EUROPEAN COMMISSION – EUROSTAT
ESSNET ON CONSISTENCY OF CONCEPTS AND METHODS OF BUSINESS AND TRADE-RELATED STATISTICS –
2011 PROJECT ON TARGET POPULATION, FRAMES, REFERENCE PERIOD, CLASSIFICATIONS AND THEIR APPLICATIONS
MBGA N° 30121.2011.004-2011.289

Deliverable No 2.8

Report on the Evaluation of Inconsistencies with Respect to Target Populations, Sampling Frames and Reference Periods, and their Application in the Member States

Task: 2.3

Author: Boris Lorenc, Martin Ribe (Statistics Sweden)

Date: November 2013

Report on the Evaluation of Inconsistencies with Respect to Target Populations, Sampling Frames and Reference Periods, and their Application in the Member States

Table of contents

Summary

This report addresses threats to consistency that arise from issues related to target populations, sampling frames and reference periods, and practices when dealing with these components of the process of producing economic statistics in the participating countries. The report is based on a set of questionnaires (WP2 Deliverable 2.6) sent to Business Register and 19 Statistical Domains representatives in 27 (at the time of conducting the survey) Member States and four EFTA countries (the participating countries, in what follows). The results, summarised in WP2 Deliverable 2.7, are analysed and evaluated in this report.

For the analysis and evaluation a definition of consistency was further developed, as well as different criteria for evaluating consistency (for a micro- and a macro-level) introduced. Variation between the statistical domains in responses concerning methods and practices of dealing with target populations, sampling frames and reference periods – where not motivated by differences in regulations – indicates potential threats to so-called horizontal consistency. Additionally, variation between responding countries in these responses indicates potential threats to so-called vertical consistency, which might result in producing European statistics whose quality is weakened by consistency issues. However, both kinds of consistencies do have effects on the European level and need to be addressed in that perspective, as inconsistencies between statistical domains in participating countries are also reflected in inconsistencies between statistical domains on the European level.

Further, the evaluation was in part also set in the context of on-going processes towards integration of statistics production into systems that aim at more effective use of resources through better standardisation, better flexibility through more unified processes, reduction of primary data collection through increased use of administrative data, improved process and product quality, among other goals. An improved consistency of produced statistics, that results from such systems 'per design', needs to be seen as one component of the overall quality of statistics, balanced against other quality components as well as against costs for statistics provision (businesses) and statistics production (producers of official statistics). Involvement of users is essential in assuring external validity and fitness for purpose of such projects.

Proposals aimed at improving the horizontal and vertical consistency of statistics in the participating countries and on the European level, based on the analysis and evaluation presented herewith, are to be found in Deliverable 2.10.

1. Goal, structure and content of the evaluation report

This report is a deliverable of the "2011 project on target population, frames, reference period, classifications and their applications" of the "ESSnet on consistency of concepts and applied methods of business and trade-related statistics" (in short "ESSnet on consistency").

As determined by the structure of the ESSnet on consistency, the 2011 project on *target populations*, *sampling frames*, *reference periods*, *classifications* and their applications is work package #2 (WP2) of the ESSnet on consistency focused on the components contained in its name, whereas *statistical units* and *variable definitions* - which also may introduce issues to consistency of statistics - are treated in WP1 and WP3 respectively of the same ESSnet. Thus, the current report as a deliverable of WP2 is limited to those consistency issues that may be related to the components: target populations, sampling frames and reference periods; we refer also to the companion Deliverable 3.8 of this WP, regarding NACE coverage, breakdowns and size-classes.

Current deliverable is based on a report of results of the inventory, Deliverable 2.7 of WP2, which in turn is based on data collected by the questionnaire presented in Deliverable 2.6 of the work package. Further, the evaluation report takes as input the analyses and considerations outlined in the direction report of the work package, Deliverable 2.1.

It was not part of the goals set for this study and this report to investigate, by itself, the degree to which the participating countries (based on responses by their Business Register (BR) and Statistical Domains (SDs) representatives) adhere to regulations governing production of European official statistics. However, some of the respondents might have listed issues which indicate, in a particular country and a SD or BR, deviations of the current practices from the regulations' intentions. Endeavour was made in the report to highlight such instances. But, as concerns the evaluation conducted within this project, it is at this point only assumed that – in cases where a deviation exists – the concerned BR/SDs will take action to align their practices with the existing regulations.

The report consists of three main parts. First, Chapters 2 and 3 discuss a theoretical basis for the analysis and evaluation. Second, Chapter 4 gives an analysis of the questionnaire results by theme, integrating into each theme responses of both the Business Register representatives and the Statistical Domains representatives. Third, Chapter 5 summarises the analyses, integrates the observations and then evaluates the observations regarding the participating countries and on the EU level. That chapter also assesses the findings in terms of the requirements laid out in the Direction Report (Deliverable 2.1).

2. Theoretical considerations

2.1. Concept of consistency

2.1.1. The idea of consistency

Like in the general theory of statistics, it will also here be understood that many sources of uncertainty are involved in production of any statistics, whose statistical properties (first and second moments) may contribute even to discrepancies relevant when considering consistency, as discussed below.

The concepts of consistency and its counterpart inconsistency have not generally been formally defined in the usual survey methodology theoretical framework. However, use of the term occurs in fundamental principles or survey quality frameworks, notably in those of Eurostat (the Code of Practice, 2011), UNSD (Fundamental Principles of Official Statistics, not dated), and the European Central Bank (2008). The Project's Deliverable 2.1 reviews the concept and its relation to the other quality framework concepts, in particular those of coherence and comparability, as well as accuracy.

Basically, consistency involves considering coherence or comparability of at least two statistics. Two questions are of relevance here: how and where to evaluate consistency.

First, how to evaluate consistency: in principle it is to be done from metadata, understanding though that these need to be complete and accurate (as discussed in Deliverable 2.1). For the purpose of this report, the metadata to be considered are those that describe target populations, sampling frames, classifications, and reference periods used to produce the statistic. Thus, here the consistency comprises what involves these concepts.

Second, where to evaluate consistency: An evaluation of consistency could pertain to a comparison of any two or more of the following:

a. a published statistic;
b. a statistic calculable on the basis of available data, in particular microdata;
c. an ideal statistic or "true value", produced from an ideal process, perfectly well-defined and error-free.

In all of these cases, an inconsistency could be seen as any discrepancy between the two or more statistics that is due, not to any real difference in the phenomena under investigation, but to differences in the above mentioned concepts characterising the production of the statistics (i.e. target populations, sampling frames, classifications, reference periods, etc).

2.1.2. Remarks on how to apply the concept

In positing categories (a) and (b), it is assumed that publication of statistics by its producer is such that there is a well-defined publication plan which unequivocally defines which statistics are published (i.e. a list of published statistics, each of which, when published - it is assumed - corresponding metadata are attached). With that, category (a) would be well defined. In addition to that, producers of statistics nowadays increasingly make less aggregated data or microdata available, providing for users the possibility to calculate their own statistics, category (b) here. This category of statistics is much less well defined which is why it is suggested here to treat it separately.

Categories (a) and (b) correspond to non-comparability or non-coherence, depending on the context, in the sense of the Code of Practice. Category (c) corresponds to the well-known notion of "total survey error" (Biemer, 2010). As already noted in Deliverable 2.1, this is related to the quality dimension of accuracy, which is different from that of consistency. However, in all categories the difference also has a random component, due to random sampling and nonsampling errors. Ideally this random component of the difference is not to be seen as contributing to inconsistency.

2.1.3. A definition

Consistency means agreement in a set of concepts (their referents), as reflected in complete metadata, pertaining to two or more produced statistics that leads to the statistics being coherent and comparable. In this report - in the context of WP2 - the relevant concepts are target population, sampling frame, coverage, classifications, reference period, and breakdowns.

Different types of consistency can be defined:

1. horizontal - consistency of produced statistics between two or more statistical domains in a participating country or between two or more statistical domains on the EU level,
2. vertical - consistency of produced statistics within the same statistical domain between participating countries, or their joint consistency with the corresponding statistics produced on the EU level,
3. temporal - consistency of a single produced statistics within the same statistical domain and participating country between different reference periods.

For the rest of the report, only the two former kinds of consistency issues will be considered relevant.

2.1.4. Remarks on the definition

The above definition implies that, to decide on existence of issues with consistency, it does not suffice to inspect the produced statistics themselves (their numerical values), but rather to investigate the complete set of metadata corresponding to these statistics. However, in order to determine importance of a deviation in consistency, it is helpful to investigate the possible numerical impact. The impact may or may not be large enough to be considered as disturbing consistency to such a degree as to require an action to be taken by the producer of statistics. To decide on this, analyses case-by-case are called for, as at the moment a general statistical approach to this field is lacking.

The definition is tailored to WP2 by specifying the set of concepts in a particular way. The definition can easily be expanded or changed to accommodate other sets of concepts that may represent a threat to consistency, like statistical units (WP1) or variables and their definitions (WP3).

The definition is similar to, and inspired by, the ones presented elsewhere in the ESSnet on consistency, for instance in the joint Deliverable 2.1/3.1 of WP2 and WP3 (the Direction Report).

However, it differs on some points:

- while it is useful, when evaluating consistency, to be aware of regulations pertaining to that field, it is not meant here to consider regulations themselves as part of the consistency definition;
- the term "the sum" is suppressed (as it is seen merely as a numerical value, which by itself cannot help decide on existence of issues with consistency as defined above).

The analysis in the External experts' report, as well as the opinions of experts within the project, indicate that current EU regulation aims primarily to improve vertical consistency within statistical

domains, while regulation of horizontal consistency (on a country level or the EU level) between different SDs is at the moment largely missing. (Parenthetically, integration through Business Registers is an important way, currently explored, of improving horizontal consistency). While importance of both types of consistency is acknowledged, any prioritisation between them is left outside the definition of consistency.

The definition of vertical consistency in (2) could have been formulated so as not to mention statistics on the EU level, on the assumption that if all the participating countries' statistics in a domain are consistent, then also the European aggregate automatically is consistent. However, there exist statistics that pertain solely to the EU level, the EU being the unit of analysis which are not sums or aggregates of similar country level concepts. For instance, one such statistics is the EU export, which is not the sum of exports of each of the MSs. As such a statistics is produced on the basis of input from each of the country, it is appropriate to request or investigate consistency of this set of statistics, which the definition describes as the joint consistency of the country statistics and the statistics produced on the EU level.

Domains other than statistical domains, and aggregation over them, may lead to more complex situations regarding consistency. As an example, geographical domains - such as provinces or other regional components of a country - are a common way of breaking down a published statistic. The aggregate on the country level is commonly seen as needing to be consistent with the geographical components that it is broken down to, implying that there is a need also for vertical consistency within a single country. Thus, horizontal and vertical consistencies may become intertwined. However, a deep analysis of the concept of consistency is outside the scope of this report. The above definition suffices for its present purposes.

2.2. Evaluation of consistency using the questionnaire

2.2.1. Inconsistencies and perspectives

Both vertical and horizontal consistency are essential for the quality of statistics, notably for comparability and coherence in the sense of the Code of Practice. Vertical consistency is a condition particularly for international comparability of the statistics, while horizontal consistency is a condition particularly for coherence.

Inconsistencies in concepts are potentially disturbing in the perspectives of both users and producers of statistics. In the perspective of users of statistics, inconsistencies in concepts may disturb by disrupting the basis for validity of interpretations, conclusions and decisions that depend on comparability or coherence of the statistics. In the perspective of producers of statistics, on the other hand, inconsistencies in concepts may disturb in partly similar ways but further by confining work to "stovepipes" and prohibiting rational co-use of common resources such as data sources, processes, staff, training, and tools. The possibility for such co-use of common resources is advantageous for both economy and quality assurance.

2.2.2. Use of the questionnaire data

Vertical inconsistencies can be detected from questionnaire data in mainly two forms:

(i) Lack of compliance to prescriptive EU rules.
(ii) Variation between countries in choices involved in implementation of less prescriptive EU rules.

In both these forms vertical inconsistencies disturb international comparability. It may be noted that some less prescriptive rules may possibly have to be so in order to truly follow varying real-world circumstances in different countries and ensure best attainable comparability.

Horizontal inconsistencies can partly be detected from the questionnaire data similarly as in case (ii) above. That is, they are detected as a variation between statistical subject-matter domains in choices made at the implementation of rules. However, the situation is more complex for horizontal than for vertical inconsistencies, as a horizontal inconsistency for a country may be more or less due to inconsistencies in the regulations that apply. Such inconsistencies in the regulations were identified and documented in the External study.

In other words, the prevalence of horizontal inconsistencies at country level appears to be a function of in principle three arguments: first, vertical inconsistencies between countries and common rules; second, horizontal inconsistencies in the common rules; and third, lack of horizontal coordination within countries. How this works in every detail can hardly be fully examined here from the questionnaire data. Nevertheless the responses give some indicators how countries prioritize between vertical and horizontal consistency.

Hence, aiming at useful conclusions, the primary attention in the use of the questionnaire data has to be given to vertical inconsistencies, although bearing in mind their implications for horizontal inconsistencies.

2.2.3. Limitations

Although the questionnaire responses may be expected to convey very useful information, there are of course important limitations to be aware of. Like in all surveys some measurement uncertainty must be expected, even if the respondents were meticulous in forming their responses. The persons at the National Statistical Institutes (NSIs) who filled in the responses may have varying knowledge and experience, potentially affecting the meaning, quality and comparability of the responses. For instance, potentially problematic points may be:

a. Particularly some statistical domains are notably wide and cover several surveys and statistical processes, and then the person or persons assigned to fill in the questionnaire responses may have varying knowledge of the different parts.
b. Some questions ask for judgments, which may depend on the perspectives of the persons filling in the responses.
c. Generally the relationship between actual problems and awareness of them is not straightforward. A high awareness of problems could result either from the problems being overwhelming, or from a high quality ambition entailing thorough special studies of actually small problems. Likewise unawareness of problems could be either because there are none, or because you never really looked for them.

Further it has to be noted that conceptual consistency is just one basic condition for the comparability and coherence that is ultimately aimed at in the use and the production of the statistics. The finally resulting comparability and coherence of the statistics also depends on the error structure (the usual uncertainties) in the production process and how it interacts with the conceptual basis.

On the whole the questionnaire responses should be useful in identifying apparent indications of consistency problems. But they could not be expected to detect and evaluate possible actual or apparent inconsistencies more completely, as for that you would need more ambitious approaches, such as compliance monitoring exercises with coordinated expert visits to NSIs.

2.2.4. Types of questionnaire data

Generally the questionnaire questions work with a few alternative response categories which the respondent has to choose between. However, they are mostly semi-open by also allowing for free-format verbal specifications from the respondent. The choices of response categories may directly be counted in frequency tables for interpretation. The verbal specifications in responses may be relevantly coded for tabulation.

For the purpose of this report, the questionnaire questions can be classified into a few types regarding the kind of characteristics of statistical processes the responses show. These may be named as follows:

- Compliance criteria
- Harmonisation parameters
- Option indicators
- Quality indicators

These types will now be defined and described.

Type: Compliance criteria (C)

Meaning: Compliance criteria are questions on characteristics that are more or less directly governed by prescriptive rules. The responses show whether the rules are complied to or not.

Presentation and interpretation: On Compliance criteria, results may be presented primarily in frequency tables on the number of countries that comply to a rule. Naturally, less than full compliance signifies a potential deficiency in comparability or coherence.

In terms of conceptual consistency: Less than full compliance can be seen as a form of vertical inconsistency.

Note: Compliance criteria primarily detect vertical inconsistencies corresponding to case (i) mentioned in Section 2.2.2, and secondarily resulting horizontal inconsistencies.

