14 0) and also in national accounts (FTE employment is the number of full-time equivalent jobs, defined as total hours worked divided by average annual hours worked in full-time jobs). FTEs are a more accurate measure of labour input but they are available in fewer countries. As the concept of 'full-time' may vary, the definition does not really make the FTE data comparable. Given the administrative origin of the data, it may not be possible to calculate FTEs in some countries. Another possibility would be to use 'hours worked' directly. This is gaining favour in employment statistics, but the comment on data availability is also likely to apply to this variable.

Legal unit - Institutional sector and sub-sector

- Purpose The institutional sector classification for legal unit is determined by the institutional sector classification of the enterprise with which it is associated.
- Sources SBR procedures
- Comments In practice, institutional sector classification is often applied directly to legal units. Then the enterprise inherits its institutional sector classification from the legal unit(s) from which it is constituted.
- Also in practice, enterprise and legal unit coincides in the majority of cases.

C5.4 Legal unit - Relationship between units

Link with Enterprise Group and Truncated Group

Legal Unit - Percentage of control of legal units by the enterprise group (direct + indirect)

- Purpose To allow the delineation of the enterprise group.
- Sources SBR procedures.

Legal Unit - Percentage of ownership of legal units (direct + indirect)

- Purpose To support calculation of control.
- Sources SBR procedures.

Link between enterprise and legal unit

Legal Unit - Identification number of the enterprise(s) to which the legal unit belongs

- Purpose To identify the legal units belonging to an enterprise.

Legal Unit - Identification number of the enterprise(s) to which the legal unit belongs

Sources SBR procedures

Links with local units

Legal Unit - Identification number of the local unit(s) that belongs to the legal unit

Purpose To identify the local units belonging to the legal unit.

Sources SBR procedures

Consolidation method (non-core): *Integration method of the legal unit in the consolidated accounts (if relevant).*

C5.5 Legal unit - Relationship with other registers/update sources

Legal unit - Value added tax (VAT) identification number

Legal unit - identification number of each other administrative source used to maintain the SBR or to compile economic statistics

Purpose For linking administrative data to the SBR; for the links to non-resident units and for the links to foreign trade.

Definition Defined by the corresponding administrative source. A separate VAT identification number may not exist in countries where a single identification system for legal units is used by several administrative sources.

Proxies Certain economic activities may be exempted from VAT and thus have no VAT number. In this case identity number from another administrative source, such as tax number, corporate registration number may play the same role.

Sources Mainly tax administration.

Comments In some countries, a VAT number may relate to part of a legal unit or to a natural person who is not an economic operator but who has a VAT number for tax reasons.

Legal unit - Balance sheet data registration identification number

(Applicable only for legal units required to publish accounts.)

Legal unit - Balance of payments register or foreign direct investment register

Legal unit - Farm register

Legal unit - Chamber of commerce and trade associations

Purpose The first link concerns the balance sheet data. Many countries use the published accounts as a source for the SBR and combining the SBR and published accounts data is likely to become very important in the future. It reduces response burden and serves the production of economic and financial statistics.

The second link concerns the BoP and FDI registers and the usefulness of these links concerns the harmonisation of statistics related to globalisation. Conventional bank settlements data are becoming more and more frequently replaced by data based on direct surveys, for which the BoP compilers are increasingly relying on SBRs.

The third link to farm register (if separate from the SBR) is important for the coverage of main agricultural enterprises and for updating the increasing rural multi-activities, where the role of agriculture as principal or secondary activity may often change.

The fourth (set of) link(s) to (often) compulsory registrations that exist in countries for various reasons, for example establishment of legal persons and liability commitments of producers. These administrative or para-administrative bodies may well register births, deaths and changes in legal structures of the legal units within their realms of responsibility. They often cover a large part but not all of the legal units in scope for the SBR.

Definition The practical arrangement of the links, either from the SBR to the associated registers, or vice versa, is a country matter.

Proxies The links can be achieved in several ways, for example:

- Recording in the SBR the reference number of the unit in the other register(s), together with the legal unit identity number.
- Adding one or more marks to the SBR to indicate that the legal unit is also present in other register(s) under the same identity number.
- In the absence of a unique identifier, the link can also be built by name/address matching and possible use of other characteristics. This is less effective, but matching names/addresses can be useful for detecting errors. For every legal unit recorded in administrative sources, the different relations should be stored separately.

Sources Administrative sources

Comments The requirement to publish annual accounts depends on national legislation, which may vary between countries. It usually applies to incorporated and publicly traded companies.

The data from published accounts may be considered free from confidentiality restrictions and thus suitable for exchange with other NSIs. However, this is not the case if the data are accompanied by data collected via surveys.

The links to balance sheet data can be used for combining SBR and accounts data, which are generally available in satellite registers.

In the EU, either the national central bank (NCB) or the NSI is a member of the European Committee of Central Balance Sheet Data Office, which supplies data for an increasing number of Member States.

The BoP register in most countries is maintained by the NCB. In this case close cooperation and exchange of information between the NSI and the NCB is vital for well targeted and good quality BoP and FDI surveys. The BoP register may alternatively be maintained by the NSI, or there may not be a

separate BoP register and the SBR is used for BoP purposes.

The link between farm and business registers is discussed in Section 3.4.1.

Several other registers (e.g. shop register, tourist establishment register, transport registers, educational institutes register, etc.) may also be important for updating the SBR, especially in determining whether or not units are active, because data in these registers tend to be updated frequently.

Harmonisation of units in the SBR and these associated registers is an important issue.

C6 Local Unit of Legal Unit - Characteristics

C6.1 Local unit of legal unit - Identification characteristics

Local unit of legal unit - Identification number

Purpose	To identify the unit and to be able to record its continuity.
Definition	An SBR-specific identity number is recommended. The identity number should remain the same while the unit is considered as continuing according to the continuity rules as defined in Chapter 6 (even though the legal unit to which it belongs may change).
Proxies	Using a unique administrative identity number (if one exists) is also possible, but may cause problems in applying continuity rules.
Sources	SBR procedure, unless an administrative number is used.

Local unit of legal unit - Name

Local unit of legal unit - Address

Local unit of legal unit - Telephone and fax numbers, electronic mail address and information to permit electronic collection of data

Purpose	To contact a unit.
Definition	The official name of the local unit of legal unit is generally the same as the legal unit that controls it, but with an additional part usually specifying location or activity. If there is only one local unit belonging to the legal unit, a separate name may not exist. The actual address of the location of the unit and the contact information should always be recorded.
Proxies	Different local units within a legal unit may use different trading names, also known as <i>signboard names</i> or <i>commercial names</i> , and provision should be made for recording these different names where they exist.
Sources	Administrative sources and surveys.
Comments	The address may be given a code referring to a national territorial nomenclature. In this case it is preferable to use a nomenclature or national geographical code that enables the unit to be pinpointed as accurately as

possible, at the level of street, section of street and building.

It is always worthwhile providing for the possibility of recording a correspondence address (for survey purposes) in addition to the physical address of a unit.

Local unit of legal unit - Identity number of the enterprise to which it belongs

Purpose	The local unit of legal unit must be linked with the enterprise to which it belongs, generally through the link between the legal unit and the enterprise.
Definition	This link can be included in the SBR by adding the identity number of the enterprise to the local unit of legal unit (and vice versa).
Proxies	When an enterprise consists of one legal unit only, itself with only one local unit, a simpler arrangement of the SBR is possible.
Sources	SBR procedures.
Comments	May be used to define the perimeter of local units of enterprises.

C6.2 Local unit of legal unit - Demographic characteristics

Local unit of legal unit - Date of commencement of activities (economic birth)

Purpose	For monitoring the demography of legal local units.
Definition	This date should refer to the birth of the unit, or other creation date of the unit according to the continuity rules.
Proxies	If the exact date is not known, it may be estimated as the administrative birth date.
Sources	Surveys, administrative sources and estimation.
Comments	The economic birth date may be different the administrative birth date.

Local unit of legal unit - Date of final cessation of activities (economic death)

Purpose	For monitoring the demography of legal local units.
Definition	The date refers to the death of the unit or the date on which it becomes permanently economically inactive.
Proxies	As explained for the corresponding characteristic of legal unit, this date may not be available with any precision. Only the fact that the local unit has ceased to exist or has ceased its activities during the reference year may be known. In this case the date should be estimated.
Sources	Surveys, administrative sources, SBR procedure and estimation.

C6.3 Local unit of legal unit – Economic/stratification characteristics

Local unit of legal unit - Principal activity code at ISIC 4-digit level

Purpose	As a stratification characteristic and in compilation of regional and small area statistics
---------	---

Definition	The activity code is determined according to ISIC Rev 4. The activities actually conducted by the unit determine the code, even when they are ancillary in the context of the enterprise.
Proxies	Employment (if available) by its various activities (if more than one) can be used, also activity descriptions.
Sources	Surveys and administrative sources.
Comments	If the activity is ancillary to that of the enterprise, this fact should be recorded.

Local unit of legal unit - Secondary activities, if any, at ISIC 4-digit level

This concerns only legal local units which are the subject of surveys.

Purpose	Secondary activities may be used in helping delineate establishments, If establishments are not maintained in the SBR, this characteristic offers the possibility of defining them for analytical purposes.
Definition	The activity codes are decided according to ISIC Rev 4.
Proxies	Employment (if available) by its various activities (if more than one) can be used, also activity descriptions.
Sources	Surveys and administrative sources.
Comments	Local units of legal units are generally large and important and their information should be obtained and updated in the SBR. The local units referred to here are divisions (or other operational units) of legal units, not of enterprises, which means that their links for example to establishments are not direct but through the legal units that constitute the corresponding enterprise.

Local unit of legal unit - Activity carried out in the unit constituting an ancillary activity of the enterprise to which it belongs (yes/no) (non-core)

Purpose	To distinguish ancillary units. This characteristic enables statistical analyses to reallocate the cost of ancillary activities to the activities for the benefit of which they are conducted.
Definition	An ancillary activity is incidental to the main activity of an enterprise. It facilitates the efficient running of the enterprise but does not normally result in goods and services that can be marketed. For enterprises that are relatively small and have only a single location, ancillary activities are not separately identified. For larger enterprises with multiple locations, it may be useful to treat ancillary activities in the same way as a secondary or even a principal product ⁸⁶ .
Proxies	Ancillary activities typically belong to certain ISIC classes.

⁸⁶ SNA2008: 5.10

- Sources Administrative sources and surveys.
- Comments Whether or not a unit is an ancillary unit should be explicitly recorded in the SBR. The absence of a value should be interpreted as *having no knowledge about this characteristic*.

Local unit of legal unit - Number of persons employed

Local unit of legal unit - Number of employees (non-core)

Local unit of legal unit - Number of employees in full-time equivalents (FTEs) (non-core)

Purpose The register should record the actual numbers of persons employed and employees, both as head counts and the latter also in FTEs. The main aim is to obtain stratification characteristics as well as statistical information. Persons employed are preferable for stratifying survey samples for very small units. The business register employment figures are used especially for small area statistics, where the business register is the only comprehensive source. In addition, employment figures can be used by employment statistics where needed.

Definition The structural business statistics definitions should be used, with the exception that the requirement to measure the numbers as annual averages does not apply. For stratification purposes and according to the BR intention is to use the situation at the end of year (including seasonally active units). As the end date approach is not harmonised the annual average can also be used as reference calculated for a certain period. The number of employees in full-time equivalents might be calculated for a full year as well as for the period in which the enterprise is active. If number of employees in fulltime equivalents is used as stratification characteristic, the calculation for the active period should be used, while for statistics covering a year the calculation should cover the whole year.

Proxies These figures can be obtained directly in some countries, while other countries may have an administrative source available only for the number of paid employees. However, the latter countries can obtain total employment by making a statistical adjustment to their figures on paid employees by adding a constant representing unpaid employment (including working proprietors), calculated e.g. according to legal form and activity:

For sole proprietors, total employment = paid employees + 1;

For partnerships, total employment = paid employees + 2.

Depending on the availability of administrative sources, more sophisticated methods have been devised in some countries. Note also that 0 means less than half a person, whether calculated as head counts or FTEs.

Sources Administrative sources, surveys, and SBR calculations.

Comments Note that the reference period used for the measurement of employment in business demography is a year, i.e. the labour force should be an annual average, though this can be approximated by using the number of persons employed at any given moment during the year if this is the only information

available. How the annual average is calculated depends on the updating frequency of the register. If the unit operates during only part of the year (seasonal, new enterprises), the average should be calculated for that period.

Both head counts and FTE have certain advantages and the latter should be recorded if possible. Head count is the number of physical persons, full-time and part-time, employed by a unit. FTE is the number of full-time equivalent jobs, defined as total hours worked divided by average annual hours worked in full-time jobs. FTEs are a more accurate measure of labour input but they are available in fewer countries. As the concept of 'full-time' may vary, the definition does not really make the FTE data comparable. Given the administrative origin of the data, it may not be possible to calculate FTEs in some countries. Another possibility would be to use 'hours worked' directly. This is gaining favour in employment statistics, but the comment on data availability is also likely to apply to this characteristic.

Local unit of legal unit - Geographical location code

Purpose The geographical location code complements the address and postal codes and can be used to derive classifications relating to the geographical location of units at the most detailed level.

Other national classifications such as administrative regions, travel-to-work areas, health region and education regions can also be derived from it.

Definition Countries should decide which code is most useful for their own purpose.

Proxies The geographical location code may refer to classification at the most detailed level used in the country. It may refer to geocoding to latitude and longitude points.

Sources Administrative sources.

C6.4 Local Unit of Legal Unit - Relationship between units

Local Unit of Legal Unit - Identification number of the legal unit to which the local unit belongs

Purpose To identify the legal unit to which the local unit belongs.

Sources SBR procedures.

C6.5 Local Unit of Legal Unit - Relationship with other updating sources

Local Unit of Legal Unit - Reference to administrative source(s) in which the local unit of legal unit appears

(where such sources exist and contain information that can be used for statistical purposes)

Purpose To provide links to administrative sources that provide data about local units of legal units for updating the SBR.

Definition Depends upon the particular situation.

Sources Administrative sources.

Comments This characteristic is country specific. There may not be any such sources in some countries, while in other countries there may be several.

Some sources may refer to a legal unit rather than a local unit of legal unit, so the reference unit should be checked carefully. This might be the case if the local unit is located at the residence of a legal unit which comprises several local units. In the case of a complex enterprise, there may be several legal units with a single local unit at the same address, probably forming one local unit for statistical purposes. In this case, all links should be stored.

Annex D: Classification of institutional sectors⁸⁷

The institutional sector classification groups together similar kinds of institutional units according to the nature of the economic activity they undertake. Corporations, NPIs, government units and households are intrinsically different from each other in their economic objectives, functions and behaviour. In relation to the three basic economic functions:

- production of goods and services;
- consumption to satisfy human needs; and
- accumulation of various forms of capital

only the production function is of interest in the context of the SBR. So, only the institutional units engaged in production are covered in the SBR and can be classified by institutional sector. Thus, only households which engage in market production are covered in the SBR (in the form of sole proprietorships). As the SBR covers only resident units, the institutional sector “rest of the world” is also not of interest.

Corporations undertake either production and/or accumulation, but not consumption. Corporations are split into non-financial and into financial ones. This distinction is made because of the special role that financial corporations play in the economy. Government undertakes production, accumulation and final consumption on behalf of the population. Households undertake consumption on their own behalf and may also engage in production and accumulation. NPIs are diverse in nature: some behave like corporations and some undertake activities similar to government but independently of it.

The five institutional sectors relevant for the SBR are:

- the non-financial corporations sector (S.11);
- the financial corporation sector (S.12);
- the general government sector (S.13);
- the non-profit institutions serving households (S.14);
- the household sector (S.15).