Type: Harmonisation parameters (H)

Meaning: Harmonisation parameters are questions on characteristics that are not entirely governed by prescriptive rules and may potentially affect comparability or coherence of the statistics. The agreement vs. variation in responses indicate the potential for good vs. disturbed comparability or coherence.

Presentation and interpretation: On Harmonisation parameters, results may be presented primarily by frequency tables on the number of countries by response category chosen. The wider spread there is in the responses, the more important is the potential disturbance in comparability or coherence, in principle. However some questions on choice of methods are of a secondary character, whence variation in responses there does not necessarily signify a disturbance.

In terms of conceptual consistency: Variation in responses between countries indicates a form of potential vertical inconsistency in concepts. Variation in responses between statistical subject-matter domains indicates potential horizontal inconsistency in concepts.

Note: Harmonisation parameters primarily detect vertical inconsistencies corresponding to case (ii) mentioned in Section 2.2, and particularly in some cases also horizontal inconsistencies.

Type: Option indicators (O)

Meaning: Option indicators are questions on characteristics that affect the possibilities of complying with new or modified rules.

Presentation and interpretation: On Option indicators results may be presented primarily by frequency tables on the number of countries by response category chosen. If several countries respond similarly here, there may be a potential for further harmonisation.

In terms of conceptual consistency: Variation in responses between countries indicates a risk for potential vertical inconsistency in concepts in cases of changed rules. Variation in responses between statistical subject-matter domains may indicate a risk for potential horizontal inconsistency in concepts in cases of changed rules. The potential importance of these risks depends on the prevalence of obstacles to adaption to changed rules, mostly not to be seen more specifically from the responses here.

Type: Quality indicators (Q)

Meaning: Quality indicators are questions on characteristics that affect the efficiency in compliance to existing or potential rules.

Presentation and interpretation: On Quality indicators results may be presented primarily by frequency tables on the number of countries by response category chosen. Íf several countries respond similarly here, the rules may be likely to work well in practice and to fulfil their aims.

In terms of conceptual consistency: Variation in responses, vertically or horizontally, may indicate a risk for corresponding inconsistencies in concepts due to variation in choices made in trade-offs between quality aspects.

2.3. Evaluation from the perspective of integration of business statistics production

This subsection gives some context and perspective in which to view the work on improving consistency of the European Statistical System.

As acknowledged in the modern approaches to understanding and improving statistics production, as well as through policies of producers of official statistics (European Central Bank 2008, Eurostat 2011, UNSD not dated), production of statistics involves balancing between different quality components (generally: relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability, and coherence) under a cost constraint.

Further, observations by a number of prominent statistical producers, nationally and internationally, indicate opportunities for quality improvement and cost reduction through integration of statistics production and standardisation of tools and processes (HLG-BAS). Developments on such integration are already taking place (GSBPM, GSIM) and have been so in at least the preceding decade.

Business statistics has in a number of national statistical offices turned out to be a field where integration is first attempted and where the integration ideas have reached furthest. Experiences of The Netherlands (Braaksma 2012, Braaksma and Buiten 2012, Zwijnenburg 2012), New Zealand (Seyb et al. 2013), Canada (Ravindra 2012, Saint-Pierre 2013) and Australia (Brinkley and Scott 2012) give indications of the paths which such integration of business statistics production might take.

The developments towards increased integration currently also envision the heightened role of statistical registers in statistics production (Bakker and Daas 2012 and the whole of the special issue which the article introduces). The Business Register is one of the cornerstones of such an integrated system (Wallgren and Wallgren 2007). The degree to which the Business Register will be used as the sampling frame is a matter of balance between the different components of statistics production and the associated costs and quality requirements. For instance, in order to reflect the target population for a certain reference point in time as well as possible, the sampling frame ought to be 'frozen' after an update as close as possible to that reference point. However, in order to facilitate field operations, the sampling frame ought to reflect the state of affairs (contact addresses, etc) as close as possible to the start of field operations. As another example, coordination of samples may be arranged so as to reduce the response burden in a population, but may at the same time deteriorate accuracy due to mutual exclusion of important units from the sample.

How to optimally perform balance of the different quality components within a cost frame may likely be best solved through an integrated approach in a systemic way. Developments towards such solutions exist, as the references above indicate. While incremental progress can be achieved through eliminating obvious treats to consistency, taking a broader view may provide a system optimised over a more full range of requirements on quality and cost.

3. Methodology of the analysis

In carrying out the evaluation, the approach and content were informed by the theoretical background and analyses presented in the preceding Section 2.

3.1. Data

The evaluation took its starting point in the results presented in Deliverable 2.7. The final data became available in June 2013, so in order to initiate the evaluation process within the time frame of the project, some data analysis was carried out in parallel to that which led to Deliverable 2.7.

Also, in preparing empirical results for the evaluation, further tables in addition to those of Deliverable 2.7 were created, particularly with the goal of gaining a perspective on variability between the member states on questions of interest. While this aspect is essential for establishing vertical consistency, it was agreed within the ESSnet project on consistency that data on individual MSs will not be available in the report. However, these data were taken into account in analyses and evaluation, without revealing in the report any particular country.

Another component added to each relevant table were the open, textual (qualitative) responses/comments made by respondents where the possibility was provided for that in the questionnaire. These were analysed by qualitative means, mainly by a thematic analysis (Creswell 1994), and integrated with the qualitative results to produce a broader view on the topics discussed.

3.2. Major themes in the analysis

In order to be able to identify major issues to consistency, a prior analysis of the content of questions in the questionnaire identified six major themes to which these questions could be grouped. These major themes are presented in Table 1, along with the questions that contribute to them. Within each theme, the questions could also be seen as belonging to one of the four types defined in Section 2.2.

A business register is a crucial means by which to achieve better consistency through a systematic way of creating, maintaining and updating target population registers and sampling frames. Thus, when evaluating the results it is important to highlight the relation between the statistical domains (SDs) on the one hand and the business register (BR) on the other. The questionnaires were constructed to provide for that, as a number of questions covered the same or similar phenomena both from the SD perspective and the BR perspective. Therefore, in evaluating the results, we take these two perspectives into account at the same time, which is reflected in the grouping of questions to themes presented in Table 1.

Note: when identifying questions by their number, we follow the numbering introduced in Deliverable 2.6, but use the simplification by which the statistical domains question numbers (Section 3.1 of Deliverable 2.6) are preceded by “SD”, while the business register question numbers (Section 3.2 of Deliverable 2.6) are preceded by “BR”. Thus, for instance, “SD 1.1” refers to question 1.1 in the statistical domains set of questionnaires, “Do you have any specific difficulties in covering the target population in your country – Please mark all applicable” (p.7 of Deliverable 2.6).

In presenting the analyses results, first those major themes are covered where both components of the questionnaires (BR and SD) are contributing with information, and then those where only one of the components contributes.

Table 1 - Grouping of the questions into major thematic areas. In parentheses, type of question (see Section 2.2.4) is given[1].

	Major theme	BUSINESS REGISTERS		STATISTICAL DOMAINS	
1	Frame coverage	BR 1.1	(C)	SD 1.1	(C)
		BR 1.2	(C)	SD 3.7	(C)
		BR 1.3	(C)	SD 3.8	(C)
		BR 1.4	(C)		
		BR 1.5	(C)		
2	Use of BR by SDs	BR 2.1	(C)	SD 3.1	(Q)
		BR 2.2	(C)	SD 3.2	(Q)
	including				
	Relations between BR and SD	BR 2.5	(H)	SD 3.10	(H)
		BR 2.6	(H)	SD 3.11	(H)
		BR 2.7	(H)	SD 3.12	(H)
		BR 2.8	(H)	SD 3.13	(H)
		BR 2.9	(H)	SD 3.14	(H)
		BR 2.10	(H)	SD 3.15	(Q)
		BR 2.11	(H)		
		BR 2.12	(H)		
		BR 2.13	(H)		
		BR 2.14	(H)		
		BR 2.15	(H)		
3	Temporal aspect	BR 1.6	(H)	SD 3.3	(H)
		BR 1.7	(H)	SD 3.4	(Q)
		BR 1.8	(H)	SD 3.9	(C)
	Including				
	BR management	BR 2.3	(H)		
		BR 2.4	(H)		
4	BR maintenance	BR 1.9	(H)		
		BR 1.10	(H)		
		BR 1.11	(H)		
		BR 1.12	(H)		
		BR 1.13	(H)		
		BR 1.14	(H)		
		BR 1.15	(H)		
		BR 1.16	(H)		
		BR 1.17	(H)		
5	Reference periods			SD 4.1	(C)
				SD 4.2	(H)
				SD 4.3	(C)
6	Sampling methods and sample coordination			SD 3.5	(Q)
				SD 3.6	(H)

3.3. Coding of Question SD 3.3

Question SD 3.3: When do you take your main sample? (Type: Harmonisation Parameter)

This question was an open ended-question, accepting any textual response and offering no response categories. In order to summarise the results, a system for categorisation of the responses was created. The open responses were assigned a three-digit code of the form `NMM` with meaning of these positions given in Table 2.

[1] An extended version of this table, with question content added, is reprinted as Appendix 1.

Table 2 - Coding for responses on Question SD 3.3.

N	Content	MM content
0	Sample taken the year two years before the reference year	Month in the year (01 - January, ..., 12 - December)
1	Sample taken the year before the reference year	Month in the year
2	Sample taken the reference year	Month in the year
3	Sample taken the year after the reference year	Month in the year
4	Sample taken the year two years after the reference year	Month in the year
5	No relation to reference year mentioned, only a time/date when sample is taken	Month in the year
6	Sample taken less or more frequently than annually (but not monthly)	01 - less frequently 02 - more frequently
7	No sampling (census or admin data)	00
8	Response in relation to time of start of the field period	Number of months before the start
9	Other	00 - difficult to understand, or claim of no regular time for sampling 01 - can be understood, but does not give a point in time 02 - gives the time of creating the *frame* 03 - gives two or more responses

So, for instance, “In October before the reference year” is assigned the code 110, and “January every year. Date of last selection - 30th of January” is assigned the code 501.

The coding was done independently by two coders, which resulted in an intercoder agreement of 68%. After that, a reconciliation of the assigned codes was performed, resulting in the codes used for summarising the results and analysis.

Results reported further below in the report, in Table 23, represent those responses where a month in which the sample is taken could be deduced from the responses; that is, the responses coded with one of 001 to 512.

3.4. Other methodological remarks

Measurement interdependence of response values. It may be noted that some interdependence between response values for different statistical domains (SDs) is likely. Namely, within the staff of responding NSIs, partly the same persons may have been involved in preparing responses for more than one SD. This means that counts of response values over SDs and the participating countries jointly must not be taken as indicators of measurement precision in the results, as the counts may be inflated compared to an ”effective” number of observations. Likewise agreement in qualitative responses must not be over-interpreted.

Country size classification. Variation of practices regarding consistency can have differential effect on European statistics that depends on “economic size” of (i.e. the volume of economic activity in) the country. In some of the analyses attention was therefore also given to the size of economy of the country. An attempt was made to broadly classify the countries as large, medium-sized or small. The group of large countries comprises the five largest in terms of combination of population and gross domestic product (GDP): France, Germany, Italy, Spain and United Kingdom. The group of small countries comprises the eight smallest in terms of combination of population and GDP: Cyprus, Estonia, Iceland, Latvia, Lithuania, Luxembourg, Malta and Slovenia. All the other participating countries are classified as medium-sized.

4. Analysis

4.1. Frame coverage

The major theme of frame coverage includes results on restrictions to coverage, cut-offs in administrative sources, issues of under coverage and over coverage, and discrepancies between target populations and corresponding sampling frames. It also includes estimation of coverage error and adjustments made to reduce its impact. Many of the topics in this set were by the questionnaires covered from both the perspective of business registers (BRs) and statistical domains (SDs), which is why in the analysis these two types of sources are taken into account simultaneously.

4.1.1. Restrictions in coverage

Question BR 1.1: Are there any restrictions to coverage for your register i.e. thresholds, with offered response categories ***Market/Non-market activity, Legal form, NACE activities, Turnover, Employment*** (Type: Compliance Criterion)

Question SD 1.1: Do you have any specific difficulties in covering the target population in your country, with offered response categories similar to, but not the same as in, BR 1.1, that is ***Market/Non-market activity, NACE activities, Active enterprises and Other*** (Type: Compliance Criterion)

The responses to the BR question generally indicate relatively low frequencies of restrictions to coverage. About 90% of the responses over all the response categories did not state any restriction. Of the remaining 10% (which were numerically few, 14 instances[2]), restrictions were mostly noted regarding NACE activities, employment, and turnover.

Table 3 - SD 1.1: Difficulties of SDs in covering the target population.

	Mark./Nonm.	NACE	Active	Other	Total	No
SBS - Annual Structural Statistics	8	3	5	1	17	7
SBS - Industry	3	1	5	1	10	15
SBS - Distributive trade	4	1	5	2	12	11
SBS - Construction	3	0	5	2	10	14
SBS - Business Demography	3	0	2	2	7	21
STS - Industry	2	0	4	1	7	21
STS - Construction	4	0	6	6	16	12
STS - Retail Trade & Repair	2	1	3	1	7	20
STS - Other Services	4	4	3	3	14	15
PRODCOM	2	0	1	1	4	17
Inward FATS	4	0	2	3	9	13
Outward FATS	4	2	4	8	18	2
FDI	1	2	0	3	6	12
ICT	0	3	2	1	6	18
R&D	1	4	5	3	13	12
Innovation Survey	2	0	2	1	5	20
Job Vacancy Statistics	2	6	1	2	11	8
Labour Cost Survey	2	5	2	3	12	14
Structure of Earnings Statistics	2	5	3	1	11	13
Total	53	37	60	45	195	265

The responses for SDs report a not as good coverage as for the BR, with about 42% of the responses over all the categories stating a restriction. Restrictions regarding active enterprises received the most responses, followed by restrictions related to coverage of market/nonmarket activities, followed by the *NACE* and *Other* categories; however the differences in frequencies were not too large. Among the SDs stating restrictions in many member states were SBS - Annual Structural Statistics (13 countries), Outward FATS, R&D, and STS - Construction (11 countries each); the SDs with restrictions stated in

[2] Here and elsewhere in this report, when Tables are not provided, they may be found by reference to tabulation of the corresponding question given in Deliverable 2.7.

least number of member states were PRODCOM (3 countries), STS - Industry and STS - Retail Trade & Repair (4 countries each).

Responses related to restrictions in coverage in the response category of *Market/non-market activities* reveal two fundamental issues: that there is not a clear definition of this distinction, and that (accordingly) this distinction is not well reflected in relevant sampling frames, like the Business Register (BR).

> *We miss clear and applicable definition of market / non-market activities, [which] is a general problem for whole business statistics.*
>
> *This difference is not identified in our business register.*
>
> *How to define market - non market activities? What is the operational definition?*

Of restrictions in coverage of *NACE* categories, those most often mentioned were - in the decreasing order of frequency - P, Q, O, R, T, A and S; also mentioned a few times were E, K, L and U[3]. (An issue with this question was that in some instances it was understood as referring to response rates.)

> *The NACE sections E/P/Q/R/S and T are not covered. [OFATS, medium-size country]*
>
> *Covering NACE Sections O to U poses problems because the population is considerable, but market activities are limited [R&D, medium-sized country]*
>
> *NACE section O [Job Vacancy Statistics, large country]*
>
> *O to S are not covered in the disseminated indicators [Job Vacancy Statistics, large country]*
>
> *it is difficult to cover section A (companies in agriculture, forestry and fishing) and section O (public administration and defence, compulsory social security) due to their small size [Job Vacancy Statistics, medium-sized country]*
>
> *Because of the exemption of the NACE section O, local units of enterprises in section O are not covered. This affects in particular public employees in the NACE sections P, Q, R and E [Job Vacancy Statistics, medium-sized country]*

The bulk of the responses regarding restrictions in coverage of *Active enterprises* concerned insufficient timeliness of the sampling frame - most often, the BR - which resulted in either under coverage of recently created units or contact attempts with units that have ceased to exist.

Restrictions in coverage of *Turnover* are in a few cases - in the context of Q BR 1.1 - reported to exist; however, as the restrictions are mainly caused by the thresholds in admin data that feed into the BR, this is further elaborated at Q BR 1.2 below.

Restrictions in coverage of *Employment* could stem from a number of sources. These included units with zero employees, employees not subject to social insurance, employees of EU institutions, and - up to recently in a country - coverage of units below former Employment Tax threshold.

Only BR from one country has reported restrictions in coverage of *Legal forms*, and it concerns independent workers having no employees.

> *Independent workers having no employees and having activities not subject to VAT are not covered. Independent workers having no employees and an annual turnover of less than 10000 Euros are probably not completely covered.*

Restrictions in coverage due to *Other causes* referred to:

- difficulties in reaching target population as it is not well defined and lacks a usable sampling frame,

[3] For a complete list of NACE Rev.2 categories on levels 1 (Sections) and 2 (Divisions), please refer to Appendix 2.

"For Outward FDI the target population is 'nonresident affiliates', while respondents are the resident investors… the target population is unknown [FDI] *"*; *"We have a general problem on finding target population. There is no source which includes information on all R&D units* [R&D] *"*

- threshold on units covered (in terms of size, monetary amount, properties, etc), which coincides with items covered in Question BR 1.2, directly below,
- difficulties identifying groups of companies due to lack of admin data (full financial statement).

Even here, occasionally there was a comment on lack of coverage due to nonresponse created by increased response burden and at the same time making the survey voluntary.

The perception that BR has about its coverage is apparently better (90% of the responses indicate no issues) than that of the SDs (58% indicate no issues) about coverage of their target population by the sampling frame that is used. It is currently not clear whether those SDs using a BR have a better perception about the coverage than those that do not rely on the BR.

Question BR 1.2: Are there any cut-off thresholds for the administrative data sources on your business register? with yes/no responses offered for specified admin data sources *VAT, Personal income Tax, Corporate Tax, Published business accounts, Social Security data, Central Bank data, Commercial data provider, Other* (Type: Compliance Criterion)

Fifteen of the participating countries' BRs reported no restrictions. Of those who did, most frequently the restriction was in the category of *VAT* admin data. The range of the thresholds, for the 13 BRs reporting it, was between the low of about 10 K EUR (AT, DE, DK, EE, FI, LU, MT, NO), the medium of 50 K EUR (IE, RO, SI) and the high of about 80 K EUR (CH, UK).

Majority of the few responses on the other sources were also expressed in terms of amounts, which generally corresponded to the VAT thresholds.

Question BR 1.3: Has an estimate been made of the coverage of your Business Register in relation to the target population?, an open-ended question (Type: Compliance Criterion)

Question BR 1.4: Do you have any issues of under coverage in relation to your target population?, an open-ended question if affirmative (Type: Compliance Criterion)

Question BR 1.5: Do you have any issues of over coverage in relation to your target population?, an open-ended question if affirmative (Type: Compliance Criterion)

Of the 11 BRs that reported having estimated the coverage, a majority reported the coverage as being "*good*", "*nearly full*", "*almost 100%*", "*99.99%*", and "*100%*". The four responses that described how the evaluation was done mentioned comparisons of annual balance sheets, tax registers, and social security registers; turnover tax statistics of the final tax declarations; and using census data. Six BRs mention no plans for evaluation of coverage.

Responses on under coverage indicated issues coinciding with those reported for Questions 1.1, above: nonmarket activities; NACE sections A, O, P, Q; and units with thresholds on VAT, employment or similar. Additionally, issues with identifying local units have been mentioned.

There is no administrative source for local units, which may lead to under coverage.

Responses on over coverage mostly indicated issues with a lag in updating the admin data that feed into the BR, resulting in economically inactive units being represented in the BR. One instance of duplicates in the BR was mentioned. This is further elaborated in next question.

4.1.2. Ability of BR frames to represent the target populations

Question SD 3.7: Are you aware of any discrepancies between your target population and the frame population as available from your Business Register?, with ticking yes, followed by an open comment, offered for *Under coverage, Over coverage, Timeliness, Quality/Accuracy of variables, and Other* (Type: Compliance Criterion)

Somewhat less than 30% of the responses indicated no discrepancies between a SD's target population and the frame population available to it in the BR (Table 4). Innovation survey and Inward FATS have been in the largest number of participating countries reported as not having discrepancies, followed by R&D, Business Demography and the STS modules. Of the discrepancies, under coverage and over coverage were most frequent, followed by timeliness; quality and other forms of discrepancies were reported less frequently. The SDs reporting discrepancies in the largest number of participating countries were LCS, SES, the STS modules, and the Job Vacancy Statistics.

Table 4 - Question SD 3.7: Awareness by SDs of discrepancies between the target and the frame population.

	Undercov.	Overcov.	Timeliness	Quality	Other	Total	No
SBS - Annual Structural Statistics	6	6	3	3	2	20	9
SBS - Industry	6	6	5	4	2	23	7
SBS - Distributive trade	6	7	3	2	2	20	8
SBS - Construction	6	6	5	3	1	21	8
SBS - Business Demography	6	5	6	3	2	22	10
STS - Industry	8	8	5	2	3	26	12
STS - Construction	8	7	3	3	3	24	10
STS - Retail Trade & Repair	7	9	4	2	2	24	12
STS - Other Services	6	8	4	5	2	25	10
PRODCOM	7	6	3	2	2	20	6
Inward FATS	5	4	1	1	1	12	13
Outward FATS	6	3	2	1	3	15	8
FDI	9	2	5	1	2	19	4
ICT	5	6	6	2	2	21	8
R&D	9	5	5	1	1	21	10
Innovation Survey	5	6	4	2	2	19	15
Job Vacancy Statistics	9	9	5	1	0	24	5
Labour Cost Survey	8	12	8	5	2	35	6
Structure of Earnings Statistics	6	11	6	2	3	28	6
Total	128	126	83	45	37	419	167

Under coverage

Of primary note was the general issue of a time lag between a unit becoming such that it enters the target population (e.g. unit established, or crossing a threshold, or establishing a foreign affiliate, or initiating substantial R&D activity) and that event becoming registered in the relevant sampling frame - the BR or the specifically developed frames by SDs. While this aspect is treated further under Timeliness below, it may be noted even here that the time lag varied from fairly short to rather considerable, and that the respondents deemed the impact as varying from negligible, over medium, to potentially serious.

> *22% of enterprises active in 2011 according to an administrative source on compuls*[ory] *social contributions were not [present] in the 2009 BR used to draw the sample for the 2011 survey waves.* [Job Vacancy Survey, large country]

Further of general note was the issue of thresholds on administrative data that feed into the sampling frame, or into the BR if BR is used as the sampling frame, which may have an impact on under coverage in the statistics produced.

> *There are enterprises who don't declare their turnover to the Tax Office (BR main source of information of activity), but report later in their Annual Report to the Commercial Register. Due to this we are finding some active enterprises long after our frame has been compiled.* [SBS, small country]

> *Based on comparison with enterprise level OFATS data from [two other countries], there is evidence to suggest that coverage is not 100 percent. For these two UCIs, estimated coverage of the target population in 2009 was 75 - 88 percent of enterprises and 98 - 99 percent of employment.* [IFATS, medium-sized country]

> *A very rough estimate of under coverage is 2-3 percent of the total production value.* [PRODCOM, medium-sized country]

Additionally, in one medium-sized country the BR was reported to show deficiencies in coverage of NACE sections O to S (with further consequences for public employees in NACE sections P, Q, R, and E), leading to under coverage of 9.1% of the employees.

In some cases, the responses indicated under coverage by design of the survey (a threshold on volume or ownership) rather than actual under coverage in the sampling frame or the BR.

For some specific SDs, in particular R&D, there was an additional issue of identifying the target population – which is defined by the activity to be measured itself, thus not exactly known in advance. Information about units carrying out R&D activities is generally not recorded on the BR. Instead, R&D statistics production units are creating their own sampling frames based on previously accumulated information and best existing sources, however in spite of that under coverage can occur. Some NACE sections are canvassed less intensely than others, and threshold rules for the sampling frame are applied (e.g. above 10 employees, start-ups excluded); in one case, the sampling frame did by design include only previously known R&D performers, which may lead to serious under coverage if not compensated for by updating the frame. In one country it is estimated that on the level of investment in specific technological fields under coverage could be severe.

> *Impact on total spending is [marginal] but estimates on the level of investment in specific technological fields could be severely affected* [R&D, large country].

Specific studies were carried out in a number of MSs to estimate the actual extent of under coverage.

Over coverage

The comments generally agreed between the SDs in that while there naturally is an over coverage due to units ceasing activity since updating of the sampling frame, this is seen as having minimal, negligible effect on consistency of produced statistics.

> *Frame population is reflecting the January status of active units. Some of them are not active at the time of survey conducting. (Low impact)*

> *10% of sample units drawn from the 2009 SBR for the 2011 survey waves were no more active in 2011 (according to an administrative source on compulsive social contributions for 2011)*

> *total over-coverage rate for the last survey is 5.6%*

> *The reasons [...] can be put into 2 categories: 1. The local unit has less than 10 employees in practice (5% of the sample) 2. The local unit does not exist. The reasons for this non-existence can be bankruptcy, merger, liquidation or discontinuance of business. (1,8% of the sample)*

Note: from the latest response it can be inferred that there is a corresponding issue of under coverage of those businesses whose employment has risen above a threshold, which was not sufficiently highlighted in the preceding subsection on under coverage. Its impact may likely be of the same order of magnitude (for instance, around 5% of the sample size in the given example) and potentially of low impact on some important economic variables (e.g. total country production or turnover), but likely of higher impact on some specific breakdown categories or when studying the population of only start-ups or gazelles.

A major reason of under coverage and over coverage is thus related to the timeliness aspect, the next offered response category.

Timeliness

The comments generally agreed between the SDs and indicated a varying level of concern with effects of the time lag. Here, responses that indicated longest lags or most important consequences are presented.

> *BR has a time lag of 2 years* [large country]
>
> *Business Register is two years old* [medium-sized country]
>
> *BR data is collected with a lag of 2 years* [small country]
>
> *At the time of sampling data available from Business [R]egister relate to the year t-2 which results into over and under-coverage errors* [medium-sized country]
>
> *Investment Register provides information for [...] statistics with a 18 month delay* [large country]
>
> *For local units, the information is available only for the n-1 year in the Business register while the target population is that of the n-year. Similarly, the employees sampled are those who worked at the end of the previous year, while the target population is that of the n-year* [large country]
>
> *The group register in the NSI Statistical Business Register is updated with a 10 month time lag. The annual FDI survey is sent to respondents in April so the group register has to be adjusted with information from other sources* [medium-sized country]
>
> *The timeliness of the EGR is a large issue for us as at the moment the most uptodate frame is not ready in time for the start of our sampling processes* [large country]
>
> *The poor timeliness of the Business Register strongly affects the ability of the survey staff to effectively contact active enterprises at the current address* [large country]
>
> *There is sometimes a time lag in regard to the ID number of the units, which can make it difficult to combine data* [medium-sized country]
>
> *Delay in business register is usually a few months* [medium-sized country]

Another response highlighted issues of choice between different administrative sources and time of their availability, which when not satisfied may lead to under coverage and over coverage. (The response was given in the context of Q SD 3.2, but referred to from Q SD 3.7.)

> *The file includes a size class variable based on employment at the end of the second quarter of the year before, so naturally one year later some shifting in size classes will have occurred. As balance sheets (that have actual employment numbers) generally are not publicly available until at least six months after the end of the fiscal year, the register of employers of the national social security office is the best source available to us at the actual moment when we have to do our sampling.*

This situation depicted in this response impacts sampling frame formation (through a threshold value) and breakdown issues (thorough shifting of size classes).

Finally, a response highlighted incongruence when striking balance between quality components in production of statistics.

> *Up to now, since we needed to coordinate with another survey, the sample was drawn a bit early for the [SD] (before the end of the reference year). This meant that there was no BR for the reference year available at the time. In the future, the sample selection will be done closer to the end of the reference year (no coordination with other surveys).* [Emphasis added]

More on timeliness is given in the major theme concerning the temporal aspect (Section 4.3).

Quality

Quality has apparently been understood broadly by the respondents, thus including timeliness, already reported on above. A number of the responses thus reiterate the issues of a time lag between the event and it being recorded in the frame used for sampling. An additional time issue, not raised before, is of the point in time in which a change occurs.

> *In case of change in activity it is not always clear from which point in time this is to be effected. Thus diversions in classification can exist for certain reference periods*

Some additional issues were mentioned.

> *NACE codes are of moderate quality in low size classes (no profiling and little feedback due to small sample size)*
>
> *There are some misclassification problems*

Apparently, these comments aim at accuracy of the administrative (and thereby sampling frame) data.

Other

The responses reiterated the types of issues already given in the preceding response categories.

> *1. Relatively high degree of misclassification particularly of small units (natural persons and legal persons without employees, which are not VAT payers). 2. Delay in the availability of administrative data, which do not allow to identify inactive or newly active units (again, this issue refers especially to small units)*
>
> *The administrative database that is used to produce the sample is known to have under-coverage of low-paying jobs and jobs from very small businesses. It is not possible to quantify the scale of under-coverage, but research involving targeted samples of these under-represented jobs has shown that the impact on earnings estimates is small*

Question SD 3.8: Do you adjust for discrepancies between your target population and frame population?

Table 5 - Question SD 3.8: Adjustment by the SDs for discrepancies between the target and the frame population.

	Yes	No	Not req.	Total
SBS - Annual Structural Statistics	7	5	5	17
SBS - Industry	8	6	5	19
SBS - Distributive trade	8	5	4	17
SBS - Construction	8	6	3	17
SBS - Business Demography	6	7	5	18
STS - Industry	7	9	5	21
STS - Construction	6	9	8	23
STS - Retail Trade & Repair	8	6	5	19
STS - Other Services	9	8	7	24
PRODCOM	4	7	5	16
Inward FATS	3	5	5	13
Outward FATS	2	7	7	16
FDI	3	7	2	12
ICT	9	4	4	17
R&D	5	5	4	14
Innovation Survey	6	5	6	17
Job Vacancy Statistics	7	5	3	15
Labour Cost Survey	10	6	3	19
Structure of Earnings Statistics	9	7	2	18
Total	125	119	88	332

The open responses indicated actions and procedures for improving the sampling frame so that it better corresponds to the actual target population. This included measures prior to or in the data collection phase, and also those after data collection. The former concern actions to eliminate from the sample those units that make over coverage, and to bring into the sample or frame those units that should have existed in the sampling frame but have not (mostly by using additional sources of

data). The latter concern using some weighting or other probabilistic, modelling approach to correct for inadequacy of the sampling frame. About a third of those SDs who responded reported doing some form of adjustment, whereas the rest reported doing no adjustment.

4.2. Relations between the business register and the statistical domains

This major theme covers the analyses concerning a broad area of use of the business registers (BRs) in the responding statistical domains (SDs), the BRs' relation with the SDs, including mutual updating and feedback.

4.2.1. BR as a resource for sampling

Question BR 2.1: Which domains use the Business Register as a source for sampling?
(Type: Compliance Criterion)

Table 6 - Q BR 2.1: Use of the Business Register by SDs as a source for sampling.