The distinction between corporation and government is quite fundamental, but also quite difficult to apply. It refers to the distinction between market and non-market production. Corporations produce for the market and aim to sell their products at economically significant prices. Government units organize and finance the provision of goods and services to individual households and to the community as a whole. They may produce most of these goods and services themselves but the products are either provided free or at prices determined by considerations other than purely market forces. Households are primarily consumers but some of them are also producing goods and services for the market. When production takes place in the household but not in separate legal entity owned by the households, it is described as an unincorporated enterprise. It is then part of the institutional unit household.

As mentioned above NSIs are behaving mainly as non-market producers, in that they are producing goods and services not for the generation of income or profit. However, some NSIs behave like corporations and should thus be classified under S.11 or S.12. Other NSIs that are controlled by government are to be classified under S.13. The remaining NPIs, those that

87 “2013 ESA regulation” [2.143]

produce goods and services but do not sell them at economically significant prices and are not government controlled, are classified in a separate institutional sector called non-profit institutions serving households (NPISHs). They are mainly non-governmental social institutions, such as churches, social clubs, charitable associations, etc.

The five institutional sectors can further be divided into subsectors, depending on the national needs and intended analysis. In the European System of Accounts (ESA 2010) for almost all sectors subsectors have been created based on the type of institutional unit and control criteria. The sectors S.11 (non-financial corporations) and S.12 (financial corporations) are subdivided by whether the units are public, foreign controlled or nationally private controlled. The first category would include public corporations, the second one resident units belonging to foreign controlled multinational enterprise groups, whereas the last category comprises all the other institutional units classified in S.11 and S.12. The financial sector S.12 is further divided by type of institutional units into nine subsectors, from central bank, other monetary financial institutions, to insurance companies and pension funds. The general government sector is split into four subsectors: central government, state government, local government, and social security funds. And lastly, also the household sector is split into different types of households. Only the subsectors S.141 (employer households) and S.142 (own-account workers) are covered in the SBR.

The classification of institutional sector is quite different from an activity classification, such as ISIC. Therefore, the activity coding is not of much help for the coding according to the institutional sectors. Only in certain cases does an activity code lead directly to a sector code. Examples are central banks, insurance companies, pension funds and the activities of public administration. The legal form of an institutional unit is also a criterion that is partly of help: the main examples are the legal forms of sole proprietorship which are allocated to the household sector S.14. However, the legal form of non-profit associations is not per se of help as the NSIs can be classified to S.11, S.12, S.13 and S.14 according to their behaviour and control.

The main challenge for the implementation of the sector classification is the correct allocation of units to sectors S.12 and S.13. Sector S.15 can be implemented automatically based on the legal form and NPIs need to be coded manually. So, if the allocation to the sectors S.12, S.13 and S.14 is fixed, the remaining units that are not sole proprietorships can be allocated to sector S.11.

For the allocations to sectors S.12, S. 13 and S.15 additional information is required that is usually not available in the SBR and in the administrative data sources used in the SBR. The complexity of these sectors lies in deciding whether a unit is a market or a non-market producer and whether the unit is government controlled or not. Further, especially as concerns the financial sector, there are various types of units that need special attention, such as holdings, head offices, special purpose entities, captive financial units and artificial subsidiaries.

Working closely with the experts of national and financial accounts in the NSI is recommended in order to achieve a high quality and coherent classification according to the institutional sectors.

Figure D1 provides an overview of the relationships between the different kinds of institutional units and the institutional sectors.

Figure D1: Institutional Sector as Determined by Legal Description and Market/Non-Market Production

Type of producer		Market producers		Non market producers	
Standard legal description		<i>Goods and non-financial services</i>	<i>Financial intermediation</i>	<i>Public producers</i>	<i>Private producers</i>
Private and public corporations		S.11 Non-Financial corporations	S.12 Financial corporations		
Cooperatives and partnerships recognised as independent legal entities		S.11 Non-Financial corporations	S.12 Financial corporations		
Public producers which by virtue of special legislation are recognised as independent legal entities		S.11 Non-Financial corporations	S.12 Financial corporations		
Public producers <i>not recognised as independent legal entities</i>	Those with the characteristics of quasi-corporations	S.11 non-Financial corporations	S.12 Financial corporations		
	The rest			S.13 General government	
Non-profit institutions recognised as independent legal entities		S.11 Non-Financial corporations	S.12 Financial corporations	S.13 General government	S.15 Non-profit institutions serving households
Unincorporated household enterprises	Partnerships <i>not recognised as independent legal entities</i>	S.14 households ⁸⁸	S.14 Households		
	Sole proprietorships	S.14 households ⁸⁹	S.14 Households		
Head offices whose preponderant type of activity of the group of corporations controlled by them is the production of:	goods and non-financial services	S.11 Non-Financial corporations			
	financial services		S.12 Financial corporations		

⁸⁸ but “enterprises” for business statistics

⁸⁹ but “enterprises” for business statistics

Annex E: Examples of Statistical Business Registers

Annex E1: SBR Development at Statistics Denmark

Annex E2: Development and Maintenance of the SBR in Costa Rica

Annex E3: Statistics Canada's Statistical Business Register

Annex E4: Business Register at the National Statistics Office, Georgia

Annex E5: Reengineering the Business Register in Malaysia

Annex E1: SBR Development at Statistics Denmark

E1.1 Summary

The development of Danish business statistics from a situation where they covered only activities in the agricultural sector (which used to be the main activity in Denmark) to a situation where they cover all business activities, and the move from separate lists of businesses updated by individual statistics divisions to a comprehensive business register, are good examples of how an NSI can become an important player in developing a business register that is useful for statistics as well as for public administration and the businesses themselves.

So long as the statistics are based on censuses the need for a business register is not obvious, but when it is acknowledged that:

- the NSI does not have the resources to conduct censuses;
- the results of censuses are out of date before the results are published; and
- the response burden is high;

the need for a business register is clear. This was already obvious to the Nordic Council when it was formed in 1952, but it was not until 1959 that the first Danish Business Register was established.

Due to Statistics Denmark's (SD's) co-operation with other public institutions it has been possible to influence the development of different administrative registers. The breakthrough came in the end of the 1960's with the establishment of four central administrative registers. Based on this, SD was able to establish a statistical register as well as an administrative register based on regular electronic updating routines. A special law was passed to make a distinction between the statistical register and the administrative register, and it was commonly accepted. The basic information in the two registers may be the same (names and addresses, etc., as well as the administrative identification numbers), but some statistical information is available only to statisticians.

Also technical developments have increased the possibilities for a much more flexible and comprehensive business register and for identification of inconsistencies. In 1993 SD established a relational database and this was later the basis for a complete separation from the Statistical Business Register of the Central Administrative Business Register (CVR or ABR) that was first run by SD, then by the tax authorities and now by the Ministry of Businesses. SD is proud of the fact that, first, its work in linking administrative units to legal units was found to be important for public purposes and, second, the tax authorities, among others, found it useful for the ABR to include local kind of activity units, linked to legal units, for administrative purposes.

It is interesting that, because of its importance to the economy, the agricultural sector itself (organisations and the farmers) was very interested in the agricultural statistics, making it possible to maintain an annual census with data collected locally until 1970, after which the Real Estate Register of the municipalities formed the basis for the Danish Statistical Agricultural Register, separate from the Statistical Business Register. This Real Estate Register was used to update the Statistical Agricultural Register until 2008.

Still more administrative registers have been developed since 1970, and SD is, by law, required to participate in the preparatory work. SD determines whether the new register contains information of interest to statistics and tries to influence the legislator to take

statistical needs into consideration. One recent example is monthly reporting of various wage and employment figures to the tax authorities.

SD has worked hard to remove “shadow registers” kept by individual statistical divisions. The latest developments have been to incorporate (1) the agricultural register into the general business register, including the use of two administrative registers covering the agricultural sector (the Central Household Animals Register and the General Agricultural Register) and (2) the use of the digital maps and transformation of information on forests from these maps to categorise LKAUs.

In the following sections these experiences are elaborated.

E1.2 First Establishment of Danish Business Register

From 1896 to 1958 seven censuses were carried out covering only the non-agricultural industries. The same form was used for all activities and sizes of units. This did not produce economic statistics that were fully acceptable to users. The SNA also indicated that the establishment was the relevant statistical unit to use. The municipalities distributed and collected the census forms, carried out the primary validation and sent the forms to SD. Even after this, SD still had to contact a lot of businesses by telephone.

In Denmark the first business register was established on the basis of the business census in 1958. It was called a *Statistical Address List*. It was a list of companies with activities in industry, building and construction. The list was updated manually by information from address directories made available to SD free of charge. The register from which samples could be drawn from an updated population was ready in 1961. At this time only punched cards were used, which made use of the list very resource intensive.

E1.3 Breaking Point One – Precondition for Second SBR

By the end of the 1960s four electronic administrative registers had been established and electronic data processing was introduced at SD. This was a revolution. The four registers were as follows.

- *Wage Earner and Employer Register (1965)* - with the purpose of keeping information on the wage earners' supplementary pension.
- *Central Person Register (1968)* - every person in Denmark received a personal identification number (id) and it has to be used by government in all contacts with the public.
- *Value Added Tax register (1967)* - almost all businesses were obliged to register by the tax authorities in an electronic register when their turnover exceeded a certain threshold. For several years, the register was used to produce statistics on turnover.
- *Income Tax Register (1969)*. It became compulsory for all employers to withhold income tax for all employees each time wages and salaries were paid. The collection of payments to the wage earners' supplementary pensions could be collected at the same time. Later the register was extended to cover dividends, etc.

The Income Tax system, which was built to manage the flow of money, was not connected to the VAT system. These two new systems introduced the possibility of producing statistics on wages and numbers of employees.

One condition for the new systems to function was the establishment of a personal id-number system. Besides the personal id-number the Central Personal Register also included an

address register. Each occupied house and dwelling was in the address register. Each road was assigned a municipality code and a road code. The house numbers (and for flats also floor and flat numbers) that were already known were registered in the register. As businesses were not in the address register this created some challenges. However, as businesses had house numbers on roads, SD decided simply to extend the official address register with these extra house numbers.

The address register is still not complete. It will be fully incorporated into the building and construction register. At the same time all businesses will be given an address that is official. There are some challenges concerning businesses not connected to buildings (where x-y coordinates will be used) and concerning owners who are not living in Denmark (where addresses are still needed). The current system is usable, but complete integration is not expected until in the middle of 2016.

SD did not use the register of companies (run by the Ministry of Businesses). The reason was that this register was updated only after a voting at a general meeting and information on share capital could be delivered by VAT.

The law on SD in 1966 gave SD the right to access to the new registers. SD received information from the VAT register each month, including information to calculate turnover, and it received information from the Income Tax Register, including payments to the wage earners' supplementary pensions, from which number of employees could be calculated.

E1.4 Content of Second SBR

In the first instance SD did not make any attempts to integrate the two tax systems. Two separate files were kept. However, it soon became very obvious that it would be of interest to link information on turnover and number of employees. Also, the number of units (initially administrative units, later legal units) depended upon which system (the VAT System or the Income Tax System) was the basis for the calculation. This was not satisfactory.

It quickly turned out that it was difficult to link these two registers, and that a matching system had to be developed. Luckily, both the VAT register and the Income Tax Register kept links between two or more administrative units that were connected to the same legal unit. Each of the registers also identified one administrative unit as the main unit. This information was used to create a legal unit in each register.

If the legal unit was a single proprietor business the personal id-number was registered in both registers. So in case of a single proprietor business the two registers' identification numbers could be connected with virtually no problem. This was also the case for partnerships. The addresses did not always correspond in the two registers, but this was not a significant problem. A decision simply had to be taken about which should be used.

For other legal forms matching on addresses was the starting point. Also matching on telephone number was used. As not all businesses were registered in both registers the result of matching was a group of matched units and two groups of non-matching units. During the following updating routines all units which had a new or changed address were re-matched with the non-matched group from the other register, and the quality of the matching became better and better.

E1.5 Statistical Versus Administrative Business Register

In 1967 the Ministry of Administration had asked SD to begin to establish a general administrative business register. SD undertook responsibility for establishing a business

register both for the public and for statistical purposes, but making a clear distinction between the public part and the statistical part. Identification information such as the legal form and the activity code was the same. SD suggested introducing a common id-number that could be used to identify each legal unit, but this turned out to be impossible at that time, especially because it would have required a change in the two main source systems (the VAT System and the Income Tax System).

A law on the Administrative Business Register passed by the Danish Parliament in 1975.

Further development of the two Danish Business Registers is discussed below, but before getting to this, it is relevant to talk about the special cooperation concerning the number of employees at local kind of activity units (LKAUs).

E1.6 Work to Establish LKAUs

SD wanted to make employment statistics distributed by municipality (or even smaller areas). When the Income Tax System was designed it was not possible to attribute the employees to a local unit so SD had to conduct a survey for this purpose. As businesses already withheld individual taxes for each employee the only missing information in administrative sources was where the employees worked. In most cases the employer had only one local unit and all employees could be attributed to that location. This location was usually at the address registered by the tax authorities. It was also possible to see where the employees lived using administrative registers. If a group of employees lived far from the location identified by the tax registration of the business, it indicated that at least one more location was used by that employer. Thus, SD could limit its survey to those employers where these kinds of differences showed up.

A system to collect information on LKAUs was established together with the Tax Authorities and the business organizations. The system came into action in 1980. Each LKAU belonging to one legal unit with more than one LKAU was given a three digit number. Every November SD sent a list of LKAUs with addresses and activity codes to the legal unit. The legal unit had to complete the list by adding changes to addresses and activity codes by deleting LKAUs that no longer belonged to the legal unit and by adding new LKAUs. In the event of a takeover SD asked for information from the legal unit that had taken over the LKAU.

When a legal unit submitted information on income and withheld tax for each employee to the Tax Authorities, it had to indicate the three digit number of the LKAU at which the employee was employed at the end of November. The employer could also give information about whether or not the employee had been employed for the whole year. Where an administrative unit corresponded exactly to one LKAU, the legal unit did not need to fill in the LKAU list or the three digit number of the LKAU for the employees.

This working-place based system also included registration of correct addresses for those legal units that had only one LKAU but at a different address than that registered in the VAT register.

E1.7 Work to Establish a Coherent Business Register

During the 1970s and 1980s a simple record system was built. Even though LKAUs were registered for those legal units that had more than one, it turned out that there were some problems with the locations of the one-to-one units. An investigation unveiled the reason: for some units the address registered in the VAT register was not the address where the activity took place but the home of the owner or the place where he/she kept his/her account. This

meant that, as the SBR was updated every month, good information from the working-place system described above was overwritten. As a result it was not sufficient to keep records on those LKAUs where more than one was connected to a legal unit. There was a need for a complete list of LKAUs. At the same time as a result of technological developments, a move towards use of relational databases became obvious.

E1.8 Common Identification

The year is now about 1990. A committee had been set up by SD as far back as 1970, with members from several ministries, to determine how centralisation of business registration could be achieved. Although its report explained the need for common identification numbers, and although this same message was repeated by a similar group in 1982, this did not move the VAT and Tax Authorities before these two directorates were joined. Even then the situation was not ideal as the only agreement that could be obtained was that the VAT and the Tax numbers should be converted into one.

As SD already had done the work required to link the two systems, it was 'just' a matter of agreement about converting the identification numbers to the new number. In the case where there was a one-to-one situation the adjusted Tax number was used. In case of several administrative identification numbers being used by a legal unit, these were all converted to the new numbering scheme, but with still the main Tax number being identified as the main number. Now the VAT and Tax numbers were the same even though, in some cases, one number was used for VAT registration and two numbers were used for Tax registration. The numbers are referred to below as Tax numbers.

SD took the new id-number into use in 1987 in the Administrative Business Register. In 1992 the new number was also in use in the joined VAT and Tax register.

E1.9 Third Establishment of Danish Business Register

SD was then able to drop its matching procedure. Instead a relational database was created, including the administrative units and their relations and their relations to the *legal unit*, which was still the SD unit used in the Administrative Business Register still being maintained by SD. The relational database also included the *enterprise* and the *LKAU*, and the relation between these, and the relation between the enterprise and the legal unit, and in some cases the relation between LKAU and an administrative unit.