	Live BR	Frozen BR	Other[4]	Total	Not used	Not known
SBS - Annual Structural Statistics	3	23	0	26	0	0
SBS – Industry	5	21	0	26	1	0
SBS - Distributive trade	5	22	0	27	1	0
SBS - Construction	4	22	0	26	1	0
SBS - Business Services	5	21	0	26	1	0
SBS - Business Demography	5	19	3	27	1	0
STS - Industry	7	18	0	25	1	2
STS - Construction	6	18	0	24	1	3
STS - Retail Trade & Repair	6	20	0	26	1	2
STS - Other Services	6	19	0	25	1	3
PRODCOM	5	14	0	19	2	4
Inward FATS	2	17	1	20	2	4
Outward FATS	2	13	1	16	4	4
FDI	3	7	0	10	9	5
ICT	4	18	0	22	2	4
R&D	3	18	0	21	4	2
Job Vacancy Statistics	3	15	1	19	3	3
Struct. of Earnings/Labour Cost	0	21	3	24	0	3
Vocational training	0	15	0	15	3	9
Other	2	3	0	5	1	2
Total	76	344	9	429	39	50

The SDs most frequently reported by the BR to be relying on the BR as the sampling frame were SBS and STS, as well as LCS. On the opposite side, FDI, Vocational training and OFATS were the SDs relatively least frequently reported to rely on the BR.

Likely, some of these latter SDs (not using the BR) are either total surveys (censuses) (perhaps above a certain threshold), or are in stronger need to keep own sampling frames due to e.g. specific target populations, that is, where the BR does not suffice (R&D and FATS being the examples).

When responding to what was meant by the *other* response category, the respondents commonly indicated that their SD did a census, that is, that there was no sampling. However, one may note that the BR could still have been used as a frame that contained the population to be contacted and surveyed in totality.

[4] For the response category *other*, only those responses were counted where no other of the response categories was marked.

Question SD 3.1: What [BR] frame do you take your main sample from? (Type: Quality indicator)

Table 7 - Question SD 3.1: What [BR] frame do SDs take their main sample from.

	Live BR frame	Frozen BR frame	Other	Total
SBS - Annual Structural Statistics	6	15	3	24
SBS - Industry	6	15	4	25
SBS - Distributive trade	6	14	3	23
SBS - Construction	5	16	3	24
SBS - Business Demography	3	5	7	15
STS - Industry	7	14	7	28
STS - Construction	4	16	7	27
STS - Retail Trade & Repair	7	14	6	27
STS - Other Services	6	16	6	28
PRODCOM	9	8	3	20
Inward FATS	1	12	9	22
Outward FATS	2	5	13	20
FDI	3	2	13	18
ICT	8	14	2	24
R&D	2	12	10	24
Innovation Survey	3	16	5	24
Job Vacancy Statistics	4	10	4	18
Labour Cost Survey	6	14	6	26
Structure of Earnings Statistics	8	12	4	24
Total	96	230	115	441

The SDs most notably reporting not using the BR were FDI, OFATS and the Business Demography. On the other hand, STS, the other SBS annexes than Business Demography, ICT and Innovation survey have to a largest degree reported their use of the BR (frozen or live). These responses coincided to a large extent with those of the BR (in Table 7), with the exception that Business Demography was by the BR not seen as deviating from the other SBS annexes in its use of the BR, whereas the Business Demography respondents did report lower use of the BR in comparison to reports of the other SBS annexes.

A considerable proportion of the responses in the rather voluminous *other* category (with larger count than the *Live BR frame* category) came from FDI, IFATS, OFATS and R&D. These responses (as well as some given by the other SDs) concerned use of other sampling frames (not the BR), like registers of transactions that the central banks keep, or other – specifically developed – registers for the purposes of the SDs. In some of the cases, the BR was used as an aid amongst others in creating own frame. Some respondents who opted for the *other* category indicated in their response that they do not do sampling but rather conduct a census. Additional reason for choosing this category was to indicate that their sample was a subsample of another survey (e.g. IFATS sampling from SBS).

But, if the BR was used as a sampling frame, it was predominantly the response category *Frozen BR frame* that was chosen. However, about 20% of the responses indicated use of the live frame, and somewhat more prominently so in PRODCOM, ICT and SES, as well as in LCS and the STS and SBS modules. In one participating country (medium-sized), the live frame is used extensively; in three participating countries, it is not used at all, and in a few more it is used very seldom (that is, in only one or two of the SDs).

Question SD 3.2: If you do not sample from your Business Register ... why not? with the same response categories as in the Q SD 3.7 on coverage, namely with ticking yes followed by an open comment, offered for ***Under coverage*****, *Over coverage*****, *Timeliness*****, *Quality/Accuracy of variables*****, and *Other*** (Type: Quality indicators)

Table 8 - Question SD 3.2: Reasons for not sampling from the BR?

	Under cov.	Over cov.	Timeliness	Quality	Other	Total
SBS - Annual Structural Statistics	0	0	0	0	2	2
SBS - Industry	0	0	0	0	2	2
SBS - Distributive trade	0	0	0	0	2	2
SBS - Construction	0	0	0	0	2	2
SBS - Business Demography	0	0	0	0	0	0
STS - Industry	0	0	0	1	5	6
STS - Construction	0	0	1	1	2	4
STS - Retail Trade & Repair	0	0	0	0	4	4
STS - Other Services	0	0	1	0	2	3
PRODCOM	1	0	0	0	1	2
Inward FATS	0	0	0	0	6	6
Outward FATS	1	0	0	2	10	13
FDI	4	1	2	3	9	19
ICT	1	1	1	1	1	5
R&D	1	2	1	1	5	10
Innovation Survey	0	0	1	0	1	2
Job Vacancy Statistics	1	1	2	1	1	6
Labour Cost Survey	2	2	1	1	4	10
Structure of Earnings Statistics	0	0	0	0	3	3
Total	11	7	10	11	62	101

The first four response categories (Under coverage, Over coverage, Timeliness, Quality) generally mentioned the same issues already presented in conjunction with the Question SD 3.7, in the frame coverage area above. These comments were here comparatively few.

The largest proportion of responses was however in the category Other. The responses concerned a range of reasons: doing a census rather than a sample survey (FDI, IFATS, SBS), prohibited access to BR due to national legislation (in several MSs, two of them large), existence of sources (admin or other registers) with information more suitable to the SD in question (Labour Cost/Structure of Earnings, SBS, STS) than what BR is, as well as the BR not having the information needed in order to be a useful sampling frame (IFATS, OFATS).

Question BR 2.2: Are steps taken to move surveys towards using the Business Register? (Type: Compliance Criterion)

With nonresponse from 5 BR respondents, 25% (six of 24) of the responses reported that steps are taken to move surveys towards using the BR. These generally report on the potential or ongoing developments to either introduce the BR as the sampling frame in an SD or to develop a broader system for economic statistics production in which the BR would be a central element.

4.2.2. Provision of changes from BR to SDs

Question BR 2.5: Do you provide information on changes to the register for those surveys using frozen extracts from the Business Register? (Type: Harmonization parameter)

Question SD 3.10: Is information on changes to the Business Register since you selected your sample available to you? (Type: Harmonisation Parameter)

Table 9 - Question SD 3.10: Whether information on changes to the Business Register since SDs selected their sample is available?

	Yes	No	Total	Not applicable
SBS - Annual Structural Statistics	21	2	23	0
SBS - Industry	21	3	24	0
SBS - Distributive trade	20	2	22	0
SBS - Construction	20	2	22	0
SBS - Business Demography	19	4	23	2
STS - Industry	23	3	26	1
STS - Construction	20	4	24	3
STS - Retail Trade & Repair	21	4	25	2
STS - Other Services	23	4	27	1
PRODCOM	15	3	18	2
Inward FATS	16	1	17	3
Outward FATS	9	2	11	7
FDI	14	0	14	4
ICT	18	4	22	0
R&D	19	2	21	1
Innovation Survey	20	3	23	1
Job Vacancy Statistics	13	3	16	1
Labour Cost Survey	21	3	24	1
Structure of Earnings Statistics	19	5	24	0
Total	352	54	406	29

Twenty of the 29 BRs (around 70%) reported providing changes to the SDs that use the frozen BR sampling frame.

On the corresponding question to the SDs (Table 9), 87% percent (352 of 406) of the volume of responses reported that changes since sampling in the frozen frame are available to them.

Accordingly, in both cases a high proportion of responses indicate availability of the required information, but the proportion among the SDs is higher than among the BRs (87% and 70%, respectively). The difference perhaps hinges more on the active vs. passive way of availability of data ("providing" vs. "available") then on the substantial difference, as they both describe the same reality.

Table 9 shows that information on changes since sampling is mostly available, but respondent comments indicate that even so may not always be considered useful or made use of. Cf. further the following Question (SD 3.11).

> *It is available on request respectively the units in the Business Registers can be viewed via specific applications. [Job vacancy stats]*
>
> *The national register is updated daily and all frozen frames are available to us on request. The EGR updates are not available to us in order to meet our needs. [OFATS]*
>
> *This information is available, but the samples are either selected from a Reference List or are Censuses of their target populations. [R&D]*
>
> *Business Demography produced directly from a full extract of the business register. [SBS IX]*

Question BR 2.6: What information is fed back to the survey areas and how frequently? (Type: Harmonization parameter)

Table 10 - BR 2.6: Whether information is fed back to the survey areas and its frequency.

	Weekly	Monthly	Quarterly	Other	Total	Not provided
Births	0	6	2	9	17	1
Deaths	0	6	4	9	19	1
NACE	1	6	2	10	19	1
Employment	0	6	2	11	19	1
Turnover	0	3	2	8	13	4
Structure	0	3	4	8	15	3
Name/address	1	4	1	10	16	1
Other	0	0	1	3	4	1
Total	2	34	18	68	122	13

Question SD 3.11: What information is available from Business Registers relating to your frame and how frequently? (Type: Harmonisation Parameter)

Table 11 - SD 3.11: Whether information from Business Registers relating to SDs' frame is available and its frequency.

	Weekly	Monthly	Quarterly	Other freq	Total	Not available	Not known
Births	48	130	34	130	342	8	13
Deaths	46	133	36	124	339	8	11
NACE change	49	79	40	188	356	6	18
Employment change	18	79	76	180	353	9	13
Turnover change	13	37	51	193	294	29	21
Structure change	43	93	20	134	290	36	22
Name/address change	60	110	22	132	324	18	11
Other change	1	14	3	37	55	11	19
Total	278	675	282	1118	2353	125	128

In both sources of information (the BR, Table 10, and the SDs, Table 11), changes involving new and closed units, NACE and employment were the ones more available to SDs from the BRs, than the changes in turnover and in structure; name/address change was somewhat in-between the two mentioned groups.

Regarding the Other frequency in both tables, open responses often specified the frequencies as continuously, daily, annually or on request. That this category received about half of all the responses in each of the tables indicates that some of these frequencies should have been offered as response categories.

The tables generally indicate notable variations between countries in the updating frequency of the information in question.

Question SD 3.12: Do you update your sample frame for the following changes and if so from which sources? (Type: Harmonization parameter)

Table 12 - SD 3.12: Whether sample frame updated by SDs for the changes and its sources.

	BR update	Returned survey	Admin data	Other source	Total	Not updated
Births	215	87	60	43	405	130
Deaths	232	166	71	35	504	87
NACE	232	189	42	36	499	82
Employment	160	122	82	36	400	127
Turnover	105	116	75	40	336	158
Structure	195	150	38	35	418	106
Name/address	235	173	55	32	495	80
Other	21	13	4	33	71	59
Total	1395	1016	427	290	3128	829

BR was the most important source of updates (45%) to sampling frames of the SDs; the other important source was the returned surveys (32%), while the admin data (14%) and other sources (9%) played a less prominent role. In many cases the Other Sources category contained responses (repeated by the SDs) of doing censuses (and thus no sampling frame), using a live BR frame (which then is updated by BR), particularities about the information that is updated, or use of a specific admin register (the content of the Administrative Data category). Those responses that contained a source were for instance "*Data Exchange*" and "*Business press*".

An additional number of responses (one fifth of the whole volume) indicated that no update is made of the sampling frame. It is with available data not possible to determine why this was so, but possibly it was for some of the same reasons that appeared in the Other category above (e.g. doing a census and thus no sampling frame, or using a live BR frame which then is updated by the BR itself).

With respect to what sampling frames information did BRs update the SDs with, there was a considerable level of agreement between the data from the BRs (Table 10) and the SDs (column "BR update" in Table 12): ceasing of a unit's existence and change in a unit's NACE code were in both perspectives among the most frequent ones, while turnover was the least provided one.

4.2.3. Provision of changes from SDs to BR

Question BR 2.7: Do you receive feedback from the following surveys on units in your register and how frequently? (Type: Harmonization parameter)

Table 13 - BR 2.7: Feedback received from SDs on units in the BR and its frequency.

	Weekly	Monthly	Quarterly	Annually	Other	Total	Not known	Not Received
SBS - Annual Structural Statistics	3	0	0	18	4	25	0	0
SBS - Industry	3	0	0	19	4	26	0	2
SBS - Distributive trade	3	0	0	19	3	25	0	3
SBS - Construction	3	0	0	19	3	25	0	3
SBS - Business Services	3	0	0	19	4	26	0	2
SBS - Business Demography	1	0	0	5	5	11	0	10
STS - Industry	4	8	3	5	5	25	0	2
STS - Construction	3	7	4	5	4	23	0	4
STS - Retail Trade & Repair	3	9	3	7	3	25	0	3
STS - Other Services	3	8	4	6	3	24	0	3
PRODCOM	3	4	1	7	5	20	0	5
Inward FATS	0	0	1	14	3	18	0	5
Outward FATS	1	0	1	9	3	14	0	8
FDI	0	0	0	5	4	9	1	11
ICT	2	1	1	6	5	15	1	9
R&D	2	0	1	7	5	15	1	9
Job Vacancy Statistics	2	0	1	4	5	12	0	11
Struct. of Earnings/Labour Cost	2	2	1	11	3	19	0	5
Vocational training	1	0	1	4	1	7	1	14
Other	0	0	1	1	0	2	2	2
Total	42	39	23	190	72	366	6	111

Question BR 2.8: What information is fed back to the Business Register from surveys?
(Type: Harmonization parameter)

Table 14 - BR 2.8: Information fed back to the BR from SDs.

	Ceased	Empl. & turnover	Location	NACE	Struct. change	Business name	Other	Total
SBS - Annual Structural Statistics	17	16	18	24	18	12	0	105
SBS - Industry	18	15	17	25	17	13	0	105
SBS - Distributive trade	17	13	16	23	15	12	0	96
SBS - Construction	16	13	16	23	15	12	0	95
SBS - Business Services	18	14	17	25	16	13	0	103
SBS - Business Demography	6	4	5	8	6	4	3	36
STS - Industry	14	12	14	21	14	11	0	86
STS - Construction	13	11	13	20	12	10	0	79
STS - Retail Trade & Repair	14	11	13	21	12	10	0	81
STS - Other Services	13	11	12	19	11	9	0	75
PRODCOM	8	5	9	17	7	7	0	53
Inward FATS	8	2	6	5	9	6	5	41
Outward FATS	7	2	5	4	7	4	4	33
FDI	4	1	2	5	6	4	3	25
ICT	9	2	7	9	3	7	1	38
R&D	8	2	8	9	3	6	1	37
Job Vacancy Statistics	5	2	8	7	3	4	2	31
Struct. of Earnings/Labour Cost	9	8	9	13	8	7	1	55
Vocational training	3	1	4	4	4	3	1	20
Other	1	2	1	2	1	0	0	7
Total	208	147	200	284	187	154	21	1201

Question SD 3.14: For the following list of variables, please indicate if you feed back information to your Business Register on changes and how frequently.
(Type: Harmonization parameter)

Table 15 - SD 3.14: Feedback given to the BR from SDs on changes on specified variables and its frequency[5].

	Immediately	Weekly	Monthly	Quarterly	Annually	Other	Total	Not fed back
Name	108	3	19	10	74	39	214	154
Location	111	2	33	10	78	41	234	131
NACE	122	2	45	10	111	35	290	99
Employment	46	1	39	8	113	32	207	161
Turnover	30	1	19	16	127	19	193	182
Structure	99	3	38	7	54	28	201	151
Ceased	107	2	38	9	66	35	222	136
Other	7	0	4	1	15	16	27	84
Total	630	14	235	71	638	245	1588	1098

Turning to feeding back information from SDs to the BR, a considerable proportion of the responses - about a third - indicated not feeding back information to the BR. However, this tendency was more expressed in the SD responses (1098/(1098+1588)=41%) than in the BR responses (111/(372+111)=23%), possibly indicating that SDs have a better insight into occurrences of information that could have been fed back to the BR but actually was not fed back. Among the SDs feeding back information to a relatively large extent were the SBS (except for Business Demography) and STS, whereas Business Demography and Vocational training survey were among the SDs providing feedback to the BR to a relatively low extent.

[5] For this table, a prior data editing step was performed by which, if a respondent provided more than one frequency category (e.g. that the feedback was provided both “weekly” and “monthly”), only the highest-frequency response was taken into account as all the lower-frequency responses are a direct consequence. Also, responding with “other” frequency” led to taking only this response category into account.

Annually and *Immediately* were the two dominant responses on frequency of updates (in the BR responses, Table 15, the *Immediately* responses are to be found in the Other category), however the intermediate frequencies like *Weekly*, *Monthly* and *Quarterly* were present to a considerably degree - between a quarter and a third of the responses - indicating that there is variability in SDs' practices of giving updates to the BR.

Similar to the results above (Table 15), the information often reported fed back was the NACE code change, whereas changes in turnover/employment among those least reported fed back.

Question SD 3.15: If you do not feed back information on changes to the Business Register please briefly explain why not. (Type: Quality indicators)

The open responses, which were relatively few, mostly belonged to the category where feedback still was given to the BR, but that it occurred with a non-constant, varying frequency (i.e. the updating was *ad hoc*).

4.2.4. Updating of the BR and resolution of conflicts

Question BR 2.9: Is the information received from surveys updated immediately or periodically? (Type: Harmonization parameter)

Table 16 - BR 2.9: Frequency of updates based on SD feedback.

	Immediately	Daily	Weekly	Monthly	Quarterly	Annually	Other	Total
SBS - Annual Structural Statistics	9	1	0	1	1	12	2	26
SBS - Industry	9	1	0	1	1	12	2	26
SBS - Distributive trade	8	1	0	1	1	12	2	25
SBS - Construction	8	1	0	1	1	11	2	24
SBS - Business Services	9	1	0	1	1	11	2	25
SBS - Business Demography	5	0	0	1	0	1	3	10
STS - Industry	9	1	0	5	3	5	2	25
STS - Construction	8	1	0	5	3	5	2	24
STS - Retail Trade & Repair	8	1	0	5	3	5	2	24
STS - Other Services	8	1	0	5	3	5	2	24
PRODCOM	7	1	0	3	2	6	3	22
Inward FATS	6	1	0	0	0	6	3	16
Outward FATS	6	0	0	0	0	4	4	14
FDI	5	0	0	0	1	3	5	14
ICT	6	0	0	0	0	4	2	12
R&D	5	0	0	0	0	5	2	12
Job Vacancy Statistics	4	0	0	1	0	2	3	10
Struct. of Earnings/Labour Cost	6	1	0	1	0	7	1	16
Vocational training	3	0	0	1	0	1	3	8
Other	0	1	0	0	0	1	0	2
Total	129	13	0	32	20	118	47	359

Updating of the BR *immediately* with information provided by SDs was reported most commonly, however, that response category was given for just about half of the BR-SD pairings. The other prominent frequency was *annually*, reflecting either that a process of updating the BR takes place annually, or that a majority of the SDs in the questionnaire study were annual surveys, which then as a rule provide their input annually. The intermediate frequencies were far less prominent.

Question BR 2.10: Do you allow any of the following survey areas to directly update the business register? (Type: Harmonization parameter)

Table 17 - BR 2.10: Permission for direct update of the BR by SDs.

	Allow	Do not allow	Total
SBS - Annual Structural Statistics	9	19	28
SBS - Industry	8	20	28
SBS - Distributive trade	7	21	28
SBS - Construction	8	20	28
SBS - Business Services	8	20	28
SBS - Business Demography	6	16	22
STS - Industry	6	20	26
STS - Construction	5	21	26
STS - Retail Trade & Repair	5	21	26
STS - Other Services	5	21	26
PRODCOM	8	18	26
Inward FATS	5	20	25
Outward FATS	2	21	23
FDI	2	21	23
ICT	3	21	24
R&D	2	22	24
Job Vacancy Statistics	1	23	24
Struct. of Earnings/Labour Cost	3	21	24
Vocational training	1	23	24
Other	4	9	13
Total	98	398	496

Question SD 3.13: Are you allowed to directly update the Business Register for the following changes? (Type: Harmonization parameter)

Table 18 - SD 3.13: Allowed to directly update the Business Register for the following changes?

	Name	Location	NACE	Empl/Turn	Structure	Ceasing	Other	Total	Not allowed
SBS - Annual Structural Statistics	4	7	9	6	6	5	3	40	101
SBS - Industry	4	7	9	5	4	4	2	35	113
SBS - Distributive trade	4	6	8	5	5	4	3	35	100
SBS - Construction	4	6	8	4	4	4	2	32	111
SBS - Business Demography	7	8	7	6	7	8	3	46	89
STS - Industry	2	5	5	3	4	4	1	24	144
STS - Construction	3	6	6	4	5	4	1	29	147
STS - Retail Trade & Repair	5	7	8	5	5	5	2	37	126
STS - Other Services	4	6	7	5	4	4	2	32	140
PRODCOM	2	4	7	1	3	3	1	21	108
Inward FATS	2	3	3	2	2	1	0	13	108
Outward FATS	2	2	2	2	3	2	1	14	106
FDI	2	2	3	2	2	3	0	14	101
ICT	0	0	1	0	0	0	2	3	150
R&D	3	4	2	3	3	3	3	21	120
Innovation Survey	3	4	2	3	2	3	3	20	140
Job Vacancy Statistics	1	2	0	0	0	0	2	5	109
Labour Cost Survey	2	3	1	2	0	2	2	12	157
Structure of Earnings Statistics	0	1	0	0	0	0	2	3	150
Total	54	83	88	58	59	59	35	436	2320

By both kinds of sources (the BRs and the SDs), direct update of the BR is predominantly not implemented, as reported for approximately 80% of both the *sd × country* combinations in the BR questionnaire data and the *sd × country × change type* combinations in the SD questionnaire data (Table 17 and Table 18, respectively). While it is beyond the available data to say something about the reasons for this, it likely has to do with ensuring a certain level of quality of the data prior to its inclusion into the BR.

Question BR 2.11: How do you resolve conflicts of information between survey area feedback and administrative sources? (Type: Harmonization parameter)

To facilitate an overview of the responses to this open-ended question, a thematic analysis was carried out, resulting in three main categories: existence of a resolving process, priority given to a specific source (the survey or the administrative data), and use of a rule for deciding which of the two sources to give priority to.

Responses mainly stating use of a resolving process: eleven responses were interpreted as providing this content. Some examples follow.

> *The difference is solved by statisticians in a special tool - Coordinating Database and BR team gets the valid result to put it into the BR.*
>
> *Further investigation from other sources (i.e. internet) depending on the variable*
>
> *… It is also possible that (structural) conflicts are treated by chain-management. This is a group of specialist which guard all statistical processes between the input of the BR until the output of National Accounts.*
>
> *We check all available information and in very few cases we recontact the enterprise to understand and complete the information.*

Priority of a source: in this category of responses, eight were interpreted as giving priority to survey data and three as giving priority to administrative sources.

> *In the BR the information from statistical surveys has a priority before administrative sources.*
>
> *Feedback from surveys has priority over data from administrative sources. Exceptions are possible, if the information from an administrative source is more up-to-date and considered to be reliable.*
>
> *Survey area feedback is used only if administrative sources are not available (e.g. NACE, structure of enterprise,…).*

Specific rules: five responses were interpreted as having that content.

> *Information on location received by survey will usually be updated. Information on legal form and business name will not be updated in any case [*if it comes from a survey(?) – added comment*].*
>
> *The BR staff analyse and solve them if they are related to big enterprises. For small units the most recent information is the correct one.*
>
> *The priority rules are established for sources with the same date.*

While the practices vary, the content of the responses indicates a strong dedication to ensuring as high a quality of data in the BR as possible.

Question BR 2.12: Do you give priority to a specific survey / domain when resolving conflicts in NACE classification? (Type: Harmonization parameter)

Question BR 2.13: Do you give priority to a specific survey / domain when resolving conflicts in employment values? (Type: Harmonization parameter)

Question BR 2.14: Do you give priority to a specific survey / domain when resolving conflicts in turnover values? (Type: Harmonization parameter)

Table 19 – Priority given to a specific survey when resolving conflicts.

Conflicts in	Yes	No	Total
NACE	18	10	28
Employment	16	12	28
Turnover	16	12	28

To at least one of the three questions BR 2.12-2.14, a response was provided by 24 participating countries. The responses indicate that if priority is given to a specific SD when resolving specific conflicts in data, it is mostly SBS or STS. This was more prominent in the case of conflicts in NACE, and somewhat less prominent in the case of conflicts in turnover and employment, where reference to administrative sources was instead reported more often.

> *SBS Annual structural survey [for NACE and turnover; for employment: statistics on earnings and labour costs survey] (data from annual balance sheets has priority before all other surveys)*
>
> *The declaration of the activity from the enterprise is the starting point for the coding. The data of social security are the top of the priority [for employment].*

In a number of responses, indication was given of more frequent surveys having priority over infrequent ones, as well as specific surveys having priority over general economic surveys.

> *Surveys, which cover a certain economic sector are given priority over surveys, which cover the economy as a whole; Monthly surveys are given priority over quarterly and annual surveys.*
>
> *Usually the STS data has priority over annual data.*
>
> *Monthly survey data has a priority.*

Sources of information for updating the BR were reflected in a number of responses.

> *We use two administrative sources that give the information on turnover by economic activity. Information on employment is only taken by a Linked Employer-Employee Social Security Register. Turnover is updated only by admin sources.*
>
> *Employment values we calculate from Social Security source data. Annual turnover is from administrative sources data: Published business accounts, Personal income tax, Corporate Tax, VAT.*
>
> *The SBS normally is the only one that have detailed information on the value added to be able to resolve conflicting NACE classification. [For turnover:] Corporate tax data is given the first priority, followed by SBS. The remaining once will have the VAT estimated figure.*

In summary, the responses reflect use of various means and priorities for resolving conflicts in information sources feeding into the BR.

Question BR 2.15: How are changes in units' NACE codes dealt with on your register?, with response categories ***Updated at a fixed point*****,** ***Updated immediately*****,** ***Held as a separate variable*****, and** ***Other*** (Type: Harmonization parameter)

Multiple responses were allowed, generating 35 responses. *Immediate* updates were reported somewhat more frequently than those carried out at *fixed points* in time. If choosing both *fixed point* and *immediately* – which three countries did – this was accounted for by importance of the update or by dependence on the source of data:

> *Immediate updates concern important cases only.*
>
> *Fixed point updating is done based on administrative data transfers.*

If time reference of the NACE codes is saved (which six responses indicated doing), this tends to be more associated with immediate updates of the BR than with updates of the BR at fixed times.

The responses also reflected management of the "parallel" registers, the live one and their frozen extracts. There were also clear indications of the difference of treatment, in this respect, between the administrative use and the statistical/sampling use of NACE codes. This use likely is correlated – however, without sufficient information in this study to make firmer judgements about that – with the existence of live BR (used for keeping track of changes in the target population) versus the frozen BR (used for sampling and estimation purposes).

> *All the changes are recorded in the BR with the reference date and the validity date. The BR database contains all the historical data.*
>
> *The SBR is a live register with information about dates for changes in reality. The frozen versions are drawn at given points in time, showing the information as it is in reality ultimo the reference period at the point in time when the frozen version is drawn.*
>
> *Nace codes in frozen frames remain unchanged except in certain cases (if unit's number of employees is 50 or more, or if unit is important in its domain). In the 'live' register, the Nace can be changed at any time.*
>
> *NACE codes - "statistical"- are changed annually. Administrative NACE code updated immediately from administrative sources.*
>
> *NACE codes are updated immediately, but for samples once per year NACE codes are held as a separate variable "frozen" NACE*
>
> *Annual updates of NACE and in case of organization changes of units immediately updates of NACE.*

4.3. Temporal aspects

This section evaluates results on questions on the major theme of temporal aspects. It encompasses issues like timing and updating of data in registers, frames and samples.

4.3.1. Time lags and updates in the BR

This topic is among those central to the risk for inconsistent practices resulting in deficiencies in comparability and coherence. For instance, with a strong dynamic of change in company structure and other properties or measures of companies' performance, any difference between the reference periods that the used sampling frames have may strongly impact comparability of statistics produced on the basis of these different frames.

The questions in Business Registers (BR) and specific Statistical Domains (SD) questionnaires relating to temporal aspects are specified in Table 1. The evaluation is presented first for the questions separately, followed by a summary of the findings.

Question BR 1.6: Has an estimate been made for the lag between the administrative sources recording the event and the event being recorded on the business register. If so what is the average lag in months? with ticking yes on a period offered for specified events ***Business start-ups, Business closures, Changes in employment/ turnover, Changes in activity (NACE), Central Bank Data, Changes in structure*****, and** ***Other - please specify*** (Type: Harmonisation Parameter)

Table 20 - BR 1.6: Estimate of lag between admin source and BR recording of an event.

	< 1 month	1-3 months	4-6 months	7-12 months	> 12 months	Total	Not known
Start-ups	11	9	2	3	1	26	1
Closures	10	7	2	4	2	25	2
Empl./turnover ch.	2	7	3	7	6	25	2
NACE change	7	8	2	6	3	26	1
Structure change	7	10	1	3	2	23	3
Other	3	0	0	1	0	4	5
	40	41	10	24	14	129	14

The response distribution in Table 20 shows that of the countries able to provide an estimate of the lag, about two-thirds of the responses – over all the categories of events – indicate a lag of 3 months or less between administrative source recordings and BR recordings. Reported time lag of recording *start-ups*, *closures* and *structure* changes was lower than that of *NACE* changes, which in turn was lower than that of *employment and turnover* changes.

However it is also seen from some of the respondent comments that the mentioned lag is just indirectly related to the comparability of the statistics, as the lag between the actual event and the BR recording may be larger than the recording lag reported here. This may reduce the benefit to comparability of a possible requirement on the BR recording lag specifically.

> *Changes in employment or turnover are directly transmitted into [the] BR; changes in structure are mostly provided by [the] Economic Chamber, from which data is received in a weekly basis, register maintenance however is 4-6 month delayed.*
>
> *The record of changes in structure is quite good but not regular. These changes in employment or Nace code need at least 12 month to be integrated.*
>
> *The time-lag for Business closures depends on witch administrative source the business reports to. If they pay VAT and have employees (or are in a survey) we can detect the closure earlier than registered in the matrix.*
>
> *Changes in employment / turnover are difficult to specify because of the variety of frequencies of how employment and turnover is updated in BR. Large enterprises are maintained on daily basis as a result of delivery from administrative source, while smaller enterprises are more dependent on administrative sources for changes.*

In other words, when responding to the question, it seems that perhaps some respondents have taken into account a wider chain (e.g. Event -> Admin register -> BR) rather than the intended one (Admin register -> BR). Employment/turnover changes are generally visible in annual business reports, but - when/if they are put in an admin register - there is little reason that these specific pieces of information are less speedily transferred to a BR than for instance structure change or NACE change.