The working place division became a part of the Business Register division.

Business demography was a new important area. Therefore the new Business Register system also included information about demographic events at LKAU level, including relations between LKAUs involved in the same event.

The process and system that followed an LKAU over time was very complicated, using information about changes in addresses and activity codes, about starting and stopping legal units, and about employees and turnover.

E1.10 Fourth Establishment of Danish Business Register

No more than six years passed before the Business Register was revised again. The reason was the joining of VAT and Tax by the Tax authorities in 1992. Also the revision of the law on the Administrative Business Register (ABR) resulted in a wish to renew business registration more fundamentally to include other administrative registers as well and to

require that the public administration should not ask for the same information more than once.

SD is proud of the fact that, firstly, its work in linking administrative units to legal units was found important for external purposes, and secondly, that, among others, the Tax authorities found it useful for administrative purposes for the ABR to include administrative local kind of activity units.

Formally, legal units and LKAUs were transformed to the new ABR as legal units and legal LKAUs. SD kept legal units, enterprises and LKAUs together with the legal LKAUs and Tax units. Unfortunately, demographic registration was also moved to the ABR. In practice this could not function properly as a legal unit cannot accept an LKAU-identification number change. Thus, after a while SD took on this statistical work and included it in the statistical part of the Business Register.

E1.11 Further Development

The SBR has been further developed by including enterprise groups. A system for registration of owners of legal units has been taken into use and the data exchange with European Group Register has been incorporated.

SD has also worked hard to remove shadow registers kept by individual statistical divisions. One of the latest developments has been to incorporate the agricultural register into the general business register. This includes use of two administrative registers covering the agricultural sector (the Central Household Animals Register and the General Agricultural Register), also use of digital maps and transforming information they contain on forests to categorise LKAUs. The reason for this approach is that, if any one division discovers the need for a change, it is carried through to Business Register and thus the Business Register is up to date and available for use by all divisions. Also the statistics produced by different divisions are consistent in the sense that they start from the same base.

A few years ago a new system for employers to report the income of their employees to the Tax authorities came into use. Every month the employers have to report the income paid to each employee and by which legal LKAU the employee is employed. The information is used to make very good employment statistics, but the information is also used to see if a legal LKAU has been taken over by a new employer so that the corresponding statistical LKAU can be transferred to the new enterprise.

Hopefully, by 2015 or 2016 it will be compulsory for the businesses to report ownership of other businesses to the Ministry of Businesses, and this information will be available to SD. SD has asked for a very detailed ownership groupings (below 5 %, 5-10 %...) but the actual groupings have not yet been decided.

E1.12 Lessons Learned

The Danish experience suggests that three things are very important:

- identification numbers;
- keeping things simple, having information on administrative registrations and keeping track of registration history; and
- cooperation with public administrations and business organisations.

Common identification

The most important thing is the identification numbering system. The best situation, of course, is if one common identification number system for legal units is used by all organisations. Then different administrative identification numbers can be connected to this single legal unit identification number.

Keep things simple

It is important to keep things simple. This means not mixing everything together but keeping things separate. In particular, administrative information should be kept *untouched* and connected to administrative units. Business Register processes can then transform this information into relevant information about statistical units.

It is also very important to know when information has been updated, from what source, and to what point in time the information relates. If information is received about a change in activity code it is important to know if this change relates to the current year or the previous year.

Cooperation with administrative sources

Cooperation is very important. Even though it may seem that nothing is happening, progress will be made over time if the arrangements are good. The nomination of the SD General Director as a member of the board of *Datacentralen* (a publicly owned data processing centre that has been chosen as the technical operator for the new administrative registers) was a very important decision for SD and emphasised the central role given to SD in this context.

As previously mentioned, a committee was set up in 1970 by SD with members from several ministries to report on centralisation of business registration. It was already clear that common identification of the units was central. Many subjects were discussed, among others the legal form, which, in some cases, was regulated by law, and local units.

Later, in 1982, and over the period 1992-1995, the role of the Central Administrative Business Register was discussed again, and SD played a central role when this register was separated from the Statistical Business Register in 1999. It was decided that SD would keep this Central Administrative Business Register from where identification information could be drawn and delivered to whoever was interested on a service basis.

Updating frequency

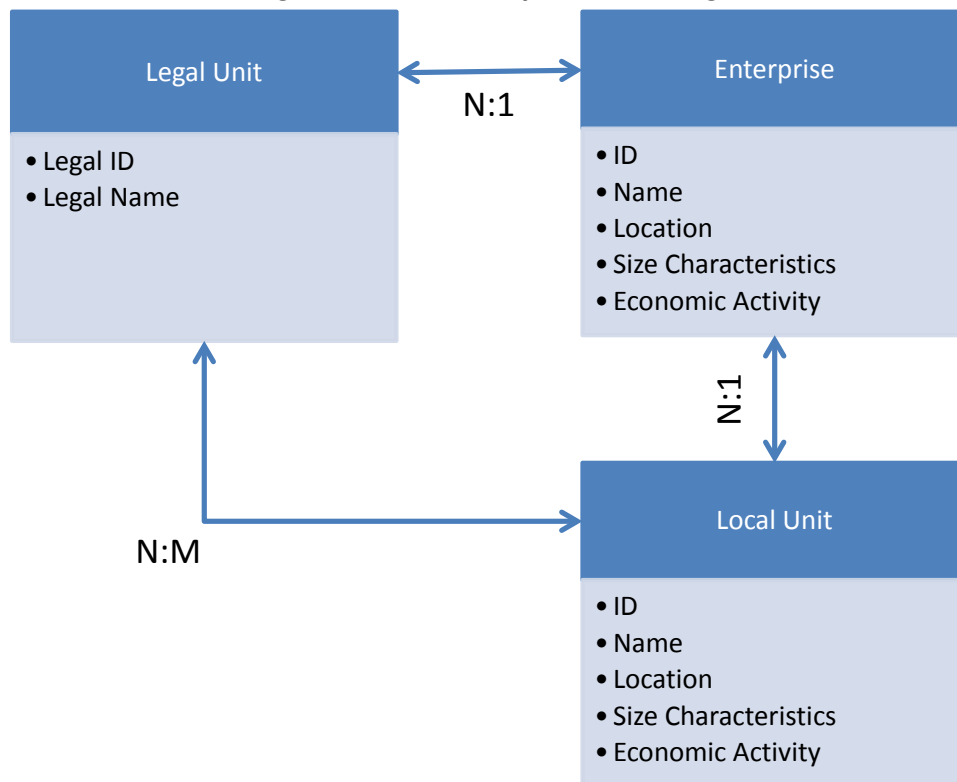
As already mentioned, the 1966 law about SD gave SD access to the various registers. SD used to get information from the VAT Register each month and from the Income Tax Register, including the payment to the wage earners' supplementary pension, each quarter. Now this information is delivered each week. SD also gets information about each employee connected to each legal LKAU each month, and about monthly turnover, exports, imports etc. Since 1999 when the current Central Administrative Business Register (CVR) was established, SD has been receiving updating information when the CVR is updated, though this is dependent to a large extent on self-registration and updating.

Annex E2: Costa Rica’s SBR programme

E2.1 Overview of Costa Rica’s SBR Programme

The Register of Enterprises and Establishments (REE) in Costa Rica is a structured registry of the resident institutional units (*enterprises*) in the private sector in Costa Rica, and their local units (*establishments*), engaged in the production of goods and services. It includes information on the characteristics of each unit, such as identification (ID) characteristics, location, economic activity and size. Institutional units are enterprises associated with one or more legal units. They have unique ID numbers assigned by the National Institute of Statistics and Census (INEC). Each enterprise is further divided into one or more local units (establishments). Each legal unit of an enterprise can be further associated with one or more of the enterprise’s local units. On the other hand, every local unit of an enterprise is associated with at least one legal unit. Thus, two or more local units of the same enterprise can be associated with the same legal unit. This is illustrated in Figure 1 below.

Figure 1. Structure of Business Register



The REE provides an up-to-date sampling framework with the broadest possible coverage for compiling economic statistics in the country. It enables analysis of business dynamics of the different economic activities that comprise the Costa Rican economy in terms of births, deaths, mergers, temporary closures and linkages.

The main results from the process of maintaining and updating the REE database are published every year. In 2013, for example, there were 43,847 enterprises and 10,775 local units in the REE.

Costa Rica’s REE programme at the INEC is the responsibility of a team comprising a project coordinator, an IT specialist, two professional staff, and five data collectors.

E2.2 Background

The creation of the REE was a direct response to the need to improve and broaden the scope of the country's economic statistics, and particularly to address the lack of adequate information on the structure and dynamics of enterprises located in the national territory.

The first statistical business register that was created in Costa Rica was based on the 1990 national census. At that time there was no system in place for updating the register, and it became quickly obsolete in view of the dynamics that characterized its statistical units.

A project proposal for the creation of the REE was drafted in 2007 as a response to the lack of an adequate source of information on the structure and dynamics of Costa Rican businesses. A Framework Cooperation Agreement was then signed in 2008 between the Central Bank of Costa Rica and INEC to create the REE, and was implemented between March 2008 and April 2009.

The field operation for the development of the REE was carried out in 39 districts (the smallest geographic unit in Costa Rica) containing more than 50 percent of the business that existed in the country at the time⁹⁰. Information for the remaining districts was obtained from administrative records of the Social Security administration. This approach was in line with other countries' experiences, which show that the use of administrative sources facilitates the process of obtaining data for updating and maintaining the register in a timely and cost-effective manner by leveraging resources already available. The project also included an assessment and analysis of the coverage and quality of administrative records.

E2.3 Updating Process

The process for updating the REE relies on three mechanisms. First, telephone interviews are carried out to obtain basic characteristics such as identification characteristics, location, size, and economic activity. This mechanism is supported by a customized information system developed in-house that allows updating of the REE database during the interview itself.

The second updating mechanism consists of the use of administrative records from both internal and external databases, including:

- **external sources:** Ministry of Finance, Agency for the Promotion of Foreign Trade, Electoral Register, National Registry and Social Security Register;
- **internal sources:** Consumer Price Index and Construction Price Index databases, as well as the National Enterprise Survey.

A third updating mechanism, implemented since 2012, is the collection of data based on an enumeration exercise that covers one district at a time. This enumeration of local businesses is supported by the use of hand held *Personal Digital Assistants* (PDAs) - computer devices - to update all the key characteristics in the REE. The main purpose of this exercise is to assess the coverage of the REE and the increase or decrease in the number of businesses in the district.

⁹⁰ This estimate is based on data from the Social Security administration.

In addition to the three methods described above, special data collection projects that are carried out by INEC for other government agencies provide inputs to update the records of the REE.

Approximately 40 percent of the records in the REE are updated every year. Priority is given to those enterprises that have not been interviewed for a long period of time. This process of continuously updating the REE is crucial in providing an accurate picture of the current economic situation of the country and an adequate sampling framework for statistical surveys. Having a central business register avoids duplication of work and the proliferation of fragmented datasets collected by different users or agencies on an ad-hoc basis.

E2.4 Unique Identifier

Costa Rica has a unique identifier for each natural person and for each legal entity. The ID number for natural persons (“*cédula física*”) has nine digits and is required for all official administrative procedures related to social security, obtaining a passport, etc.

The ID number for legal persons (“*cédula jurídica*”) is assigned by the National Registry to legal entities for paying taxes, complying with social security regulations, engaging in buying and selling operations, etc. It has ten digits; the first four digits identify whether it refers to a corporation, an association, a cooperative, a foundation, a foreign enterprise, etc.

E2.5 Administrative Records

Social Security Register

The social security register (SSR) consists of each month’s payroll, with details on the number of employees, salaries, and economic activity. It includes information obtained from employers, self-employed persons, individual employees, and persons employed under a special (collective) contract.

The SSR is used to improve the coverage of the REE by means of a data integration process. The first step is to reclassify data from the SSR database of employers that refer to geographic areas and economic activities (according to ISIC Rev. 3) to INEC’s statistical classification of territories and ISIC Rev. 4.

Once this reclassification is done, an assessment of the SSR database is carried out to determine which characteristics are relevant for the REE.

Finally, the SSR database is linked with the REE database using either the unique ID numbers (for natural or legal persons) or the passport number in the case of foreign persons. Non-matching records are analyzed in detail to check whether the enterprise is already included in the REE under another ID number (to avoid duplicates), or whether it is inactive, etc. Those records from the SSR database that are not found in the REE database are added as new records, and are then immediately updated to the extent possible by means of follow-up phone interviews.

Records from the Ministry of Finance

Administrative records from the Ministry of Finance contain data on the taxable income of natural and legal persons, over a specific period of time, collected by the Tax Revenue Administration. This information is used for two specific purposes: first, to update the data of large taxpayers and improve the coverage of the REE by adding new records for large enterprises whose ID numbers are not found in the REE; second, to identify large or complex

enterprises based on their level of taxable income and their number of employees by economic activity.

Register of the Agency for the Promotion of Foreign Trade

The Trade Intelligence Division of the Agency for the Promotion of Foreign Trade (PROCOMER) provides INEC its register of exporting enterprises every year, within the framework of a joint programme to carry out a Census of Exporting Enterprises. The goal of this programme is to measure the number of jobs created by the exporting sector and to assess the share in total exports of different categories of enterprises by size.

The PROCOMER register of exporting enterprises is used to improve the coverage of the REE on the basis of a comparative analysis that allows identification and addition of new records for exporting enterprises not found in the REE.

National Registry and Electoral Supreme Court databases

The National Registry is the institution responsible, among other things, for the cadastre and registry of real estate, industrial, and other property, as well as for the official registration of legal persons. On the other hand, the Electoral Supreme Court is in charge of civil registration, i.e., registration of all events that are relevant from the point of view of civil law (births, marriages, divorces, deaths, emission of personal ID number, etc).

These two databases are used for cross-validation purposes. Specifically they are used to verify the existence of the unique ID numbers (for natural and legal persons) of the enterprises included in administrative records, and to verify whether the legal name of an enterprise in the REE corresponds to the name assigned by the National Registry or the Electoral Supreme Court.

Consumer Price Index and Construction Price Index databases of INEC

The list of enterprises providing data to INEC for the Consumer Price Index and Construction Price Index programmes is periodically used to update the REE. With the help of an information system, the data collected through the price surveys are compared to the information in the REE database, thus avoiding the need to conduct additional phone interviews.

INEC's National Enterprise Survey Framework

The National Enterprise Survey is conducted on a quarterly basis on a sample of enterprises from the private sector engaged in various economic activities throughout the country. It collects data on the number of enterprises, jobs, hours worked, and salaries. Its main purpose is to provide information on the situation of Costa Rica's labour market from the point of view of the enterprises.

It also provides an opportunity to update the information of enterprises that belong to the same enterprise group. More specifically, in some cases it is possible to obtain from a single interview, data on the various individual enterprises that belong to the same group, even if not all of them are included in the sample. This also helps to improve the coverage of the REE by identifying the enterprises that are not yet included in its database.

The REE is updated with data collected through the National Enterprise Survey with the help of a special module, which allows comparison of basic data of the surveyed enterprises with the current contents of the REE database.

E2.6 Cooperation Agreements with Sources of Administrative Records

An inter-agency cooperation agreement between the INEC and the Ministry of Science and Technology provides for the electronic exchange of data and mutual technical assistance with respect to the collection, use, and maintenance of statistical data for the *National Survey for the compilation of Indicators on Science, Technology and Innovation*.

Within the framework of an agreement between INEC and the Institute for Research in Economics (IICE) at the University of Costa Rica (UCR), INEC provides the sampling framework for the *Survey on Business Expectations*, of which the purpose is to determine the expectations of business executives regarding the current economic situation in their respective industries. The sample includes businesses from agriculture, manufacturing, construction, wholesale and retail trade, and other service activities.

A number of other inter-agency agreements have been established between INEC and the Ministry of Finance, the Agency for the Promotion of Foreign Trade (PROCOMER), and the Ministry of the Economy, Industry and Trade. These agreements, which reflect the good relationships that exist with these institutions, allow exchange of information to improve the operation of the REE programme.

Annex E3: Statistics Canada's Statistical Business Register

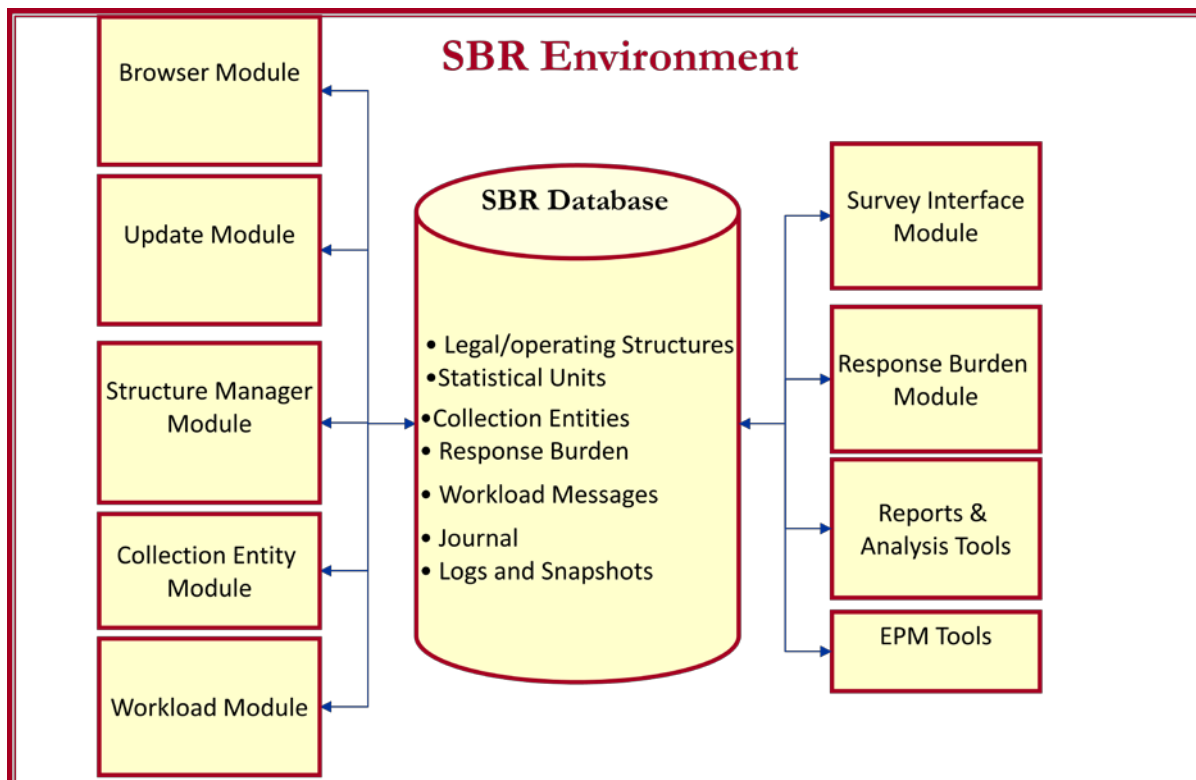
E3.1 Introduction

When Statistics Canada redesigned its statistical business register (SBR) system over the period 2007-08, it built a Microsoft Windows application installed on a client PC using a service-oriented architecture. There are five major components comprising the system, namely:

- VB.Net is the programming language for the Windows Forms user interface, the business layer and the data layer;
- SQL Server 2005 is the underlying database that both stores and manipulates the data;
- SAS is used to crunch and manipulate input data from external sources;
- the system is message based and uses BIZTALK for routing the messages; and
- Web services are used to manage both security and access to the data.

All users access the SBR by means of the same common interface with a privilege administration tool as the control mechanism to manage this access. The SBR includes information on legal and operating units and their structural relationships. It maintains a journal/log of all updates applied to these units as well as to stratification characteristics and information on reporting arrangements.

Seven modules within the register manage different aspects of this information.



E3.2 Browser Module

The Browser Module allows the user to browse and search for information on a given enterprise. It displays information such as the business structure, collection entities, the response burden, and the history of updates contained in the Register (Journal, Log and Snapshots).

The Log contains all the updates performed on a given variable. The Journal records significant events (e.g., amalgamation, dissolution) concerning a given business. The monthly Snapshots show the image of the business structures at regular points in time in the past (previous months).

E3.3 Update Module

This component serves to control and manage all updates, both manual and batch, that need to be applied to the database. The SBR receives requests or signals for updates from various sources such as subject matter, collection and external administrative sources. Although each may follow a slightly different process, in general a request is vetted to determine if human intervention is required (Workload) and if accepted, it is applied to the database. Whether a request is rejected or accepted, the systems will send out notification to ensure that the affected parties are aware of the status of the request.

E3.4 Structure Manager Module

The Structure Manager Module is used to show complex structures via the Browser. It manages these enterprise structures and the links that exist between production entities. It manages and controls the parent–child links, propagates attributes within the structure, and checks the coherence of the structure once changes have been made.

E3.5 Collection Entity Module

This module is responsible for generating and updating collection entities (containing information used in contacting respondents) based on the information contained in a survey control file received from the survey's sampling process. It also manages the manual customization of collection entities that is performed by survey managers based on pre-established business rules.

E3.6 Workload Module

When an update signal arrives at the SBR that requires manual investigation/verification, the Update Module generates a signal to the Workload Module indicating the need for review by an analyst/profiler. The Workload Module manages, prioritizes and assigns the signals to the analysts/profilers. After manual review of a signal, the analyst/profiler either implements or cancels the corresponding change request.

E3.7 Survey Interface Module

The function of this module is to produce two key SBR products that are necessary to SBR partners in conducting their surveys: the frozen frames for sampling; and respondent information file for collection.

The monthly standardized frozen frame is the result of extracting all units that comprise the total business population. It contains a list of all units of production with their tombstone information, the industrial classification, the detailed geographical code, the size variables (such as revenue and employees), and other information to satisfy sampling procedures. Survey methodologists use this file primarily for the generation of survey samples. It is also used by subject matter divisions as an input to their edit, imputation and estimation system.

From the units selected for a given survey by survey methodologists, the SBR provides the respondent information file to collection staff. This file contains the information required in order to carry out data collection, such as the contact name, address and telephone number.

E3.9 Response Burden Module

This module presents all information relating to respondent burden for economic surveys. It displays information about all contacts that the NSI has had with any given enterprise. The response burden tool displays this information by survey, enterprise, contact name, and questionnaire. It also calculates the actual burden per enterprise. Finally, it provides extractions concerning exclusion orders and cases that require specialized treatment. A central frame used by the entire economic survey program means a truly comprehensive view of response burden and thus facilitates its management.

E3.10 Reporting and Analysis Tools

These tools produce the reports needed to manage survey operations and analyze sub-populations. The Survey Frame Assessment (SFA) tool extracts and presents changes that have occurred in a survey's population between two reference points. It includes births, deaths and changes to industry codes and size indicators such as revenue. The SFA tool can dynamically display all changes that occurred as of the previous day for selected characteristics by operating entity. This is of particular importance as changes can be reviewed immediately prior to the production of the monthly frozen frame. Other tools include the analysis of updates to the SBR and demographic analysis of the business population.

Annex E4: Business Register at the National Statistics Office of Georgia

Activities for the development of the Business Register (BR) at the National Statistics Office, Georgia (Geostat) started with evaluation and identification of gaps and specific problems. The evaluation resulted in identification of several serious gaps, such as an incomplete database, several missing characteristics, outdated software (Paradox database developed in 1990s, with no possibility of adding new fields), lack of database structure, and inefficient updating procedures (receiving data on paper and manual data entry).

At a later stage it was necessary to identify possible internal and external sources for updating BR and to obtain full access to the relevant administrative sources. In this regard a number of working meetings were held with various public institutions that provide administrative data. In addition, in response to the Geostat's initiative, the Government adopted a resolution, according to which the administrative authorities are obliged to provide regularly required information to Geostat. According to a Government Resolution, approved in February 2011, and amended in July 2011, the Revenue Service (Tax Office) is obliged to submit monthly data to Geostat on active tax payers (that are enterprises). In this context, an enterprise is considered active if it has indicated;

- turnover more than zero in any kind of declaration (VAT, revenue, etc);
- number of persons employed more than zero in any kind of declaration (VAT, revenue etc.);
- wages or number of persons employed in a wage notification;
- profit or loss in a profit declaration; and/or
- any kind of tax payment with the exception of property tax.

Active enterprises report data on turnover and number of employees to the Tax Office on a voluntary basis. Thus, the Tax Office does not receive these data from all enterprises. It submits the data it does receive to Geostat.

As a result of these activities, Geostat has direct access to administrative sources. In most cases, high quality statistical products are obtained by combining administrative and survey data.

Other sources for updating the BR include various business statistics surveys. Quarterly and annual business surveys provide information on contact details and type of economic activities of the surveyed entities.

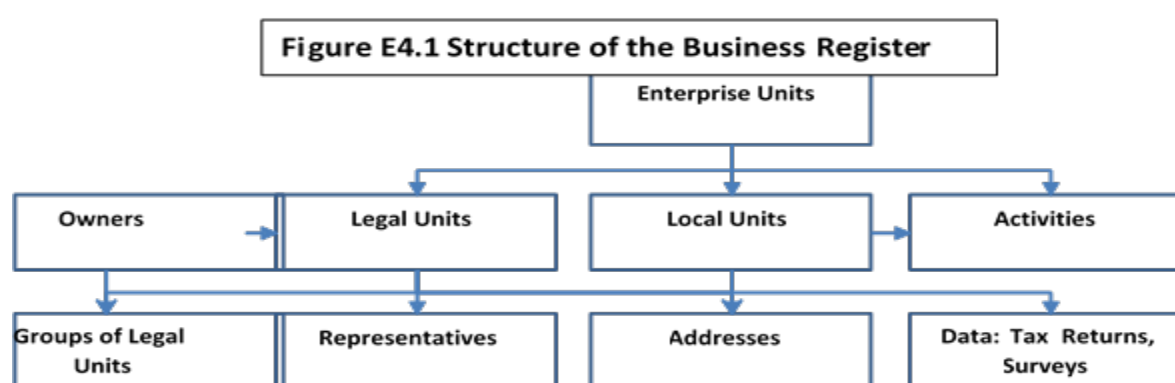
International experts were engaged in the above mentioned activities in order to ensure international recommendations were taken into account and relevant methodologies for the formation of BR in line with international standards were introduced. In this respect cooperation with Statistics Sweden is noteworthy and greatly contributed to the development of the BR.

The Cooperation Agreement between Geostat and Statistics Sweden was signed in June 2011. The project is funded by the Swedish International Development Agency (SIDA) and covers a four year period. One of the components and goals of this cooperation is improvement and development of the BR in Georgia. As a result of this cooperation with the Swedish experts, a BR maintenance strategy has been developed that covers use of administrative sources, use of survey results, assessing needs of users, planning of updates, and rules for dealing with BR information.

Various administrative and internal data have been collected, unified and processed, resulting in the formation of a new database, namely the BR. At the moment the database contains approximately 581,000 entities (both active and non-active entities). It includes all registered economic entities, including physical persons. There is almost full coverage of legal entities, but full coverage of local units, local kind-of-activity units and enterprises has not yet been achieved.

Based on information received from the Revenue Service (Tax Office), Geostat has developed a methodology for creating a sampling frame for business surveys. This sampling frame covers active enterprises from the non-financial corporation sector. It may be used for all business surveys in Georgia.

The structure of the BR, as shown in Figure E4.1, is in accordance with international standards and recommendations.



One of the major factors in the development of the BR was the software. In this regard new user-friendly software has been developed. The database is operating in MS SQL Server and interface in Web application (PHP).

It was also important to establish a flexible and automated system for updating the BR. The main sources are administrative sources, namely the National Agency of Public Registry and the Revenue Service (Tax Office).

- From the Public Registry, Geostat receives information on newly registered and liquidated (closed) companies, as well as information on changes.
- From the Revenue Service (Tax Office), Geostat receives information on active taxpayers.

Update procedures are automated. The information is received from administrative sources monthly, using automatic interfaces that the Geostat has with the databases of these institutions. Received data are cleaned and structured. The statuses of the companies are identified. This information received from the administrative sources is automatically reflected in the BR.

Geostat also uses results of various business surveys to update the BR. It conducts special surveys that are directly aimed at updating the BR and improving its quality.

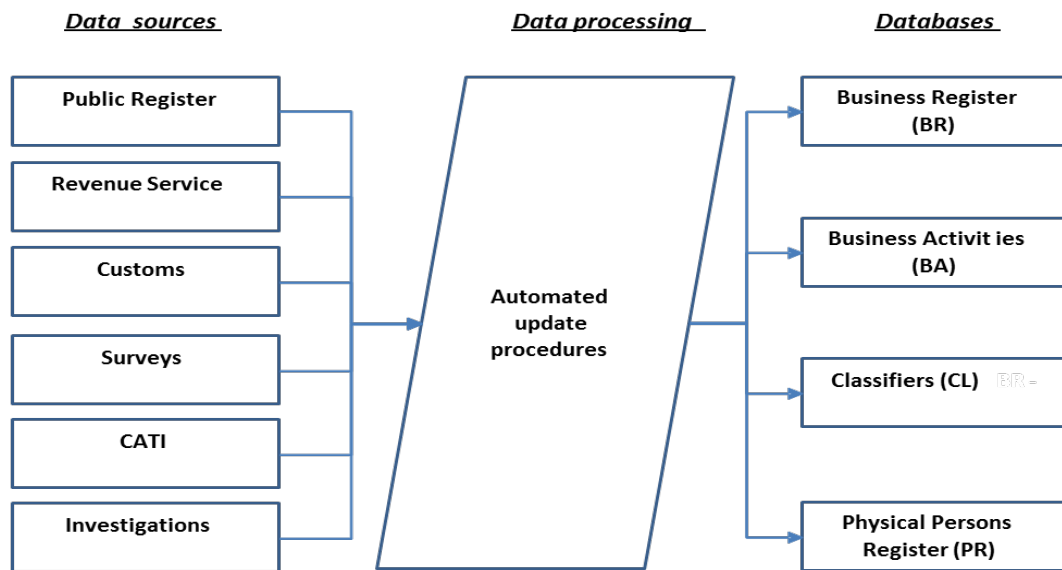
Figure E4.2 shows the current BR updating system.

In the data received from administrative sources the actual address and type of economic activity are specified only for a limited number of enterprises. Furthermore, in some cases, the information on economic activity is incorrect. This was a major problem for the BR and

hence for all business surveys. Thus, since January 2013 the Geostat has been conducting a monthly survey, using computer assisted telephone interviews (CATI), of companies included in the BR to update information about the status, kind of economic activity and actual address of every local unit of the enterprises.

Relevant documentation is also essential for the users of the BR and is a part of the overall quality of the service the BR provides. Thus, Geostat created a metadata document for the BR that contains comprehensive information about the BR and brief description of the characteristics, production, updating and maintenance procedures, etc. In addition, it contains information about the rules for dealing with BR.