Question BR 1.7: For new units how frequently do you receive and update ...from administrative data (Type: Harmonisation Parameter)

Table 21 - BR 1.7: Frequency of update for new units.

	Daily	Monthly	Quarterly	Annually	Other	Total	Not used
VAT	2	11	2	9	3	27	1
Personal Tax	0	2	2	13	1	18	10
Corporate Tax	0	1	2	10	2	15	14
Business Accounts	0	1	2	14	1	18	12
Social Security	0	9	6	6	3	24	8
Central bank	0	3	1	10	0	14	13
Other	6	5	0	3	9	23	5
Total	8	32	15	65	19	139	63

As expected, some sources are by nature limited to annual data, such as taxation and business accounts. However for VAT data the updating frequency differs strongly between countries, with both weekly and annual updating being common. Other frequencies were also mentioned, amongst them *weekly* (three participating countries) and *biannually*.

> *Monthly for big enterprises and quarterly for the rest.*
>
> *Economic Chamber data is received and updated in a weekly to monthly basis. In "Other" the ... Company Register is referred to, which is the most important source.*
>
> *For structural information the information based on social security data is almost daily available. For statistical data only once a year.*
>
> *VAT: information from tax authorities on turnover is received daily, but the information about the administrative unit is received weekly.*

Question BR 1.8: For existing units how frequently do you receive and update ...from administrative data (Type: Harmonisation Parameter)

Table 22 - BR 1.8: Frequency of update for existing units.

	Daily	Monthly	Quarterly	Annually	Other	Total	Not used
VAT	1	11	2	10	4	28	1
Personal Tax	0	2	2	13	1	18	10
Corporate Tax	0	1	2	10	2	15	14
Business Accounts	0	1	2	15	1	19	12
Social Security	0	9	3	6	2	20	8
Central bank	0	3	1	11	0	15	13
Other	7	5	0	4	10	26	3
Total	8	32	12	69	20	141	61

Distribution of the responses, in Table 22, was for this question similar to that for the preceding question (on new units), enabling similar conclusions. Namely that some sources are by nature limited to annual data, such as taxation and business accounts, and that for VAT data the updating frequency differs strongly between countries, with both weekly and annual updating being common.

> *Information related to existing units is available from legal administrative sources (commercial register, trade register,...) on a daily basis.*

4.3.2. BR as a resource for sampling

Question SD 3.3: When do you take your main sample? (Type: Harmonisation Parameter)

Table 23 - SD 3.3: Month of taking the main sample (only responses that contained a specific time).

Month:	01	02	03	04	05	06	07	08	09	10	11	12	Total
SBS - Annual Structural Statistics	2	1	1	0	1	1	0	1	0	0	7	4	18
SBS - Industry	3	2	2	0	0	0	1	0	0	1	5	4	18
SBS - Distributive trade	3	1	1	0	0	1	0	0	0	0	6	4	16
SBS - Construction	2	2	1	0	0	0	2	0	0	1	5	4	17
SBS - Business Demography	1	0	1	1	1	1	0	0	0	0	1	1	7
STS - Industry	3	1	0	1	0	0	1	0	2	1	2	6	17
STS - Construction	2	0	1	0	0	1	0	0	1	1	4	5	15
STS - Retail Trade & Repair	2	1	2	0	0	0	1	0	0	2	4	6	18
STS - Other Services	4	1	2	0	0	0	1	0	0	1	2	5	16
PRODCOM	7	2	0	0	0	0	1	1	1	0	4	4	20
Inward FATS	3	0	1	0	1	0	0	0	0	0	3	3	11
Outward FATS	1	0	2	1	0	0	1	1	1	1	0	1	9
FDI	1	0	3	2	0	0	1	0	1	1	1	0	10
ICT	5	2	4	0	0	0	0	0	0	1	3	3	18
R&D	2	1	1	1	0	0	0	0	2	0	3	2	12
Innovation Survey	2	1	3	3	1	0	1	1	1	1	3	3	20
Job Vacancy Statistics	2	2	0	1	0	0	0	0	0	1	3	1	10
Labour Cost Survey	2	1	3	0	0	0	0	0	1	1	4	1	13
Structure of Earnings Statistics	2	3	1	1	0	1	0	0	3	1	3	1	16
Total	49	21	29	11	4	5	10	4	13	14	63	58	281

Table 23, and also respondent comments, generally indicate notable variations between countries. Cf. further the following Question (SD 3.4). However, months at the end or beginning of a calendar year (November – January) were contained in 60% of these responses (170 of 281 responses).

> *Initial sample in December R-1, then monthly updates.* [STS]

> *We select our sample in July because our Business Register is updated in this date. The units selected in sample begin to give data in September, but we don't use it for calculating our index until January. In January we updated our weights.* [STS C, D]

> *At the end of 2011, but for Production index we do a continuous updating of panel. When a unit d[ies] we try to choose another similar unit and, in some cases, we incorporate new units.* [STS A]

The annual update is made during February, because we have to wait on some administrative information coming during January (it is a synthesis of some information collected during the year before). [R&D]

There is not a specific rule. Normally is taking from the NSI Statistical Business Register updated in August/September. [PRODCOM]

The panel of enterprises is set up at the beginning of the base reference year and then is updated as soon as the latest PRODCOM is available. [STS A]

In addition to the 281 responses that could be assigned to some month according to the methodology outlined in Section 3.3, there were more than one hundred (113) responses that could not be assigned to a month as they referred to no specific point in time, referred to creating the frame rather than drawing the sample, or a number of other reasons. They still provide, however, an insight into variety of procedures and reasoning about them regarding timing of drawing samples by the SDs.

The sample frame is continuously renovated. Major revisions are made with the availability of detailed censal data; the sample units are renovated by sixths every quarter and the censal areas of the new dwellings entering the first time in the sample are refreshed and updated in order to cover the new buildings, change of use of premises, etc. [STS B]

The DIRID is updated with information of the last survey when finished. Besides, usually at the end of January we get the provisional files updating units which received public support or grants for R&D activities. Last week of March is the deadline for receiving the final files. With this information, the DIRID is brought up to date. [R&D]

Frozen frame is created annually by 1st November, but updates to it (in economically significant cases) are made each month. For BD sample done usually T+15 after the frozen frame is compiled. [SBS IX]

Frame is updated every month. Corresponds with monthly statistics. For quarterly statistics the frame of the 3rd month is taken. [STS]

Question SD 3.4: Why do you take your sample at this time? (Type: Quality Indicator)

Table 24 - SD 3.4: Reasons for taking the sample at the time it is taken.

	Timeliness	Coverage	Coordination	Technical	Other	Total
SBS - Annual Structural Statistics	9	7	10	5	1	32
SBS - Industry	5	6	11	6	1	29
SBS - Distributive trade	7	6	10	6	1	30
SBS - Construction	6	7	10	5	1	29
SBS - Business Demography	2	5	2	3	0	12
STS - Industry	10	6	5	7	8	36
STS - Construction	13	5	8	10	3	39
STS - Retail Trade & Repair	13	6	10	10	4	43
STS - Other Services	9	8	8	11	2	38
PRODCOM	15	7	6	2	1	31
Inward FATS	2	0	9	4	2	17
Outward FATS	4	5	5	3	4	21
FDI	9	5	7	2	1	24
ICT	15	6	9	7	2	39
R&D	8	7	7	6	2	30
Innovation Survey	11	4	9	9	2	35
Job Vacancy Statistics	10	4	5	3	3	25
Labour Cost Survey	12	10	8	4	3	37
Structure of Earnings Statistics	13	10	6	2	3	34
Total	173	114	145	105	44	581

Table 24 shows that all the provided response categories were used to a considerable extent by the respondents: *timeliness*, *coverage*, *coordination* and *technical reasons*. The table may have to be read with some reservation, since the distinction between the response alternatives may not be entirely

clear, as seems somewhat indicated by respondent comments. For instance, there are specific points in time when a sampling frame is updated. At that moment, the frame is at the peak of its coverage, as the target population will change as the time passes. Thus, both timeliness (as soon as current version of the frame has been produced) and coverage (a 'fresh' frame has relatively the best coverage), and perhaps also technical reasons might be invoked when providing reasons for the decision when to take a sample from the frame. A selection of the participating countries' responses might illustrate this.

On timeliness reasons:

> *The time of drawing the sample is within the main reference period of the survey. The sample should be as current as possible.* [ICT]
>
> *The time for drawing the sample is chosen for reasons of timeliness of BR and organisation of data editing process.* [STS C, D]
>
> *Latest data available for sampling.* [PRODCOM]
>
> *VAT-Group estimation was finished in August.* [STS D]

On coverage reasons:

> *We wait to have the most complete information for the year concerned.* [SBS]
>
> *In that date is available the final frame for the previous year.* [PRODCOM]
>
> *Just after register has been updated in September.* [SBS]

On technical reasons:

> *Survey starts a few weeks later.* [Innovation stats, R&D]
>
> *The programs for sample rotation had to be completed.* [SBS III, STS C]
>
> *To make sure that the target population enterprises have time to implement the correct standard method for the specific year in their wage-systems.* [Labour Cost Survey, Structure of Earnings Survey]
>
> *This enables us to have up-to-date data and gives us enough time to finalize the preparation of the letters ...* [ICT, Innovation stats]
>
> *Business register software constraints mean that selection for multiple surveys need to be timetabled.* [Job vacancy stats]

The respondent comments also generally indicate that the reasons for choosing the sampling time are often related to specific conditions in countries. However, some slight tendencies in the response pattern in the table can be noted. For instance, timeliness reasons are on the whole relatively frequent for STS and coordination reasons for SBS, which seems natural. The balance between timeliness reasons and coordination reasons may be seen as indicative of a trade-off between the objectives of timeliness and horizontal consistency, respectively.

Question SD 3.9: Is the data ... from your Business Register of sufficient timeliness ... (Type: Compliance Criterion)

Table 25 - SD 3.9: Perception of timeliness of BR data.

	Yes	No	Total
SBS - Annual Structural Statistics	22	1	23
SBS - Industry	24	0	24
SBS - Distributive trade	19	1	20
SBS - Construction	22	0	22
SBS - Business Demography	21	4	25
STS - Industry	24	2	26
STS - Construction	24	1	25
STS - Retail Trade & Repair	25	1	26
STS - Other Services	23	4	27
PRODCOM	18	1	19
Inward FATS	21	0	21
Outward FATS	11	8	19
FDI	15	2	17
ICT	21	2	23
R&D	20	2	22
Innovation Survey	23	1	24
Job Vacancy Statistics	11	6	17
Labour Cost Survey	20	5	25
Structure of Earnings Statistics	19	5	24
Total	383	46	428

Table 25 shows that the respondents mostly consider the BR to be sufficiently timely for use in the statistical domains – over 90% of the responses indicating the case being such. However, in some statistical domains, in the first place Job Vacancy Statistics and Outward FATS, there is a notable minority that does not consider the BR to be sufficiently timely.

> *We have a lot of enterprises, especially in NACE 46.1, we do not reach and the mail returns, because the address is unknown.* [SBS III, STS D]
>
> *There is an almost 2 year gap between the time reference in the frame and the year we are surveying for. For a small country and particularly in recent years a 2 year gap is too wide as there may have been significant changes in important enterprises.* [STS C]
>
> *The frame is updated for t-1 year.* [PRODCOM]
>
> *Time gap between the R&D survey reference year and the Business Register reference year is usually 1 year (Last edition R&D survey year 2011, last Business Register year 2010).* [R&D]

Question BR 2.3: How frequently are frozen extracts created from your Business Register? (Type: Harmonisation Parameter)

Table 26 - BR 2.3: Frequency of creating frozen extracts of the BR. (The sum exceeds the number of responding BRs as it was allowed to provide response in more than one frequency category.)

	Annually	Quarterly	Monthly	Other	Total	Not created
Number of responses	20	2	9	7	38	0

Table 26 shows that frozen extracts are mostly created annually but that creating monthly extracts is also common. From the respondent comments it appears that different frequencies of frozen extracts may be used for different statistical domains.

> *Frozen frame is created annually by 1st November, but updates to it (in exceptional, economically significant cases) are made each month.*

... three versions for reference t are generated in different periods of time (t +6 m, t+8 m, t +10 m).

BR makes one frozen extract per annum. Other: SBS take a copy of BR's live frame twice per annum.

Frozen extracts from Business Register are created in time for each domain using Business Register.

Frozen extracts ... are created twice per year annually on 10th of October and on 1st June. The last one updates economic activity and status of activity.

In principle the choice of frequency of taking frozen extracts involves a trade-off between horizontal consistency and timeliness. Namely, annual "freezing" favours horizontal consistency, as different surveys can pertain to the same unchanged business population. On the other hand, more frequent freezing favours timeliness in some respects. The impact on timeliness in statistical results is only partial though, as more up-to-date information is usually collected in the surveys to complement the older information from the frame.

The results thus seem to indicate that there is some inconsistency between countries in the trade-off between horizontal consistency and potential for timeliness. This may concern the entire system of business-related statistics, from both users' and producers' perspectives.

Question BR 2.4: For those surveys using a frozen extract from your Business Register, how current is the extract? (Type: Harmonisation Parameter)

Table 27 - BR 2.4: Recency of the frozen BR extract.

	Within month	Within quarter	Within year	Other	Total	Not known	Not used
SBS - Annual Structural Statistics	8	1	12	1	22	0	1
SBS - Industry	6	1	13	1	21	0	3
SBS - Distributive trade	6	1	13	1	21	1	3
SBS - Construction	6	1	13	1	21	1	3
SBS - Business Services	6	1	12	1	20	1	3
SBS - Business Demography	5	0	9	2	16	0	5
STS - Industry	7	0	9	0	16	1	3
STS - Construction	7	0	8	0	15	2	4
STS - Retail Trade & Repair	8	1	9	0	18	1	3
STS - Other Services	7	0	9	0	16	2	4
PRODCOM	6	1	6	1	14	2	4
Inward FATS	1	1	11	3	16	4	3
Outward FATS	1	0	8	2	11	4	3
FDI	1	0	5	2	8	5	5
ICT	2	0	13	1	16	2	4
R&D	2	0	10	1	13	3	5
Job Vacancy Statistics	5	1	6	1	13	2	3
Struct. of Earnings/Labour Cost	3	2	14	1	20	1	2
Vocational training	3	0	9	0	12	3	3
Other	0	0	2	0	2	0	2
Total	90	11	191	19	311	35	66

Corroborating findings reported for the preceding question, Table 27 shows that the frozen extract is mostly not more recent than from within the year. However for SBS in a notable minority of countries frozen extracts are from within a month or quarter.

Frequency of frozen extracts (question BR 2.3) affects the possibility for the extract used to be recent. Nevertheless, the results for the two questions combined indicate that some countries taking annual extracts may do this at such a time that the extract is recent when used in annual surveys like SBS.

For the same reasons as those explained for the preceding question (BR 2.3), how current the extract needs to be may be seen as a trade-off between horizontal consistency and potential for timeliness.

The results thus seem to indicate that there is some inconsistency between countries in the trade-off between horizontal consistency and potential for timeliness. This may concern the entire system of business-related statistics, and also specifically particular statistical domains, from both users' and producers' perspectives.

> *We understood this question as follows: frozen extract is taken within the last year from the survey point of view.*
>
> *Not all the drawings from the SBR are registered. A system has been produced so that the statistical division can draw information from the frozen versions themselves.*

4.4. BR maintenance

This major thematic area contained questions on maintenance and procedural aspects of keeping the Business Register (BR).

Question BR 1.9: What information is available to allow NACE coding of units on your business register? (Type: Harmonization parameter)

In two of the countries, the BR reported use of seven sources of information to allow NACE coding, and in as many countries (two) the BR reported use of one source: in one case NACE codes from admin data, and in the other case a survey. Most often, four or five sources of information were reported by the BR.

Of the offered response categories, most frequently chosen were the administrative sources containing NACE coding of the units and survey returned data (with three-quarters or more of the BRs providing each of these responses), followed by internet searches and business descriptions (between two-thirds and a half of the BRs). Open responses clarify that a procedure is in several participating countries put in place whereby units classify themselves into a certain NACE code, which is then fed into the BR by the administrative source. In some countries this is verified by a special survey to the units, whereas elsewhere a coding is independently done by the BR using available sources like

> *Officially published documents and information on Internet* [A small country]

When this question was understood as referring to existing units, update from surveys (statistical domains) was mentioned as well as other *daily maintenance work*.

Question BR 1.10: How is employment derived on your register for new units? (Type: Harmonization parameter)

BR in four of the countries reported use of three of the offered sources of information for deriving employment, whereas two BRs did not choose any of the offered sources. Most frequently, one or two sources of information were reported. These were the response categories "administrative source" (chosen by three-quarters of the BRs) and "survey source" (chosen by half of the BRs).

Open responses indicate that in some cases, for this and the next two questions, the responding BRs did not realise that the questions were on new units, thus reporting their usual (for businesses already established) procedures of keeping information on these variables up to date in the BR. However, in some other cases, it might be that procedures for obtaining these data for new and for already existing units are equivalent.

However, the BR in one participating country reports of a special survey:

> *A survey which covers the new entries is carried on 3 month after they enter the BR. After 12 month the administrative sources provide the information.*

which apparently fills the time lag existing between the event of creation of a new business and the data on its size and activity being reflected on administrative sources. Another approach of filling this gap is through business's own estimation:

> *Every new unit during their legal registering procedures has to fill a survey about the number of planned employment, turnover and investments.*

In some further cases, estimation was reported needed:

Estimation based on Periodic Tax Return data.

Further, some of the responses indicate – by appearing in exactly the same form as responses to both questions – that this question and the next one were understood as if referring to the same thing.

Question BR 1.11: How are employees derived on your register for new units?
(Type: Harmonization parameter)

BRs in three of the countries reported use of three of the offered sources of information for deriving employees for new units; no BRs omitted to choose any of the offered sources. Most frequently, one source of information was reported, and that was an "administrative source" – almost all of the BRs (26 of 29) used this source. A "survey source" was used by approximately half of the BRs, while the two other offered sources ("imputation", "other") were infrequently mentioned.

As mentioned, most of the open responses are repetitions of open responses to the previous question, BR 1.10. Additionally, some of them might apply to the process of updating the data for existing units, rather than referring to new units.

Question BR 1.12: How is turnover derived on your register for new units?
(Type: Harmonization parameter)

BRs in two of the countries reported use of three of the offered sources of information for deriving turnover for new units, while another two BRs omitted to choose any of the offered sources. Almost all of the responding BRs reported use of an "administrative source", whereas reports of using a "survey source" was given by approximately half of the BRs; "imputation" was to a limited extent reported as used, and practically no "other" sources.

Some of the open responses are repetitions of responses to the previous question, BR 1.10. Those that are not, imply estimation based on some model, however again there is little indication that this is valid/tenable (due to lack of data) for new units.

Question BR 1.13: Do you record institutional sector codes for units on your BR?
(Type: Harmonization parameter)

Question BR 1.14: How do you calculate institutional sector codes for units on your BR?
(Type: Harmonization parameter)

Most of the BRs reported recording institutional sector codes on the BR – only four of the 29 BRs didn't report doing so. (One of these still provided a response on the next question, Q BR 1.14.) Among the offered ways of calculating institutional sector codes, the one mostly used (by 19 of the 25 BRs that record the codes) was the rule-based method. Which specific rules and which input data are used varied considerable between the BRs, however at least the NACE code and the legal form were repeatedly mentioned. The process was in some cases automated, but often with an added possibility of manual coding.

Basic characteristics involved in it are: NACE Rev.2 activity code, type of ownership and ownership structure, and source of financing

The Institutional sector (IS) is based on a combination of NACE code and Legal form. A table prepared by the NA is used in the BR to automatically allocate the IS

Automatically classified on the basis of their legal form, NACE code, market activity (turnover) and information about ownership/control.

Specific rules have been encoded in the script which automatically calculates the correct institutional sector code, taking into account NACE, legal form, form of ownership, number of persons employed (for sole proprietors). In certain exceptional cases coded manually, using other sources.

Variables ownership, legal form and industry code are used

Calculated from NACE code, legal form and employees/no employees

Inter-institutional working group is created to specify codes of unusual cases

Responses in the "Other" category included those indicating that someone else (central bank, national accounts) derives the codes, or that there is a manual process in place which apparently cannot be called rule-based. Surveys and administrative sources (like those on the financial sector) are additional ways reported to provide the information for coding.

Question BR 1.15: Do you record inactive units on your statistical business register? (Type: Harmonization parameter)

Question BR 1.16: How do you identify inactive units on your statistical business register? (Type: Harmonization parameter)

As all the BRs later provided a response on how they identify inactive units (Q BR 1.16), it is safe to say that all the BRs do record inactive units. Regarding the ways of doing that, as per the responses provided, amongst the offered response categories the most frequently chosen ones were rules on administrative variables, and admin variables in other contexts.
Employment and turnover were the two variables relied on when identifying inactive units on the BR:

Employment = 0 or turnover = 0

Employment and turnover = zero for at least two years

When there are no more employees, or turnover is less than 10.000€ p.a.

A threshold for real active enterprises which means that the activity should be corresponding to ½ a fulltime worker.

Monthly turnover and number of employees=0, except for non-profit institutions, foundations, sole proprietors

It is interesting to note the additional variations to the rule reported in the responses above, for instance the length of period of zero employment and activity of at least two years, a turnover cut-off that is non-zero, or an activity cut-off of a ½ fulltime worker; further, in some cases, like non-profit institutions, foundations, and sole proprietors, the zero employment and activity does not suffice for identifying inactive units.

Direct use of administrative registers was also reported:

If the Legal Unit register itself as ceased.

Status of insolvency and legal cessation date received from Administrative BR.

When a legal unit is not in VAT declaration data base.

Information received from businesses, directly or through surveys, makes a majority of "Other" open responses.

Question BR 1.17: Does your Business Register record a date stamp against the following variables? (Type: Harmonization parameter)

The NACE code variable was reported to be most commonly recorded with some date (only one BR not recording any date), closely followed by the three other variables asked about (Employment, Employees, Turnover, with between three and five of the BRs not recording any date for these variables).

Interestingly, a full record of changes (both date the variable refers to and date of change) was reported as kept by about a quarter of the BRs. In the other cases, only either the date of change or the date the variable refers to was recorded.

Yet another date, referring to a variable, potentially relevant to be kept was reported by a BR:

Reference date of source

which need not be the same as the data that the change was recorded in the BR.

4.5. Reference periods

Question SD 4.1: Do you experience any specific difficulties in meeting the reference period requirements in your country? (Type: Compliance Criterion)

Question SD 4.2: What period do you collect survey data for? (Type: Harmonisation Parameter) (Omitted for STS, Job Vacancy Statistics and LCS.)

Question SD 4.3: Do you adjust for units returning data for different periods [than the nominal reference period]? (Type: Compliance Criterion) (This was question SD 4.2 for STS, Job Vacancy Statistics and LCS, as the preceding question was there omitted.)

The Statistical Domains (SDs) mostly do not report experiencing difficulties in meeting the reference period requirements. Those 13% of the responses that indicated experiencing difficulties covered a number of issues.

Mostly for annual surveys, the issue of an accounting year differing from the calendar year was reported as a prominent issue. In such cases, the responses also mostly indicated that the difference was disregarded, that is, the assumption was made that the submitted data are a sufficient substitute for the corresponding calendar year data.

> *Some firms use financial year which is different from calendar year. In this case financial year is used as estimate for calendar year* [Inward FATS, medium-sized country].

Data on an accounting year that differs from the calendar year are mostly - based on the responses - assigned to that calendar year in which the accounting year ends.

> *In that case, we ask them for the data for the period which finishes in the calendar year of reference* [SBS, large country]

In what proportion of businesses the accounting year is not the calendar year varied between the participating counties from one percent to 25 percent (each of these percentages reported by a separate participating country, one small and one medium-sized).

> *25% of enterprises end their financial year not the end of December, 6% of enterprises have a financial year smaller/larger than 12 months* [SBS, medium-sized country]

However, in some cases the question was also understood to refer to difficulties in collecting data, in turn referring to businesses' burden. In a specific case of the Construction module (Annex B) of STS, the question was generally understood in terms of requiring more frequent data collection than the quarterly pace valid for a number of participating countries, and the respondents' comments reflected the opinion that such increase in frequency would not be appreciated by businesses.

In a specific case, a broader issue was mentioned, indicating a conflict between the quality components of timeliness and consistency.

> *Due to our press release is at mid-June, some enterprises have not closed their account registers* [ICT, large country]

Among the SDs where such a question was meaningful (STS, Job Vacancy Statistics and LCS were omitted), almost three-quarters of the responses indicated the calendar year as the period for which the data were collected. The remaining three response categories offered – *financial year*, *fixed reference date*, and *other* – received 16%, 5% and 7% of the responses, respectively.

Relatively the highest proportion of responses in the *financial year* category were to be found on SBS, while in *fixed reference date* category were to be found on ICT. The latter reflects a special arrangement for reference period(s) in this SD:

Some questions for January, some for previous year, according to model questionnaire of Eurostat [ICT, small country]

The open comments indicate that collecting data for reference periods other than the calendar year may be a compromise that takes what is available, rather than forcing the businesses to change their accounting periods,

We aim to measure calendar year data, but data from respondents with financial year data other than calendar year is accepted [FDI, medium-sized country]

although even drives in the other direction were reported:

Even units with different financial year reports are obliged to submit also annual reports on calendar year [FDI, medium-sized country].

In SES, the few responses that have provided data on the representative month did mostly refer to October, but with April, May and September also mentioned with one instance each.

4.6. Sampling methods and sample coordination

This major theme covers sampling methods used by statistical domains (SDs) and coordination of sampling between the SDs. Another component of sampling practices, namely at what time SDs take their samples and for which reasons, has already been covered in analysing results that pertain to the temporal aspect (Section 4.3, in particular those in Table 23 and Table 24).

It may be remarked that sample coordination more than sampling methods may be considered an essential harmonisation issue. Namely, sample coordination has conceptual consequences regarding the comparability with respect to timing of the actual survey populations. Sampling methods on the other hand are a more technical matter which should usually not have conceptual consequences if operations are carried out competently.

4.6.1. Sampling methods

Question SD 3.5: What method is used to derive the sample from the frame population?
(Type: Quality indicator)

Table 28 – Methods used by the SDs to derive the sample.

		Simple	Systematic	Stratified	Proportional	Panel	Census	Other	Total
1	SBS - Annual Structural Statistics	0	2	15	2	0	7	0	26
2	SBS - Industry	0	2	15	1	0	10	0	28
3	SBS - Distributive trade	0	2	14	1	0	7	0	24
4	SBS - Construction	0	2	15	1	0	9	0	27
5	SBS - Business Demography	0	1	0	0	0	6	1	8
6	STS - Industry	1	4	11	5	4	11	5	41
7	STS - Construction	2	1	14	3	3	7	5	35
8	STS - Retail Trade & Repair	1	1	18	2	2	5	2	31
9	STS - Other Services	2	2	16	4	2	11	1	38
10	PRODCOM	0	0	2	1	1	13	3	20
11	Inward FATS	0	0	4	0	1	10	3	18
12	Outward FATS	0	0	3	0	0	13	4	20
13	FDI	0	0	5	0	0	7	5	17
14	ICT	0	1	20	1	0	3	0	25
15	R&D	0	0	10	1	1	10	4	26
16	Innovation Survey	2	2	19	1	1	8	4	37
17	Job Vacancy Statistics	0	2	16	1	1	6	0	26
18	Labour Cost Survey	1	0	23	2	0	6	1	33
19	Structure of Earnings Statistics	2	0	21	2	0	4	3	32
	Total	11	22	241	28	16	153	41	512

Of the SDs that do sampling (i.e. that do not conduct censuses), two-thirds reported using *stratified* sampling; the *other* category, and *proportional* and *systematic* sampling, were the majority in the content of the last third of the responses. The *other* category mainly gave further clarification about variables used for stratification or about cut-off values used.

4.6.2. Sampling coordination between statistical domains

Question SD 3.6: Is your sample coordinated with other surveys/domains and if so how?
(Type: Harmonization parameter)

Table 29 - SD 3.6: Is your sample coordinated with other surveys/domains and if so how[6]?

	Same frame	Same time	Overlap	Excluded	Total	Not coordinated	Not known
SBS - Annual Structural Statistics	165	111	31	7	314	71	42
SBS - Industry	167	101	37	17	322	88	25
SBS - Distributive trade	161	110	27	9	307	54	28
SBS - Construction	154	99	30	6	289	78	31
SBS - Business Demography	29	7	7	1	44	91	5
STS - Industry	126	65	20	9	220	185	37
STS - Construction	94	65	9	4	172	168	34
STS - Retail Trade & Repair	105	50	2	5	162	161	37
STS - Other Services	122	61	6	0	189	196	0
PRODCOM	89	42	21	5	157	154	12
Inward FATS	120	75	53	9	257	103	23
Outward FATS	51	14	11	5	81	197	16
FDI	45	25	38	9	117	139	0
ICT	67	38	12	7	124	135	55
R&D	80	61	12	5	158	165	16
Innovation Survey	63	61	20	7	151	153	46
Job Vacancy Statistics	76	35	25	5	141	130	5
Labour Cost Survey	95	38	23	14	170	257	7
Structure of Earnings Statistics	82	31	14	15	142	216	6
Total	1891	1089	398	139	3517	2741	462

This question allowed for multiple responses within each of the *SD* × *SD* × *participating country* cells. While somewhat more than a half of the responses indicated a form of coordination, almost as many responses indicated no coordination (40% of all the responses) as well as some responses expressing no knowledge about coordination (7%).

If there was a coordination of sample selection between SDs, it was mostly to the extent of using the same frame (54% of these responses, or 28% of the total of responses), followed by sampling at the same time (31%, or 16% respectively), in turn followed by positive or negative coordination of inclusion probabilities (15%, or 8% respectively).

Which SDs coordinated their sampling with which other SDs can be seen by building appropriate adjacency matrices. Two examples are given, one for sampling from the same frame (Table 30), and one from sampling at the same time (Table 31). However, some explanation of these matrices first.

The Question SD 3.6, on which the data in the two tables come from, was posed so that each of the SDs could tick which of the other SDs (18 specific SDs were offered, mostly the same SDs that were also respondents to the questionnaire, plus an *other* option) they were coordinating their sampling with. This was offered separately for each of four forms of cooperation: taking samples from the *same frame*, at the *same time*, and positive (*overlap*) or negative (*excluded*) coordination of inclusion probabilities.

[6] Data were edited prior to compiling this Table. The editing, due to potential logical incompatibility of the responses in the case of multiple responses per statistical domain, consisted of the following: occurrence of the response *not known* led to exclusion of all the other responses; thereafter, occurrence of the response *no coordination* led to exclusion of all the other responses. However, further multiple responses for a specific statistical domain were allowed; for instance, *same frame* and *same time* for a single statistical domain were deemed consistent). The *other* response category has predominantly been used by the respondents to clarify or specify already given responses, which is why it is not presented here.