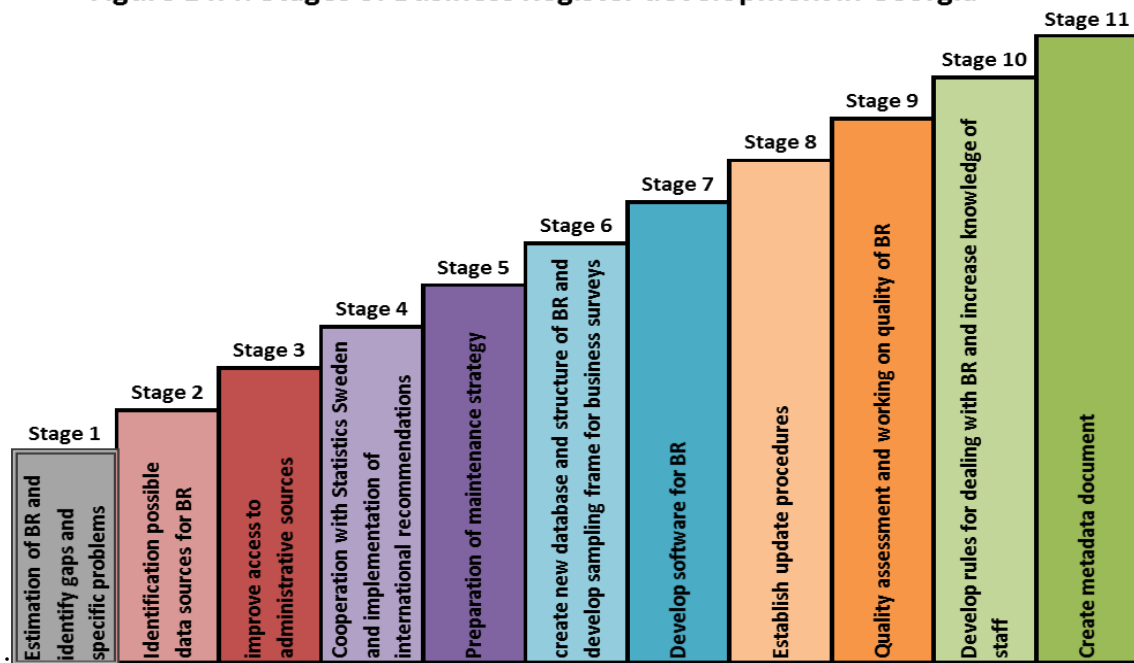
Figure E4.2: Data Flow of the Business Register



The relevant division in Geostat is continually working to maintain the BR and to improve its quality. Figure E4.4 shows the stages of BR development.

It is of major importance to share experiences with other countries and improve capacity of staff through participation in trainings and workshops.

Figure E4.4: Stages of Business Register development in Georgia



Annex E5. Redeveloping the Malaysian Statistical Business Register

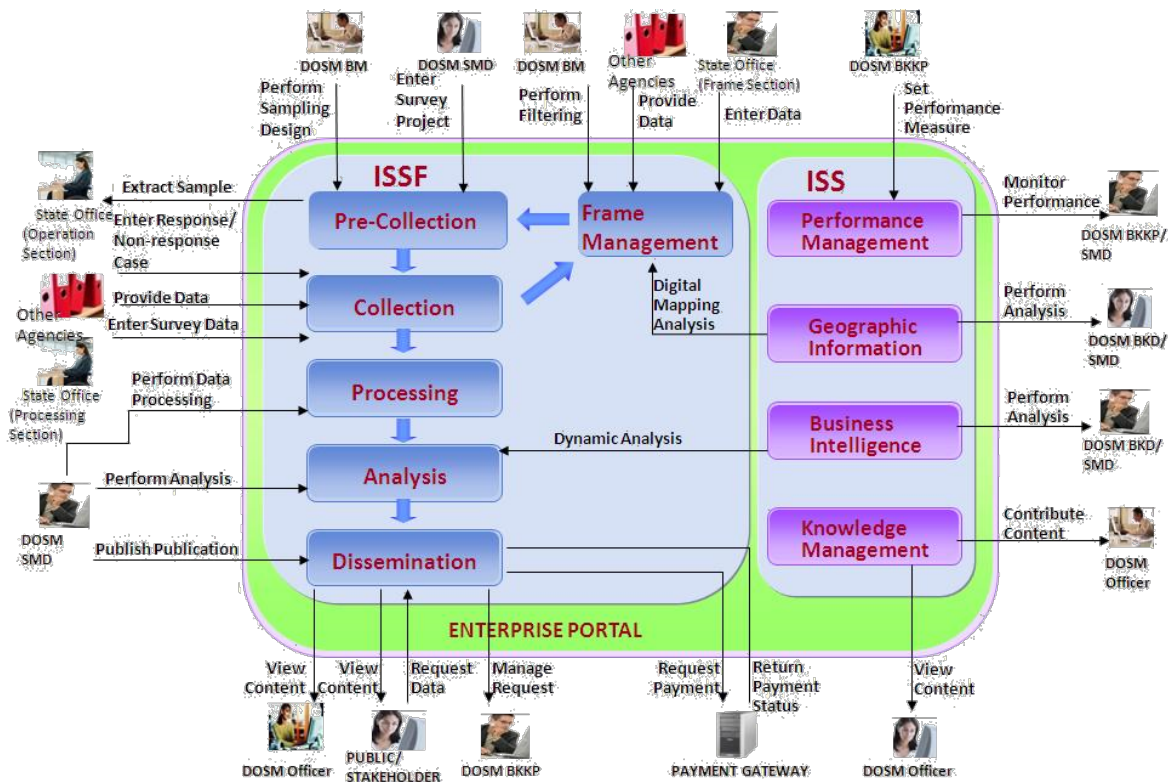
E5.1 Brief History of Business Register in Malaysia

The Central Register System (CRS) of Department of Statistics Malaysia (DOSM) was set up in 1994 in a PC based system using Dbase IV. It was upgraded in 1996 into a local area network (LAN) environment. Further enhancements were made in 2002 with an upgrade to a wide area network (WAN) environment as well as incorporation of features needed by users, requirements for new surveys and accommodation of new industrial codes.

E5.2 Streamlining Business Processes

DOSM has implemented an *Integrated Statistical Systems Framework (ISSF)* to ensure the quality in the statistical production process. It is an integrated online system to streamline statistical business processes and replace non-integrated systems. It is in line with the Generic Statistics Business Process Module (GSBPM) developed by the UNECE, which provides a basis for standard terminology on statistical metadata systems and processes (Figure E5.1).

Figure E5.1: Integrated Statistical Systems Framework



The ISSF is a flexible tool for describing and defining the set of business processes needed to produce official statistics. It can also be used as the basis for quality assessment of censuses, surveys, administrative data and data from other sources. The processes have been articulated within the ISSF module, which can be accessed by the head office and the state offices.

The ISSF provides the main business functionalities of DOSM for pre-collection, collection, processing, analysis and dissemination. The web browser will be the front end interface to this integrated system and will provide wide access capabilities anywhere and anytime through web based application. The *business register (BR)* which is known as *Establishment/Enterprise (EE) Frame* will be supporting the pre-collection and collection module.

The modules for implementing the BR are shown in Figure E5.2:

<i>Figure E5.2: BR Modules and Functions</i>	
<i>Module</i>	<i>Function</i>
Screening Agency information Business Profiles	Maintenance through list information from administrative sources
Establishment/Enterprise Duplicate Data Collection Respondent Management Control	Maintenance of establishment and enterprise information
Report Business Intelligence (BI)	To generate monitoring reports and analysis tools
Geographical Information System (GIS)- EE	To monitor the number/type of establishments in an enumeration block

E5.3 Integration of the BR with the Collection System

In the ISSF platform the BR is integrated with collection processes through the *Operational Control Information (MKO) module*. The MKO module is being designed to allow the subject matter experts and enumerators to update basic business information in the frame before clean data from the survey/census are processed. This increases the efficiency and timeliness of the updates. Comprehensive, reliable and timely BR coverage and content increase data collection efficiency and assist the state offices in performing their updating activities.

The first benefit is accessibility by enumerators of the most recent and detailed information in the BR. Having access to comprehensive and timely information facilitates interpretation of information received from the respondents. It assists the department in improving respondent relations and improved respondent relations lead to more comprehensive and better quality responses.

E5.4 Future Challenges

Enhancing the Use of Administrative Data Sources

While data from primary sources continue to play leading role in producing economic statistics, usage of administrative data is increasing. Thus efforts must be made to enhance the cooperative arrangements between DOSM and administrative agencies.

Furthermore, the DOSM must address barriers in terms of quality associated with the use of administrative data as basis for official statistics with respect to coverage, timeliness, frequency, validity, reliability, consistency, legality and confidentiality constraints.

The DOSM should seek new sources of administrative data. In this respect, one of the steps to be taken is to engage with agencies involved in the *Malaysia Corporate Identity Number (MyCoID)* introduced by the Company Commission Malaysia. MyCOID refers to the company incorporation number which is used as a single source of reference for registration and transaction purposes with other relevant Government agencies. These will be the future platform for coordinating and acquiring data from these agencies.

There is also potential to replace some survey data with administrative data obtained from other sources. These would reduce the burden on respondents and the operational costs incurred by the office.

Awareness among administrative agencies is vital in developing the uses of their data sources. This can be done through meetings, seminars, training, and conferences.

Through ISSF, the DOSM must monitor the quality of incoming administrative data.

Keeping the BR Accurate and Up-to-date

The BR should provide information that is as close as possible to the situation in the real economic world. Ways in which to do this include:

- improving access and use of identification and matching data;
- improving updating from survey feedback;
- enhancing the use of administrative data sources; and
- establishing area frame development and maintenance procedures.

Reducing Respondent Burden

The BR contains many records as it is the main medium for survey frame creation and samples selection. It should enable control of the number of establishments and enterprises that are involved in censuses and surveys. Integration between BR and MKO will assist the enumerators to better manage the respondents and ultimately to ensure that the statistics collected are timely and reliable.

Business Demography

There is a growing demand for business demographics, which involves statistics for specific events such as birth, death, survival and growth. Together with Company Commission Malaysia, DOSM is conducting research on the relevant terminology, scope, coverage and data availability.

Business Profiling

Business profiling is the method for analysing the legal, operational and accounting structure of large enterprises and enterprise groups. It provides better understanding of the complex structures of the enterprise groups in the country. Profiling can be viewed as part of a broader, coordinated strategy for improving agency's economic statistics. DOSM has taken

step in this direction by creating an enterprise module in the ISSF system that can facilitate profiling activity.

Annex F: Examples Relating to SBR Quality Assurance

Annex F1: Istat's SBR Quality Declaration

Annex F2: The Colombian Experience in the Implementation of Quality Processes in the Statistical Directory of Enterprises

Annex F3: Statistics Netherlands' Administrative Data Source Evaluation Checklist

Annex F1: Istat's SBR Quality Indicators

F1.1 Introduction

The Italian NSI (Istat) maintains an SBR, referred to as the *Business Register of Active Enterprises (ASIA)*. Based on the principle of transparency, Istat produces a declaration of ASIA quality to accompany ASIA viewed as a micro-data file. The declaration comprises a set of indicators intended to measure the various quality components. It is a synthesis of indicators, direct and indirect, that give a temporal reference to the data, sources and variables. It is easy to read and interpret in terms of users' needs. It covers several aspects:

- d) use of metadata contained in the archive's database;
- e) compilation of indicators, presented in the form of tables, graphs, histograms, etc.;
- f) calculation of synthetic indicators;
- g) selection of priority indicators.

The following paragraphs describe 50 quality indicators that can be calculated for each reference year, including the various breakdowns of these indicators by selected characteristics. They cover the basic data from the administrative files that provide the inputs to ASIA as well as the statistical data on the output side, i.e., ASIA enterprises.

Comparisons over time (two or more consecutive years) and interpretation of increases and decreases in the indicators provide additional information on quality.

In the following paragraphs

- t = reference year of ASIA data
- s = administrative source of input:
 - Chamber of Commerce (CCIAA);
 - Fiscal register (MEF);
 - Social Security (INPS);
- s_t = data from source s for reference period t .

F1.2 Quality of Inputs

Quality Criterion: Timeliness

The indicators measure the time lag in the delivery of the administrative data to the SBR for each source. The time lag is the difference between the date on which the data are supplied and the reference period to which they refer.

- 1) Temporal lag, measured in months, between supply date (s_t) and reference year t , compiled for $s_t = \text{CCIAA}$;
- 2) Temporal lag, measured in months, between supply date ($s_t = \text{MEF}$) and reference year t , compiled for $s_t = \text{MEF}$.

Quality Criterion: Coverage

- 3) Number of records received from supply s .

Measurement of the completeness of enterprise births and deaths is useful in detecting under-coverage and over-coverage. It is better to use comparisons of data from a single source over

time since simple counts of the numbers of births or deaths during a reference period do not of themselves provide much information.

The effect of update delay of the source is obtained by comparing the values of a characteristic - the number of cessations in year t (*Ncess as in indicator 4*) or the number of starts of activity in year t (*Nstart as in indicator 5*) – in the yearly supplies received at time t (s_t) and at time $t+1$ (s_{t+1}) – each case referred to the year t .

The two simple indicators to estimate the lag in the registration of the two dates in input source $s = \text{CCIAA}$:

- 4) Loss of information about cessations occurring in year t using the supply s_t :

$$1 - \text{Ncess}[\text{St}+1(t)] / \text{Ncess}[\text{St}(t)]$$

where

t = year of cessation

$\text{Ncess}[\text{St}(t)]$ = number of cessations occurring in year t for which data was received in year t from source s .

$\text{Ncess}[\text{St}+1(t)]$ = number of cessations occurring in year t for which data was received in year $t+1$ from source s .

- 5) Loss of information about starts of activity occurring in year t using supply s_t :

$$1 - \text{Nstart}[\text{St}+1(t)] / \text{Nstart}[\text{St}(t)]$$

where

(t) = year of starting activity

$\text{Nstart}[\text{St}+1(t)]$ = number of starts occurring in year t for which data was received in year t from source s .

$\text{Nstart}[\text{St}(t)]$ = number of starts occurring in year t for which data was received in year t from source s .

Quality Criterion: Completeness

The indicators measure the completeness of characteristics, including the numbers of missing values and/or number of missing values as a proportion of the total number of values, for selected characteristics.

- 6) Company name, $s = \text{CCIAA}$: Number and % of records with missing company name.
- 7) Legal status, $s = \text{CCIAA}$: Number and % of records with missing legal status.
- 8) Address(es), $s = \text{CCIAA}$: Number and % of records with missing address(es).
- 9) Principal economic activity code, $s = \text{CCIAA}$: Number and % of records with missing principal economic activity code.
- 10) Territorial (municipality, province) code, $s = \text{CCIAA}$: Number and % of records with wrong territory code.
- 11) Economic activity code, $s = \text{MEF}$: Number and % of records with missing economic activity code.
- 12) Registered head office abroad, $s = \text{MEF}$: Number and % of records with missing information about registered head office abroad.
- 13) Employees, $s = \text{INPS}$: number (%) of records with zero employees.
- 14) Economic activity code with classification not up to date, $s = \text{MEF}$: Number and % of records with classification code according to obsolete classification.

F1.3 Quality of Processes

Quality Criterion: Coverage

Information from the first macro-phase (Integration of administrative sources and identification of units)

Records coming from different sources that pertain to the same legal unit (i.e. identified by a common taxation identification number) are integrated in order to build-up a *cluster* of records for the same enterprise. The taxation register is, typically, the pivotal source, i.e., the base used to define the set of legal units and to integrate all the other sources.

- 15) Missing matchings: Number and % of records in source s (\neq MEF) not matched with $s =$ MEF by tax code.
- 16) Undercoverage due to lag: Number of *clusters* of administrative records indicating units not matched with MEF(t-1)* but then matched with MEF(t) and as % compared with number of *clusters* in MEF(t-1)
(*These are units that could have been linked and would have been included in the SBR in time (t-1) if the tax register had successfully updated such units in time).
- 17) Structure of *clusters* by number of matching sources: Number and % of *clusters* consisting of 1, 2, ... n sources.

Quality Criterion: Accuracy

Information from the first macro-phase (Integration of administrative sources and identification of units)

- 18) Economic activity code (NACE): number of records for $s =$ MEF with an old NACE classification, not coded using the new NACE codes as proportion of the total number of records with an old NACE classification.

Information from the second macro-phase (Estimation of characters)

- 19) Activity status: number of units (and/or related employees) with activity status estimated using administrative sources, then modified by deterministic (automatic) rules.

Information from the third macro-phase (Integration with statistical sources and quality control)

- 20) Number and % of units with erroneous NACE or activity status or number of employees.
- 21) Number of units (% of total errors) with erroneous NACE or activity status or number of employees then automatically corrected by standard deterministic rules.
- 23) Number of units (% of total errors) with erroneous NACE or activity status or number of employees then corrected manually by skilled BR staff (using clerical checks and on-line updating).
- 24) Number of units (% of total errors) with NACE or activity status or number of employees still to be ascertained and then accepted.

F1.4 Quality of Outputs

The following indicators are calculated by size class in terms of number of employees in order to give different priorities to corrections.

Quality Criterion: Coverage

- 25) Number of enterprises active in t and as change from t-1.
- 26) Number of reactivations in t as a percentage of active enterprises.
- 27) Number of units (in terms of enterprises and employees) by source (survey, estimation, profiling) by characteristic (NACE, employees and status of activity).

Quality Criterion: Completeness

- 28) Incomplete NACE codes: Number of units with NACE 2-digit code; Number of units with NACE 3-digit code; Number of units with NACE 4-digit code).
- 29) Company name: Number of units with company name missing.
- 30) Address: Number of units with address missing.
- 31) Postal code: Number of units with postal code missing.
- 32) Telephone: Number of units with telephone number (fax, email) missing.

Quality Criterion: Timeliness

- 33) Temporal lag, measured in months, between the dissemination date of ASIA (i.e., date when ASIA data are made available to users) and the reference year to which they refer.
- 34) Latest information (with reference date later than t when disseminating data with reference year t): Number of start dates later than t ; Number of cessation dates later than t; Number of events (by type) having starting date later than t ; Number of units with employees updated at the time later than t.

Quality Criterion: Accuracy

Direct indicators

Comparison between BR data and SBS Small and Medium Enterprise Survey data concerning variables that are pre-printed in SBS questionnaires using BR data:

- 35) Number and % of questionnaires rejected by type of error.
- 36) Number and % of units with wrong address (trend over the last x years).
- 37) Number and % of units with erroneous activity status (trend over the last x years).
- 38) Number and % of units with conflicting economic activity.

Comparison between BR data and Survey on the local units of big enterprises (IULGI) (concerning variables that are pre-printed in IULGI questionnaires using BR data):

- 39) Number and % of questionnaires rejected.
- 40) Number and % of units with modified NACE.
- 41) Number and % of units with modified and corrected NACE (by manual check).
- 42) Number and % of units with modified and erroneous NACE (by manual check).
- 43) Number and % of units with inconsistent employees
- 44) Number and % of units with inconsistent and corrected employees (by manual check)
- 45) Number and % of units with inconsistent and erroneous employees (by manual check)
- 46) Number and % of units with modified activity status
- 47) Number and % of units with modified and accepted employees (by manual check)
- 48) Number and % of units with modified and erroneous employees (by manual check)

Indirect indicators:

- 49) Status of activity: analyses of time series of active enterprises of ASIA (t, t-1 and t-2) and measure of the reliability of selected sub-populations (defined by items like registrations, register deletions, reactivations, etc.). The indicator is also calculated by cell, like sector of economic activity and region.

Formula

$$I_j = 100 - \left[\frac{\sum_k (x_{kj} * val_k)}{\sum_k x_{kj}} * 100 \right]$$

where:

x = number of units related to each item (i.e. number of register creations);

k = selected sub-population (item);

j = cell (class of employee, sector, etc);

val = subjective measure assigned by statistician (BR staff) according to the importance of each item.

Mixed Quality Criterion: Coverage, Up-to-Dateness, Transparency, Dissemination

- 50) Updates and changes referring to time t when disseminating t+1 data: number (%) of false active units (non-active) in t; number (%) of units with different NACE in t; number of units with revised number of employees in t and related employees (%).

Annex F2: Colombian Experience in Implementation of Quality Processes in the Statistical Directory of Enterprises

F2.1 Introduction

In Colombia, the process of updating the statistical business register (which is called the Statistical Directory) is supported by an information system called Directory Information System (SID)⁹¹. It has been in development since 2010. The objective of the development is to automate processes and generate rules and controls enabling improvement in the quality of the information.

The SID is composed of eight modules, which are *supplier management, information preparation, information processing, actualization operations, analysis of quality, information exploitation, management and configuration, management and quality indicators*.

Although there are quality control processes throughout the modules, the article is focused on the information preparation and processing modules, because these are where rules are implemented to ensure that internal and external information is of the best possible quality and can be used for updating.

The main advances and results obtained by implementing a set of validation and normalization rules, as well as the description of the process to combine administrative information with the statistical business register (SBR) information base are described. Finally, six indicators for measuring the quality of the SBR information are proposed.

F2.2 Description of the Business Register of Colombia

The Directory Information System (SID) aims to maintain an updated SBR by using information coming from administrative sources, updating operations (telephonic, field, and Web) and statistical operations.

The process of updating the SBR covers three main components, namely, universe, coverage and traceability of statistical units. These components are described as follows.

- *Universe*: the SBR must contain all the enterprises (and corresponding legal and local units) conducting economic activities in the country.
- *Coverage*: the SBR must guarantee a national coverage of all the statistical units addressing economic activities in every sector, this latter determined by means of the Uniform International Industrial Classification, Revision 4 adapted for Colombia (CIU Rev. 4 A. C.)⁹².
- *Statistical units*: the conceptual data model refers to three basic units: the *enterprise*, the *homogenous production unit* and the *establishment*. Nonetheless, the model is being adapted to a more general framework in which the legal units, enterprises, local units and enterprise groups, are taken into account.
 - *Enterprise*: economic entity or combination of the economic entities capable, by its own right, of possessing actives, incurring obligations and conducting economic and

⁹¹ Its acronym in Spanish

⁹² This is a classification according to economic activity.

productive activities with other entities to achieve the objectives for which it was created.

- *Homogenous production unit*: characterized by a single activity: product inputs, production processes and homogeneous product outputs.
- *Establishment*: An enterprise or an enterprise part located in a topographically delimited place in which, or from which, economic activities are conducted. Its implementation includes a set of variables classified in four categories: identification, location, stratification and management.

The traceability of the statistical units is made possible via the identification, location, stratification and management information, which is updated for each enterprise through the SID:

- *identification*: the Tax Identification Number (NIT)⁹³, social reason, acronym, commercial (trade) name, legal representative and juridical form are recorded, in addition to the identification keys for every statistical unit in the database;
- *location*: department, municipality, address, web page, e-mail, telephone;
- *stratification*: CIU code, busy staff, incomes from operational or sales;
- *management*: constitution date, status, activity initiation and cessation dates.

F2.3 Quality improvement of the administrative information

The quality of the administrative information is measured in terms of coverage, reliability, coherence, opportunity, accessibility and traceability in the SBR frame. The updating and maintenance process is based on the following procedures:

- processing and verification of administrative data;
- feedback from economic surveys;
- processing of economic unit data from censuses that are stored in the DANE (which is a database);
- own processes for verifying and validating the SID information.

The six major modules of the SID that guarantee the quality of the business register are: *information management, information preparation, information processing, updating operations, quality analysis* and *information exploitation*.

Information management

The information management module monitors information from the moment it is requested from suppliers until it is received by the DANE. The information received is checked to see whether it satisfies the minimum variable requirements to be used in the SID. The information is requested in a specific format. Definitions are provided.

⁹³ The Tax Identification Number (NIT) constitutes the identification number of those enrolled in the Tax Unified Register (RUT), allowing individualizing the contributors and users, for every effect on tax, custom and change matters, and especially for accomplishing such obligations.

A process includes a general review of every database provided, and the statistical record is generated containing a diagnosis of the information provided to support for the preparation and loading processes.

Information preparation

The information preparation module consolidates the data provided by the suppliers at the level of the standard economic units (*enterprise, establishment, and homogeneous production unit*). A consolidated file with data from the different administrative sources is obtained. The variables in the supplier consolidated file are semi-automatically aligned to those of the SID.

Information processing

The information processing module automatically normalizes, codifies, and applies rules of validation and consistency of the information consolidated from the suppliers to update the consolidated file. The following basic elements are distinguished in this processing:

- validation rules incorporated to the SID;
- unique identifier for each statistical unit;
- address normalization for every statistical unit;
- detection and elimination of duplicates;
- Creation of a catalogue to normalize words in names; for example, the word “Limitada” (Limited) could be abbreviated as Lta, Ltda, Limit, etc., but the process normalizes it as “Ltda”.

Some statistical units are sent to a base for revision when they do not satisfy the parameters. Additionally, the information received from administrative sources is checked in this module.

The SID determines in automatic manner whether the statistical unit exists in the data base, and identifies the new units and the continuing units.

The reference date and origin of the value of every variable for every statistical unit is registered as the information is updated, so every value is traceable.

Updating Operations

Information updating operations are conducted by means of call centres, Internet and electronic forms for economic units. The call centre technology platform has been modernized to improve the quality. The web form allows the enterprises to supply their information by directly electronically.

Quality analysis

There is a set of tools that enables analysis of any kind of information from administrative reports or contained in the SID.

Information Exploitation

The real time consultation module is created that enables generation of frames for the SID users.

F2.4 Results of the Information Preparation and Processing Processes

The information preparation module yields two main products: a consolidated base and a statistical report structured in 4 parts:

- a variable list, and their descriptions, coming from the file provided by the supplier (Dictionary);
- an attachment to the Dictionary in which the variable classifications are presented;
- a basic diagnosis indicating duplicate, empty and inconsistent cells (see. Figure F2.1).
- frequencies for every categorical variable, including department, municipality, and legal organization, among others).

Figure F1.1. Data base diagnosis for updating
(Source: Information System for the Statistical Business Register.)

									
DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA									
DIRECCIÓN DE GEOESTADÍSTICA - FICHA TÉCNICA									
CARACTERÍSTICAS DE LA INFORMACIÓN									
ENTIDAD QUE SUMINISTRA LA INFORMACIÓN		PLANILLA INTEGRADA DE LIQUIDACIÓN DE APORTES							
FECHA EN LA QUE SE RECIBE LA INFORMACIÓN		31/12/2012							
NOMBRE DEL ARCHIVO ORIGINAL		20121231163952_Aportes_2_2012_DANE_2.txt							
UBICACIÓN ARCHIVO ORIGINAL		c:\datos\DEE\ARCHIVOS_PLANOS_REGISTROS_ADMINISTRATIVOS\FILA\							
CANTIDAD DE VARIABLES		40							
CANTIDAD DE REGISTROS		13444948							
DESCRIPCIÓN DE LA INFORMACIÓN									
IDEN	NOMBRE_ARCHIVO	NOMBRE_DIRECTORIO	PRESENCIA DE DATOS		VALIDACIÓN DE DATOS		INFORMACIÓN DUPLICADA POR VARIABLE		
			Registros validos	Registros no validos	Registros inconsistentes	Registros consistentes	Registros duplicados	Registros unicos	Registros diferentes
1	tipo_identificacion	TIPO_DOCUMENTO	0	13444948	0	0	7	1	8
2	numero_identificacion	NIT	0	13444948	0	0	2221158	151468	2272614
3	digito_verificacion	DIGITO_VERIFICACION	4329	13440519	0	0	12	0	12
4	razon_social	RAZON_SOCIAL	0	13444948	0	0	2202938	149567	2352505
5	codigo_sucursal		11075723	2369125	0	0	2555	113	2668
6	nombre_sucursal		639808	12895040	0	0	19549	698	16247
7	clase_aportante		639185	13805663	0	0	6	0	6
8	sector_aportante		639656	12895192	0	0	7	0	7
9	tipo_persona	OJUR_ID_OJURIDICA	639656	12895192	0	0	4	0	4
10	direccion_correspondencia	DIRECCION	647518	12797230	0	0	1892205	111268	2004473
11	codigo_ciudad	MUNI_ID_MPIO	639667	12895181	0	0	599	1	596

Codification Rules

The data base information received in accordance with the reference tables is codified in this process. Codification applies to *municipality*, *department*, *economic activity*, *type of document* and *legal organization*. For example, there are cases in which the information does not come codified but named with the department or municipality where the economic unit locates. The process transforms these data into the official codes according to the Colombian Political-Administrative Division (DIVIPOLA).

Normalization rules

The system allows normalization of economic unit address (location), name, telephone number, by applying defined and updated rules, according to the word normalization catalogues.

Symbols like # or – in the addresses, are eliminated within the normalization process.

Validation rules

The completeness and consistency of information is verified as follows.

- The length of the telephonic numbers must be 7 or 10 digits, without the inclusion of the city indicative. The datum type must be numeric.
- The address must have more than 4 characters.
- The commercial (trade) name must have more than 4 characters.
- If the statistical unit is classified as active, and comes from a DANE survey, it must have information about incomes and employed persons.
- Every statistical unit must have an identification unique number, which must be within a specific length range depending on the type of the associated document.
- The e-mails must have the symbol @ and belong to a valid domain.

F2.5 Proposed Quality Indicators

In this moment, Colombia is conducting the process of defining a set of indicators whose purpose is to evaluate the quality of SBR. The indicators proposed are as follows.

Indicator 1

Name: Updating level.

Objective: to know the updating rate for every economic sector in the frame.

Type of Indicator: process quality.

Variables used in the calculations are:

- A_j : Total updated records for sector j.
- B_j : Total records expected for updating in sector j.

The formula used for the calculation is: $I_{1j} = \frac{A_j}{B_j} * 100$

Calculation Frequency: annual.

Tolerance ranges:

- Critical ≤ 70 ;
- $70 >$ Fair ≤ 90 ;
- Satisfactory > 90 .

Indicator 2

Name: birth and death tracking

Objective: to evaluate the sector dynamics based on economic unit demographics.

Type of Indicator: process quality.

Variables used in the calculations are:

- A_i : births in year i.

- B_i : deaths in year i .

The formulae used for the calculation are: $I_{2A} = \frac{A_i}{A_{i-1}}$; $I_{2B} = \frac{B_i}{B_{i-1}}$

Calculation Frequency: annual

Indicator 3

Name: Coverage.

Objective: to establish the SBR coverage compared to the information from the Tax and Custom National Directory (DIAN) database.

Type of Indicator: process quality.

Variables used in the calculations are:

- t_i : Total records in the database in year i .
- T_i : Total unique records identified in the Tax database of the National Register in year i .

The formula used for the calculation is: $I_3 = \frac{t_i}{T_i} * 100$

Calculation Frequency: annual.

Tolerance ranges:

- Critical ≤ 70 ;
- $70 >$ Fair ≤ 90 ;
- Satisfactory > 90 .

Indicator 4

Name: Employment precision.

Objective: to determine whether the information by sector kept in the SBR is approximated to the official employment statistics generated by any National Statistics Institute.

Type of Indicator: quality of the process.

Variables used in the calculations are:

- e_i total employees according information contained in the SBR in year i ;
- PEA_i economically active population in the year i , according to official data;
- L_i : lower limit established for estimating the unemployment by Statistic National Institute;
- L_s : upper limit established for estimating the unemployment by Statistic National Institute.

The formula used for the calculation are:

$$I_4 = \begin{cases} 1 & \text{if } L_i \leq 1 - \frac{e_i}{PEA_i} \leq L_s \\ 0 & \text{otherwise} \end{cases}$$

Calculation Frequency: annual.

Tolerance ranges:

- 1 Satisfactory;
- 0 critical.

Indicator 5

Name: Income precision

Objective: To determine whether the information by sector kept in SBR is approximated to the official income statistics generated by any Institute of Statistics.

Type of Indicator: Quality of the process.

Variables used in the calculations are:

- c_i : Total income according with information keep in the SBR in year i .
- C_i : Total income according to official data of national accounts in year i .

The formula used for the calculation is:

$$I_5 = \begin{cases} 1 & \text{if } 0.9 \leq \frac{c_i}{C_i} \leq 1.1 \\ 0 & \text{otherwise} \end{cases}$$

Calculation Frequency: annual.

Tolerance ranges:

- 1 Satisfactory;
- 0 critical.

Indicator 6

Name: Opportunity

Objective: to establish the level in which registers are available for being used by the SBR users.

Type of indicator: output quality.

Variables used in the calculations are:

- u_j : Number of users qualifying the access as timely to SBR (schedule);
- U : Total number of interviewees.

The formula used for the calculation is: $I_6 = \frac{u_j}{U} * 100$

Calculation Frequency: Annually

Tolerance ranges:

- Critical ≤ 70 ;
- $70 >$ Fair ≤ 90 ;
- Satisfactory > 90 .

F2.8 References

Bérard, Hélène, Pursey, Stuart & Rancourt, Eric. *Re-thinking Statistics Canada's Business Register*. Statistics Canada. 11th, Floor, R.H. Coats Building, Tunney's Pasture, Ottawa, Ontario Canada K1A 0T.

DANE, Metodología Directorio Estadístico de Empresas –DEST–, Bogotá, D. C, Colombia. Versión 5. 2013.

Eurostat. *The French Business Register: From a quality approach to a statistical register*. Projects on the Improvement of Business Register. Session 5.

Eurostat. *Business Register Recommendations Manual: The use of Administrative Sources*. Chapter 20. December 2002.

Ritzen, Jean. *Statistical Business Register: Content, place and role in Economic Statistics*.
Statistics Netherlands.

Annex F3: Statistics Netherlands' Administrative Data Source Evaluation Checklist

Quality Framework for Administrative Data Sources

NSI's are increasingly making use of secondary data sources such as administrative registers for the production of statistics. However, the information in these sources is often collected and maintained by other organizations, usually for non-statistical purposes. Since the production of high quality statistics depends on the quality of the input data, it is useful to determine the quality of secondary data sources, in a systematic, objective, and standardized way. For this purpose a quality framework has been developed by Statistics Netherlands that focuses specifically on the quality of administrative and other secondary data sources⁹⁴.

The framework consists of three high level views of the quality of a data source. These views are called: *Source*, *Metadata*, and *Data*. Each view stresses different quality aspects referred to as *dimensions* as detailed in Figures F3.1, F3.2 and F3.3.

Figure F3.1: Quality aspects related to the data source as a whole, the data source custodian, and the delivery of the data source to the NSI

Source		
Dimension	Quality indicators	Methods
1. Supplier	1.1 Contact	Name of the data source. Data source contact information. NSI contact person.
	1.2 Purpose	Reason for use of the data source by NSI.
Relevance	2.1 Usefulness	Importance of data source to NSI
	2.2 Envisaged use	Potential statistical use of data source
	2.3 Information demand	Does the data source satisfy information demand?
	2.4 Response burden	Effect of data source use on response burden
Privacy and security	3.1 Legal provision	Basis for existence of data source.
	3.2 Confidentiality	Does the Personal Data Protection Act apply? Has use of data source been reported by NSI?
	3.3 Security	Manner in which the data source is sent to NSI Are security measures (hard/software) required?

⁹⁴Administrative Data Source Evaluation Checklist <http://www.cbs.nl/nr/rdonlyres/Odbc2574-cdae-4a6d-a68a-88458cf05fb2/0/200942x10pub.pdf>

Source		
Dimension	Quality indicators	Methods
Delivery	4.1 Costs	Costs of using the data source.
	4.2 Arrangements	Are the terms of delivery documented? Frequency of delivery.
	4.3 Punctuality	How punctually can the data source be delivered? Rate at which exceptions are reported. Rate at which data is stored by data custodian.
	4.4 Format	Formats in which the data can be delivered.
	4.5 Selection	What data can be delivered? Does this comply with the requirements of NSI?
Procedures	5.1 Data collection	Familiarity with the way the data are collected.
	5.2 Planned changes	Familiarity with planned changes to the data source. Ways changes are communicated to NSI.
	5.3 Feedback	Can contact data custodian in case of trouble? In which cases and why?
	5.4 Fall-back scenario	Dependency risk incurred by NSI. Emergency measures when data are not delivered according to arrangements made.

Figure F3.2: Quality aspects related to metadata related aspects of the data source

Metadata		
Dimension	Quality indicators	Methods
Clarity	1.1 Population unit definition	Clarity score of the definition.
	1.2 Classification variable	Clarity score of the definition.
	1.3 Count variable	Clarity score of the definition.
	1.4 Time dimensions	Clarity score of the definition.
	1.5 Definition changes	Familiarity with changes that occur.
Comparability	2.1 Population unit definition comparison	Comparability with NSI definition.

Metadata		
Dimension	Quality indicators	Methods
	2.2 Classification variable definition comparison	Comparability with NSI definition.
	2.3 Count variable definition comparison	Comparability with NSI definition.
	2.4 Time differences	Comparability with NSI reporting periods.
Unique keys	3.1 Identification keys	Presence of unique keys. Comparability with unique keys used by NSI.
	3.2 Unique combinations of variables	Presence of useful combinations of variables.
Data treatment (by data custodian)	4.1 Checks	Population unit checks performed. Variable checks performed. Combinations of variables checked. Extreme value checks.
	4.2 Modifications	Familiarity with data modifications. Are modified values marked and how? Familiarity with default values used.
	4.3 Punctuality	How punctually can the data be delivered? Rate at which exceptions are reported. Rate at which data are stored by data custodian.
	4.4 Format	Formats in which the data can be delivered.
	4.5 Selection	What data can be delivered? Does this comply with the requirements of NSI?

Figure F3.3: Quality aspects related to the accuracy of the data

Data		
Dimension	Quality indicators	Methods
Technical checks	1.1 Readability	Can all the data in the source be accessed?

Data		
Dimension	Quality indicators	Methods
	1.2 Metadata compliance	Does the data comply to the metadata definition? If not, report the anomalies.
Over coverage	2.1 Non-population units	Percentage of units not belonging to population.
Under coverage	3.1 Missing units	Percentage of units missing from the target population.
	3.2 Selectivity	R-index ¹⁾ for unit composition
	3.3 Effect on average	Maximum bias of average for core variable. Maximum RMSE ²⁾ of average for core variable.
Linkability	4.1 Linkable units	Percentage of units linked unambiguously.
	4.2 Mismatches	Percentage of units incorrectly linked.
	4.3 Selectivity	R-index for composition of units linked.
	4.4 Effect on average	Maximum bias of average for core variable. Maximum RMSE ²⁾ of average for core variable.
Unit non response	5.1 Units without data	Percentage of units with all data missing.
	5.2 Selectivity	R-index for unit composition.
	5.3 Effect on average	Maximum bias of average for core variable. Maximum RMSE of average for core variable.
Item non response	6.1 Missing values	Percentage of cells with missing values.
	6.2 Selectivity	R-index for variable composition.

Data		
Dimension	Quality indicators	Methods
	6.3 Effect on average	Maximum bias of average for variable. Maximum RMSE of average for variable
Measurement	7.1 External check	Has an audit or parallel test been performed? Has the input procedure been tested?
	7.2 Incompatible records	Fraction of fields with violated edit rules.
	7.3 Measurement error	Size of the bias (relative measurement error).
Processing	8.1 Adjustments	Fraction of fields adjusted (edited).
	8.2 Imputation	Fraction of fields imputed.
	8.3 Outliers	Fraction of fields corrected for outliers.
Precision	9.1 Standard error	Mean square error for core variable.
Sensitivity	10.1 Missing values	Total percentage of empty cells.
	10.2 Selectivity	R-index for composition of totals.
	10.3 Effect on totals	Maximum bias of totals. Maximum RMSE of totals.
<p>1) R-index: Representative Index, an indicator that estimates the selectivity of the data missing by using information available in other sources (Schouten and Cobben 2007, Cobben and Schouten 2008).</p> <p>2) RMSE: root mean square error; statistical measure for the quality of an estimator. The RMSE is equal to the square root of the sum of the bias and variance of the estimator.</p>		

The framework is useful in determining whether a specific administrative or other secondary data source should be used in the compilation of statistics. The framework is applied by successively evaluating the quality indicators at the Source, Metadata, and Data levels. If a quality indicator at the Source level reveals a problem, this should be addressed first, before starting to evaluate the indicators at the Metadata level. Likewise any problems with the indicators on Metadata level should be addressed before evaluating the indicators at the Data level. Finally, if the evaluation of the indicators on Data level is successful, then the data source can be used for the production of statistics.

Quality Checklist

This framework provides the basis for the *Checklist for the Quality Evaluation of Administrative Data Sources*” developed by Statistics Netherlands. This checklist should be completed by the prospective internal user of an administrative or other secondary data source and/or an expert from the data source. For the *source part* of the checklist it is advisable to contact the official NSI contact person for the particular source (if there is one).

Annex G: Additional Concepts and Procedures

Annex G1: Register Based Census

Annex G2: Example of Time Lag in SBR Maintenance Based on an Annual Administrative Source

Annex G3: Calculating a Check Digit for an Identification Number

Annex G1: Register Based Census

G1.1 General Description

An economic census provides information about the structure and function of a production system from the national (macro-area) to the local (micro-area) level. In general it guarantees periodic and comprehensive statistics about businesses, establishments, activities carried out and employment every five or ten years. Economic census statistics are essential “for sound economic policy and successful business planning”⁹⁵, they provide an essential framework for all economic indicators (production indexes, input/output measures, labour, etc.) and they are fundamental inputs for benchmarking GDP estimates.

Basic statistical data on businesses are important and necessary for the public and private sector in policy formulation and development planning of the economy and industry at both national and local level. National and local governments use these data to monitor economic activity and the changes in national and regional economies. Census data are very useful for the individual businesses in calculating market share, locating business markets, identifying business site locations, and evaluating new business opportunities.

In many countries the censuses are regulated by law and provide for mandatory responses. In recent decades, the development, in a wide range of the countries, of the National SBR guarantees a significant improvement in methods and tools for census data collection.

At one end of the spectrum there is the traditional census, collecting data by use of enumerators and questionnaires, using no register information at all. At the other end there is the totally register-based census. Some countries use mixed mode data collection with a combination of data from registers and questionnaires (either total enumeration or a sample survey). Even countries conducting mainly traditional censuses tend to use register information to some extent, for instance as an address list.

The interaction between the SBR and the Census can be defined as *register-assisted census data collection* that combines some elements of a direct door-to-door survey and some of a classic survey by list. This technique of data collection is characterized by the following three elements:

- Enumerators are supplied with lists of the enumeration units located in their *census districts*, drawn from the SBR. Their task consists in verifying the actual status of the listed units, deleting the records of the doubled and the ceased ones, and adding new records for the possible non-listed units (for example born in the lag between reference period of the list and the date of the survey, or unregistered for any other reason).
- Some days before the survey, all the listed units receive by mail a personalized questionnaire partly completed with information drawn from the SBR. In this way, the respondents just have to verify the correctness of the pre-printed fields (rectifying them, if necessary) and complete the questionnaire with the missing information. The questionnaires are collected by enumerators,
- Enumerators are also provided with blank (non-personalised) questionnaires to be used only for non-listed units or in substitution of personalised questionnaires that are lost or damaged.

95 Alan Greenspan, chairman of the Federal Reserve Board, USA

Using this method, based on the synergy between the Census and the BR, the main benefits are:

- *reduction in burden for respondents*, by limiting the number of questions and radically simplifying the questionnaire. Besides, an overall reduction of the amount of collected data allows simplification and shortening of data processing and has a positive impact on the quality of the data itself.
- *a new approach to quality and coverage control*, since it is possible to carry out a micro-level coverage analysis, by comparing the raw data collection file with an image of the SBR as of the same date as the Census. This micro approach, instead of the classical macro approach based on a post-enumeration sample survey, makes possible the precise identification of every single unit under/over covered in both data sources. The results are that the theoretical under coverage – main problem of a direct door-to-door survey – and over coverage - main problem of a SBR – are significantly reduced.
- *creation of a basis for improvement of the SBR* in terms of new characteristics obtained, for example, secondary activities.

An evolution of the previous *register-assisted approach* is the development of a *register-based Economic Census*. With this approach the census data are reproduced entirely by integrating data from SBR and administrative sources, without any direct data collection from businesses. The potential for turning a traditional business census to a register-based one depends mainly upon the degree of enhancement achieved in the statistical use of administrative sources and on quality improvements by each administrative body in the business area.

The advantages of the register-based over a register-assisted approach are:

- a significant decrease in costs for the statistical authorities;
- the non-existence of respondents means the absence of respondent burden;
- statistics can be made available every year, providing opportunity to detect shorter term changes in the economic structures of the national and local economies.

On the other hand, besides the classical problems that arise in the use of administrative sources (the need to deal with administrative definitions and their operational rules, the timeliness in the production of data, the treatment and the exploitation of new administrative data, the enlargement of the dataset of information, the huge amounts of data linked), the key disadvantage of this approach is the absence of direct statistical information to improve SBR quality, especially in terms of under/over coverage.

There are some significant criteria for implementation of a register-based census.

- *High quality SBR in terms of coverage*. The population, in terms of enterprises and local units, is decided and cannot be changed or corrected.
- *Existence of a unified identification coding systems across different administrative sources*. In the absence of such unified systems it is extremely difficult, if not impossible, to link different registers. Integration is the key for a register based census.
- *Methodology for translating administrative characteristics into the statistical ones*. The development is required of appropriate statistical methodologies (probabilistic and/or deterministic) to translate or to estimate statistical characteristics starting from one (or more linked) administrative data sources.

The realization of a register-based census depends having the appropriate national statistical objectives and sufficient national administrative sources to support the objectives. There is no

single recipe. The actual approach depends upon the country situation including the availability and completeness of administrative sources.

G1.2 The Italian ‘Virtual’ Economic Census (VEC)

For the first time the decennial Business Census named *CIS 2011*, aiming at the enumeration of businesses, related persons employed, and other types of employment, classified by activity code, size, juridical status and other structural information about the system of enterprises was done without any direct data collection from business. Instead it was based exclusively on statistical data obtained from the integration of administrative sources.

The Virtual Economic Census (VEC) system is built around a set of basic registers containing comprehensive data on business units and individuals. The core of this system is the national SBR (*ASIA BR*), which is produced yearly by integration of administrative and statistical sources. The *ASIA BR* is considered the reference population and the official source for the Italian statistical information on the structure and the demography of the business population.

The identification and acquisition of a new set of administrative sources changed the way the process is carried out, and added new contents to the database of business units and individuals in which information is available not only at *unit (enterprise, local unit) level* but also at an *individual level*. Each person involved in the business production process is identified. In fact each unit inside an administrative source is linked to the *ASIA BR* statistical unit by means of identification codes. Persons can be linked to the business in which they assume any ownership share, to the employer for whom they are working, etc.

This system makes use of many administrative sources. They can be grouped into different types:

- *fiscal data* - VAT, income, participation in partnerships, remuneration taken from 770-Form Tax Register, statistics-based Tax Assessment fiscal survey - i.e. ‘Studi di settore’;
- *social security data* - monthly employer declarations on each employee, on outworkers - i.e. persons that are linked to a production unit and that are not employer/own account workers or employees – and on self-employed workers and family workers in agriculture, handicraft and trade);
- *Chamber of Commerce data* - list of partners or shareholders of each legal unit;
- *employment insurance data* - from the government agency responsible for insurance against work-related injuries (workplace, insurance payments).

The integration of these sources has been made possible by the existence of a tax identification code that is available – with a very high level of coverage - both for legal units and individuals, in all sources and in *ASIA-BR*. Thus, in Italy, the tax code represents an indirect unified identification code system.

Using information from these sources, a new employment data structure for the BR was set up - the LEED structure (Linked Employer-Employment Database). While in the old BR the number of persons employed represented one of the characteristics associated to a statistical unit (attribute), in the new data structure each single person is linked to each statistical unit for which it has any form of employment (according to the conceptual framework of the employment classification). In this new system, businesses and workers employed are identified by links derived by the integration and matching processes for the various administrative sources. The data structure is based on links, i.e. direct relationships between employment identification codes and enterprises, together with the basic enterprise attributes,

employment composition at enterprise level (e.g., gender and age composition, workplace) and employment attributes that differ according to the type of employment.

This new integrated approach is able to provide an in-depth analysis of the employment of the enterprises and of local units. This analysis was, and still is, the main objective of the Italian VEC. The main outputs of the Italian VEC are:

- data on employment, with a particular focus on the demographic characteristics of the workers, such as gender, age, birth location (country);
- more detailed information and job characteristics (professional status, type of contract, full/part time, etc.);
- identification of the different types of workers used in a business (employees, self-employment, family worker, outworkers, temporary workers) to provide a global picture of the labour input, both for each enterprise, and at sectorial and territorial levels;
- to face and solve measurement issues (hours worked, number of jobs, full time equivalent jobs, number of persons employed)

From conceptual point of view the change in moving to the VEC can be summarized in the two schemas in Figures G1.1 and G1.2.

Figure G1.1: Old conceptual approach: SBR-Italy (ASIA)

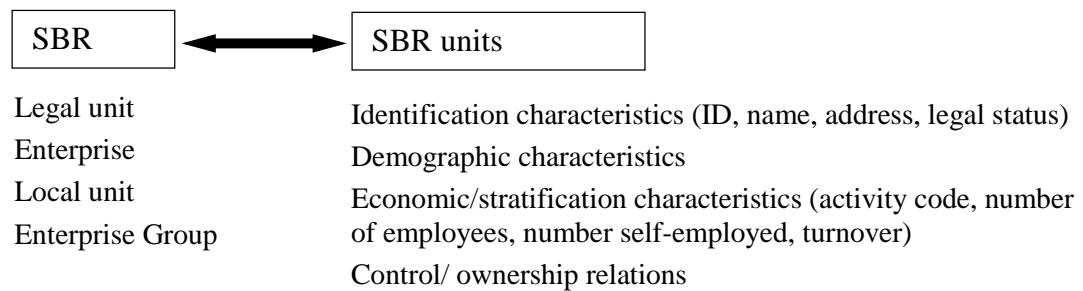
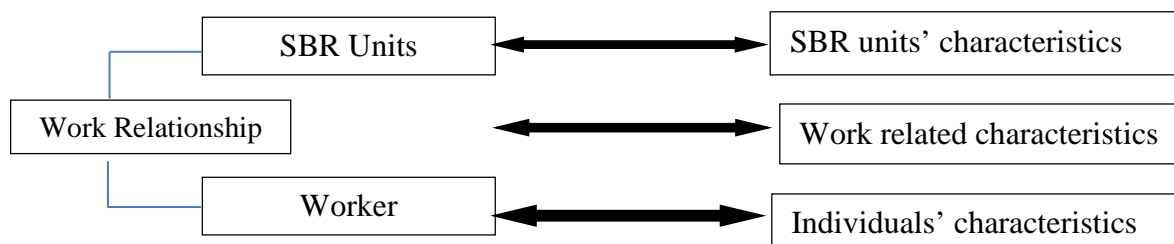


Figure G1.2: New conceptual approach: Integrated system for Virtual Census



Furthermore the VEC provided an opportunity to face and solve problems with employment definitions as regards better coherence with the international standards, and with translation of such definitions into operational rules. The global revision of the employment classification system and of employment measurement methods have been fundamental results of the VEC, providing improved coherence for the whole national statistical system.

Annex G2: Illustration of Time Lag in SBR Maintenance Based on Annual Administrative Source

The following example illustrates some issues that may arise as a result of an SBR maintenance schedule that depends upon an administrative source that is updated annually.

Suppose the SBR contains only a population of simple enterprises that are in one to one relationship with legal units and that these units are updated using a single administrative source with the following properties:

- each administrative unit is linked to only one enterprise;
- for the enterprises linked to the administrative source, an enterprise death can be detected on a continuous basis from the death of the corresponding administrative unit;
- the birth of an administrative unit is potentially a new enterprise, but not all new administrative units are relevant to the SBR (i.e., there is potential for over-coverage);
- data from the administrative source lag the actual births and deaths by one year.

Given the time lag of the administrative source, Figure G2.1 below gives the availability of data from the administrative source.

Figure G2.1 Administrative data availability

<i>Calendar date</i>	<i>Reference period</i>
2011_dec	2010
2012_dec	2011
2013_dec	2012
2014_dec	2013

For example at the end of 2013 (Calendar date = 2013_dec, i.e. December 2013) the population of active administrative units during the reference year 2012 becomes known.

Let $A(2012)$ denote the population of active administrative units during the reference year 2012.

Potential births of enterprises can be detected by comparing the populations of active administrative units for two consecutive years. Taking again as an example the situation in December 2013, the populations of administrative units $A(2012)$ and $A(2011)$ are available. Taking the difference of these two populations, the administrative units created during the reference year 2012 can be identified.

Let $B_admin(2012)$ denote the population of administrative units created during the reference year 2012.

In the situation considered here, a newly created administrative unit may not necessarily correspond to an SBR enterprise. The information to decide there should be an enterprise birth must be obtained through a control survey. For the population $B_admin(2012)$, this control survey takes place during 2014 and ends, say, in June 2014. At the end of the control survey, the administrative units which correspond to new enterprises have thus been identified.

Let $B(2012)$ denote the population of enterprises created during the reference year 2012.

Figure G2.2 below gives the availability of data from the administrative source and control survey identifying the new SBR enterprises.

Figure G2.2 Availability of data about new enterprises

Calendar date	Active units	Births detected	New enterprises
2011_dec	A(2010)	B_admin (2010)	
2012_dec	A(2011)	B_admin (2011)	B(2010)
2013_dec	A(2012)	B_admin (2012)	B(2011)
2014_dec	A(2013)		B(2012)

Based on this SBR updating process, choices can be made about when to construct new sampling frames. For example, it may be decided to create a new version of the common enterprise frame each year in December.

Taking again December 2013 as an example, the coverage as of that date would correspond to the enterprises created up to the end of 2011, less all the enterprises which have disappeared up to the end of 2013. The enterprises created during 2012 are known to be a subset of the administrative units created during 2012 but have not been yet identified through a control survey. The enterprises created during 2013 are completely missing. For the enterprises that are covered, the content (i.e. values of the characteristics of the units) obtained from the administrative source corresponds to the reference year 2012.

Annex G3: Calculating a Check Digit for an Identification Number

Introduction⁹⁶

A *check digit* is a decimal (or alphanumeric) digit added to an identification number. It is a form of redundancy check that assists in detection of errors in identification numbers (for example, bank account numbers) that have been manually entered. It is analogous to a binary parity bit used to check for errors in computer-generated data. It is computed by an algorithm from the other digits (or letters) in the identification number.

The algorithms used to generate check digits are designed to detect typical human transcription errors. In order of complexity, these include:

- single digit errors, such as 1 becomes 2;
- transposition errors, such as 12 becomes 21;
- twin errors, such as 11 becomes 22;
- jump transpositions errors, such as 132 becomes 231;
- jump twin errors, such as 131 becomes 232; and
- phonetic errors, such as 60 becomes 16.

In choosing an algorithm, high probability of detecting errors is traded off against implementation difficulty. Simple check digit algorithms are easily understood and implemented by humans but do not detect as many errors as complex ones, which, however, require sophisticated programs to implement.

Use of check digits in the Swiss Business Register

In the Swiss Business Register, an algorithm based on Modulo 11⁹⁷ is used to calculate the check digits for all identification numbers. The characteristics of the algorithm are different from those based on other modulus, such as the more common Modulo 10.

Local unit identification number (BURNR)

Local unit identification number (BURNR) in the Swiss Business Register is an 8 digit number. The last digit is the check digit number.

- Example of BURNR: 62088168 – the final 8 is the check digit.

The procedure for calculating the check digit as follows

Each of the first 7 digits of the identification number is individually multiplied by a multiplier selected from a fixed sequence (5, 4, 3, 2, 7, 6, 5) according to the position of the digit within the number.

- The digit in position 1 is multiplied by 5
- The digit in position 2 is multiplied by 4
- The digit in position 3 is multiplied by 3
- The digit in position 4 is multiplied by 2

⁹⁶ Extracted from Wikipedia: http://en.wikipedia.org/wiki/Check_digit

⁹⁷ The Barcode Solution <http://www.activebarcode.com/codes/checkdigit/modulo11.html>

- The digit in position 5 is multiplied by 7
- The digit in position 6 is multiplied by 6
- The digit in position 1 is multiplied by 5

All resulting products are added and the result is divided by 11.

- If the residual (Modulo 11) is zero the check digit is zero.
- If the residual is 1, the original 7 digit number is not used as an identification number – it is considered invalid.
- Otherwise the check digit is the number obtained by subtracting the residual from 11.

Example of calculation:

BURNR	62088168
Effective digits	6208816
Sequence of multipliers	5432765
Products	$(6*5=30) + (2*4=8) + (3*0=0) + (8*2=16) + (8*7=56) + (1*6=6) + (6*5=30) = 146$
Result	$146 / 11 = 13$ residual 3
Check digit	$11 - 3 = 8$

Enterprise Identification Number (ENTID)

A similar procedure is used for the Enterprise Identification Number (ENTID) which has 9 digits including the check digit. The sequence of multipliers (5, 4, 3, 2, 7, 6, 5, 4). The first 7 numbers are the same as for the BURNR

Example of calculation

ENTID	109322551
Effective digits	10932255
Sequence of multipliers	54327654
Products	$(1*5=5) + (0*4=0) + (9*3=27) + (3*2=6) + (2*7=14) + (2*6=12) + (5*5=25) + (5*4=20) = 109$
Result	$109 / 11 = 9$ residual 10
Check digit	$11 - 10 = 1$

Unique Enterprise identification number (UID)⁹⁸

Exactly the same procedure is used for the Unique Enterprise Identification Number (UID).

Example of calculation:

UID	CHE109322551
Effective digits	10932255
Sequence of multipliers	54327654
Products	$(1*5=5) + (0*4=0) + (9*3=27) + (3*2=6) + (2*7=14) + (2*6=12) + (5*5=25) + (5*4=20) = 109$
Result	$109 / 11 = 9$ residual 10
Check digit	$11 - 10 = 1$

⁹⁸ UID Register : <https://www.uid.admin.ch/Search.aspx?lang=en>

Annex H: Topics for Further Work and Research

H1. Introduction

The Guidelines focus on providing methodological and practical guidance to countries for establishing and maintaining SBRs. They do not, however, resolve a number of emerging issues that will pose major challenges in the future.

The continuous development of businesses, the economic globalisation and the growing complexity of global production arrangements raise a number of questions concerning definitions and identification of relevant statistical units of the SBR and international comparability of statistics derived from it. These developments also make data collection more complex and difficult. At the same time, the SBR plays a key role to play as backbone in the modernisation of statistical production and integration of economic statistics. To strengthen this role there is a need to explore how to integrate the SBR into the statistical production process as a common tool and database for all domains of business statistics, and to ensure the SBR can be used to combine information from different sources. In addition to this there continues to be a need to explore the use of administrative registers and other new or emerging electronic data sources and for further development of methods and concepts for the establishment and maintenance of the SBR.

Developing new methods and good practices is resource demanding, for which reason experiences should be shared among countries and common development projects should be encouraged. International cooperation can help reduce gaps between developed and less developed statistical systems, while capacity building and training activities will be useful to support countries with less developed statistical systems. In addition, sharing methods and practices may eventually help improve the international comparability of SBR information.

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

H2. Statistical Units and Profiling

The statistical units are the basic entities in SBRs. The most important units are the enterprise, the establishment and the local unit. However, due to the globalisation of economic production the enterprise group has also become an important statistical unit that needs to be integrated into SBRs. Legal entities in the economic world do not always meet the criteria as units suitable for statistical purposes. They need to be transformed into statistical units in a consistent way. This is especially true for large, complex and global businesses. Work in the following areas should be supported:

- Delineation of the *enterprise* statistical unit.
- Delineation and recording of *multinational enterprise groups* in the national SBR.
- Analysis of the experience of countries in profiling large and complex enterprises and enterprise groups.
- Sharing best practices in the coding of statistical units.

H3. SBR as the Backbone of Business Statistics

It is important to develop the role of the SBR as the backbone in the production of all business statistics. The SBR fulfils this role by providing survey frames and by integrating information from different data sources (surveys/censuses, statistical registers and administrative registers) and facilitating collection of economic data and production of coherent statistics as part of the thrust to modernise statistical production and services. It can also facilitate the development of a standardised production process for business statistics. Countries will need to move forward and exchange experiences made in developing integrated approaches and operational designs. Activities that will facilitate this include:

- Integration of SBR into data warehouse systems.
- Sharing of experiences in implementation of the GSPBM and the GSIM for the SBR database management.
- Development of efficient user oriented maintenance strategies.
- Balancing growing demands of user groups.

H4. Use of Administrative Data Sources

The use of administrative sources continues to be a major challenge, but has the potential for gains in many countries. The challenges include issues with coverage and definitions, which may not be in line with the statistical requirements, and timeliness. In some countries lack of access to administrative sources is also an obstacle. Potential gains include access to sources with good coverage and detailed information that can be used directly in the SBR or in validation of SBR information, reduction of response burden and increased efficiency in the statistical office. Establishing good cooperation with the owners of administrative data sources is essential to ensure sustainable solutions and to ensure that the NSI is kept informed of, and exercise influence on, future changes to administrative registers. Examples of areas where work is encouraged are:

- Sharing of best practices in cooperation with administrative registers.
- Development of register-based censuses.

H5. Use of New Data Sources

New data sources and new ways of collecting data are emerging with the potentials of improving the quality and coverage of the SBR, and in some cases also the timeliness of the register. Activities that should be pursued include:

- Sharing experiences in using new data sources, including big data.
- Sharing experiences in new electronic data collection methods.

H6. Economic Globalisation

Economic globalisation is one of the main challenges for SBRs. It influences the choice of activities to be measured and data collection becomes more complex and difficult. Economic globalisation is accompanied by the development of global production arrangements, often with complex ownership structures, where the challenge of the SBR is to collect and provide

coherent and relevant information in a continuously changing world. Work on the following issues should be encouraged:

- What are the requirements of the international statistical standards regarding, for example, the National accounts and Balance of Payments, and how can the SBR meet these requirements?
- Pilot study for the development of a supra-national SBR to deal with multinational enterprise groups (based on experience with the EuroGroup Register).
- Use of the SBR for micro-data linking.

H7. Development of New Statistics

How can information from the SBR be combined with information from other sources (administrative data, surveys or other statistical registers) to meet user needs? Possibilities include combining information from the SBR with social and/or population data, and deriving business demographic information. Questions and issues that should be addressed include:

- What are the potentials of the SBR as a part of a set of interrelated/linked statistical registers?
- Analysis of the effects of using different statistical units in business demography
- Sharing experiences in applying geocoding in the SBR

H8. Methodological Developments

There needs to be further work on various methodological issues for the establishment, maintenance and use of the SBR will be needed to take advantage of developments in techniques and data sources, and to ensure the quality and relevance of the SBR in meeting current and future user needs. Examples of particular topics where work is called for include:

- Evaluation of record matching methods.
- Evaluation of database software packages that can be used for SBR.
- Analysis of software services/methods for SBR maintenance that could be shared between countries.
- Analysis of various forms of frozen frames with respect to their use for business statistics.
- Analysis of methods of creation of identifiers and check digits.
- Use/integration of international administrative identifiers for national SBRs.
- Development of methods for adjustment of under-coverage.
- Development of efficient user oriented maintenance strategies.
- Development of input data validation methods.

H9. International Comparability

The importance of the international comparability of SBRs and derived business statistics is likely to grow in the future. Statistical offices need to consider this when developing methods and practices, while national differences in terms of legislation, institutional set-up

and structural diversities put limits on obtaining international comparability. To address these challenges work in the following areas is encouraged:

- Development of internationally comparable quality indicators.
- Development of an international quality assessment framework.
- Use/integration of international administrative identifiers in the national SBR.