Therefore, in principle, if SD A reports that it coordinates its sampling with SD B by taking samples from the same frame, then it is also the case that SD B should be reporting that it coordinates sampling with SD A by taking samples from the same frame. That is, these are symmetrical relations. In practice, there will be some deviations from a strict symmetry, but in general there will be SDs that coordinate their sampling more with the others and those that coordinate less.

This can be represented in form of matrices, where – as in for instance Table 30 – rows contain SDs that have provided information on their sample coordination, and columns contain SDs that they (the former SDs) coordinate their sampling with. So, for instance, in Table 30, the number 4 on the first row (FDI) and fourth column (IFA) of the body of the table indicates that four FDI surveys in the participating countries have ticked that they coordinate their sampling with IFATS in the same country by using same sampling frames. Correspondingly, the number 5 on the fourth row (IFA) and first column (FDI) of the body of the table indicates that five IFATS surveys have ticked that they coordinate their sampling with FDI by using same sampling frames. These numbers are not always the same as they rely on responses from different persons, but they in general reflect trends in sample coordination between the SDs.

Thus there is a fairly strong tendency of SBS Annexes to coordinate sampling with each other (and especially so of Annex I with the other annexes), but with the exception of Business Demographics (Annex IX); there is also a tendency, but somewhat weaker, of the STS Annexes to coordinate sampling with each other, as well as between SBS and STS.

Further notable coordination seems to exist between Inward FATS and the SBS and STS, as well as between ICT and SBS, between Job Vacancy Statistics and STS, and between R&D and Innovation.

On the opposite side, FDI seems generally to coordinate with few SDs, with an exception of OFATS. Further, when Business Demography (Annex IX of SBS) is asked, they report coordinating with few surveys (all in all, 29 instances were reported). However, when other SDs were asked who they coordinate with, Business Demography is chosen almost to the same level as the other SBS annexes (96 instances).

A note about the matrices: the questionnaire was sent to LCS and SES separately, but the response form contained only one response category for both surveys (“Structure of Earnings / Labour Costs”); therefore, LCS exists in the table as a survey that provided data whereas the column headed “SES” contains in fact responses to the joint response category SES/LCS. Additionally, Innovation Survey was provided as a response category in a fewer number of questionnaire variants than the other SDs, which is why in a number of cases there are empty cells in the matrices.

Table 30 - SD 3.6: A heat map for *Same frame* responses. Red end of the spectrum indicates high frequency of the responses, green end indicates low frequency.

	FDI	ICT	INN	IFA	JVS	OFA	PRO	R&D	SB-I	SB-IX	SB-IV	SB-III	SB-II	SES	ST-B	ST-A	ST-C	ST-D	Total
FDI		2	2	4	2	9	2	2	2	2	2	2	2	2	3	2	3	2	45
ICT	1			3	3	1	4	3	10	3	6	6	6	4	4	4	4	5	67
INN	1	6		4	3	1	4	8	7	2	3	4	5	3	3	3	3	3	63
IFA	5	6	5		3	4	5	9	13	3	10	11	11	4	7	8	8	8	120
JVS	2	4		2		1	4	5	4	4	5	5	5	5	8	8	7	7	76
LCS	2	6	6	3	7	1	5	4	6	5	6	7	6		7	8	8	8	95
OFA	9	2	1	4	1		2	2	4	1	4	4	4	1	3	3	3	3	51
PRO	2	5		4	4	1		5	11	5	6	6	11	4	5	10	5	5	89
R&D	1	5	9	5	4	3	5		6	5	6	6	7	2	4	4	4	4	80
SB-I	4	7		14	6	4	9	9		10	15	16	16	7	11	11	14	12	165
SB-IX	2	2		1	2	1	1	2	4		3	3	3	1	1	1	1	1	29
SB-IV	4	7		12	6	4	8	9	15	10		14	14	8	13	10	10	10	154
SB-III	4	7		14	7	4	8	9	16	9	13		14	8	10	11	12	15	161
SB-II	4	9		13	6	5	11	10	16	9	13	14		8	11	14	12	12	167
SES	2	5		3	6	1	5	4	6	5	6	6	6		6	7	7	7	82
ST-B	2	5		6	4	2	6	5	7	5	11	6	6	4		9	8	8	94
ST-A	3	6		7	7	3	13	7	11	6	8	9	13	6	9		9	9	126
ST-C	1	6		7	5	3	5	5	11	5	6	12	7	4	9	8		11	105
ST-D	3	6		8	7	4	7	7	12	7	8	9	9	5	10	9	11		122
Total	52	96	23	114	83	52	104	105	161	96	131	140	145	76	124	130	129	130	1891

Responses on coordination between SDs through sampling at the same time – Table 31 – allow for much of the same conclusions as those on coordination using same sampling frames. Thus there is a strong coordination between the SBS annexes (save for Business Demography), and likewise but somewhat weaker between the STS annexes. There is also a moderate degree of coordination between the SBS and STS. Additionally, there is coordination between Inward FATS and the SBS annexes, between Innovation Survey and R&D, and between FDI and Outward FATS.

Table 31 - SD 3.6: A heat map for *Same time* responses. Red end of the spectrum indicates high frequency of the responses, green end indicates low frequency.

	FDI	ICT	INN	IFA	JVS	OFA	PRO	R&D	SB-I	SB-IX	SB-IV	SB-III	SB-II	SES	ST-B	ST-A	ST-C	ST-D	
FDI		1	1	2	1	5	1	1	2	1	1	1	1	1	2	1	2	1	25
ICT	0			1	2	0	3	2	5	2	3	3	3	2	3	3	3	3	38
INN	1	5		3	3	1	4	8	6	3	3	4	5	3	3	3	3	3	61
IFA	3	2	3		1	2	3	5	9	1	7	8	8	3	5	5	5	5	75
JVS	1	2		1		0	2	2	1	2	2	2	2	2	4	4	4	4	35
LCS	1	2	2	1	3	0	2	2	2	2	2	2	2		3	4	4	4	38
OFA	4	0	0	2	0		0	0	1	0	1	1	1	0	1	1	1	1	14
PRO	1	2		1	2	0		2	6	2	3	3	6	2	3	5	2	2	42
R&D	1	3	6	3	4	1	4		4	4	4	4	5	3	4	3	4	4	61
SB-I	3	4		10	3	1	6	6		4	15	16	16	5	5	5	7	5	111
SB-IX	1	1		0	0	0	0	0	2		1	1	1	0	0	0	0	0	7
SB-IV	2	4		8	3	1	5	6	14	3		14	14	5	5	5	5	5	99
SB-III	3	4		10	3	1	6	6	16	4	14		15	5	5	5	7	6	110
SB-II	2	4		9	3	1	5	6	15	4	13	14		5	5	5	5	5	101
SES	1	2		1	2	0	2	2	2	2	2	2	2		2	3	3	3	31
ST-B	1	2		2	4	0	4	2	4	2	5	4	4	3		10	9	9	65
ST-A	1	3		2	3	0	5	2	4	2	3	3	4	3	11		9	10	65
ST-C	0	1		2	3	0	3	1	4	1	3	5	3	1	7	8		8	50
ST-D	1	2		2	2	0	4	2	4	2	4	4	4	2	9	9	10		61
	27	44	12	60	42	13	59	55	101	41	86	91	96	45	77	79	83	78	1089

For the remaining two response categories (the positive and negative sample coordination), the number of responses were fewer (398 and 139, respectively) but reflect the same tendencies as for the former two categories, so we omit them here for the sake of space.

5. Assessment

First part of this section summarises main findings obtained in the analysis. Thereafter, next part provides an assessment based on an integration of findings (Section 4) and theoretical considerations presented earlier (Section 2).

5.1. Summaries of findings

5.1.1. Summary of findings on frame coverage

The analysis indicated a number of issues regarding coverage of target populations by sampling frames used by the SDs. The issues seem to be more apparent to respondents in SDs than to those in BRs, perhaps due to a closer involvement of each particular SD in its sampling frame and target population than that which a general sampling frame like BR would be involved with in a particular target population.

Main issues of under coverage concern:

a) under coverage of businesses with specific properties (e.g. in specific employment size or turnover value intervals, etc), which are generally due to restrictions in administrative sources used to update the sampling frame;
b) under coverage of certain NACE activities, which as a rule are “by design” (due to conflicts of regulations or to established practices in the participating countries);
c) under coverage of newly established units, which is a general phenomenon of a time lag due to the chain of events needing to evolve in order for a unit to take a place in the sampling frame; it can be more or less severe, in for instance a large participating country, a separate

study indicated an under coverage in a SD of 22% due to this reason; this is not a deficiency in the SD *per se*, but indicates a weakness in a BR that is to cover the whole field of active businesses;

d) there is also a time component of the restriction in (a), as the dynamically changing properties of the businesses (like increasing or decreasing turnover, employment, etc) – which make the business enter or leave the target population – are registered on the sampling frame with a time lag;
e) under coverage of market activities due to insufficient clarity of the concepts used, that is, inability to distinguish between market and non-market activities;

Main issues of over coverage concern:

a) over coverage due to continued existence, in the sampling frames, of units that have ceased with their activity, caused by the same reason as the under coverage above: a time lag in the chain of events;
b) over coverage of businesses with properties that per definition exclude them from target population, due to businesses passing out of the relevant target population (e.g. have decreased their turnover to below a certain value);
c) over coverage due to inability to distinguish between market and non-market activities;

Generally, the responses indicate that attaining adequate consistency involves trade-offs. For instance, coordination of sampling with other surveys may lead to a, from a particular SD's perspective, less optimal time for sampling than with no coordination at all. Also, administrative data may be advantageous to data quality, process quality or production economy, but disadvantageous to timeliness by coming relatively late.

One may distinguish between "fast" and "slow" admin data: the former are the ones that propagate rather fast from the event itself (a creation of a company) into the admin register; the latter are those for which there is a longer delay between the event itself (closing of an accounting year, annual accounts results; change in NACE activity) and this reaching and admin register (and thereafter the BR).

In some – but not a clear majority – of the participating countries the BRs conduct analyses of their coverage; likewise, in some – but only a minority – of the SDs responding on the question, is some form of clerical or statistical adjustment being made to correct for known deficiencies in their sampling frames.

5.1.2. Summary of findings on relations between BRs and SDs

Uniform use of a business register as a common sampling frame for statistical domains was found to be hampered by:

a) the need of some of the SDs – such that have to cover some specific target populations – for sampling frames of units currently not existing in the BR; examples of that are records of transactions that central banks keep, or businesses engaged in R&D. In such situations the SDs might generate their own register, either with starting from the BR and building upon it by additional sources of information or doing this completely independently of the BR (using external sources like business information firms, media outlets, etc);
b) lack of sufficient quality (completeness) of the BR, where the issues of timeliness (a time lag leading to both under coverage and over coverage) and coverage (of activities, size classes, etc) - already presented in Section 4.1 – are dominant;
c) prohibited access of SDs to BR due to national legislations (in several participating countries, two of them large);
d) perceived unsuitability of the BR to be used as a frame for conducting censuses; however, this result might have just been an unintended consequence of a specific formulation of the question that was used (Question SD 3.2: "If you do not sample from your Business Register…", which holds for all the SDs that do a total survey, a census; more appropriate would have been to query about use of the BR)

Most commonly, a frozen version of the BR was reported to be used. From the given responses it is not possible to discern which patterns of SD properties or system properties govern the decision about

whether to use live or frozen frames, nor how – with the use of a live frame – the threats to consistency between the SDs due to a slightly varying frame are controlled.

However, if the target population of a SD constitutes a subset of the population in the BR, then inconsistencies do arise between that survey and another one, produced using the BR as the frame.

A moderate level, about a quarter of the participating countries, report of a drive to move more SDs towards using the BR as the sampling frame was reported by the BRs; in some of which cases it was in the context of integration into wider systems for production of economic statistics.

After a sample has been selected, there is considerable interaction reported to occur between the BR and the SDs in updating each other on changes regarding the sampled units that have become known to them through their respective channels (administrative sources and surveys).

Regarding the direction from BRs to SDs, it can be estimated that in about three-quarters of the pairings BR-SD information is provided or made available by the BR to SDs. In the opposite direction from SDs to the BR, a corresponding rough estimate is about two-thirds of the pairings where the information is provided by the SDs back to the BR. While NACE code and business closure updates seem to be exchanged in a relatively high volume in both directions, the BR provided relatively more of updates on turnover and employment changes, whereas the SDs provided relatively more of updates on location change. Structure changes appear to be relatively the ones shared the least, probably due to difficulties in finding out about these changes in the first place.

There is, in both directions, a considerable amount of variation in frequencies with which the updates are provided, with outcomes spanning from *immediate* to *annual* and no clear pattern indicating most common choice(s). Without a more developed theory and practice of sample updates and register updates, this variation is not unexpected.

Updates from SDs to the BR represent an additional input to BR, besides administrative sources. Updated information improves quality of the register as well as represents an economic benefit for subsequent surveys to use this updated information. As to the frequency with which this information is used, the prevalent one was *immediately*, followed by *annually*; the latter of the two responses in part likely reflecting the fact that most of the SDs included in the questionnaire study were annual surveys.

Direct updating of the BR by the SDs is disallowed in about four-fifths of the BR-SD pairings; that is, the BRs in most cases do check the information before updating the BR. In that process, if conflicts arise between survey data and administrative sources, then in general they are solved either by giving priority to only one of them (surveys somewhat more preferred than administrative data), or by following a specific rule, or initiating a resolving process.

5.1.3. Summary of findings on temporal aspects

Some components of the temporal aspect have already been covered before, as integrated into the aspects of coverage and of relations between the BR and the SDs. This major topic, however, goes deeper into the working, in time, of processes that provide frames and samples on the basis of which sampling, data collection operations and estimation are to be carried out.

In interpreting these results, one is aided by a representation of a generic chain of transactions that uses the BR, which questions on the temporal aspect were aiming to illuminate (Figure 1).

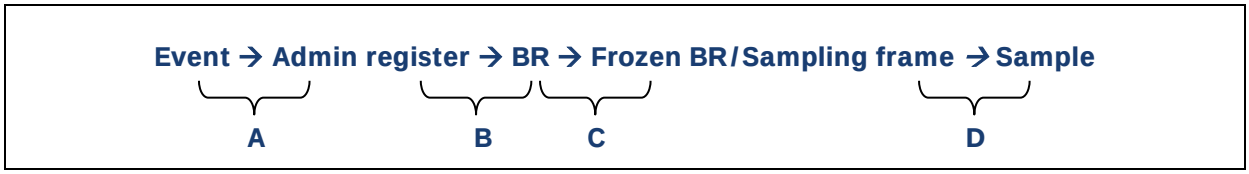

Figure 1 – Chain of activities leading to a sample for an SD. The italicised component is optional, as that component does not apply to SDs that use live frames.

Step A in Figure 1, from the event (like creation of a new business, its employees being these specific persons at a specific point in time, and so on) to the administrative source, was already presented in Section 5.1.1. There were not sufficient data in the responses to draw more specific conclusions about the length of this lag, but it may be assumed that there are considerable differences between the participating countries – in some cases, the lags may be considerable.

From responses to the question BR 1.6, the time lag in step B in Figure 1 – between the event being registered in an administrative source (VAT, social security, etc.) and that event being recorded in the BR – was about three months or less, with some variation naturally occurring between the participating countries. Frequency of updating of the BR with information arriving in step B was for both new units and existing units, and for a number of administrative data variables (VAT, personal and corporate tax, social security data, etc.), mostly annual, but also with a notable number of sub-annual frequencies occurring. This might indicate that either by design or by established practices information that comes to the BR with a time lag of not more than three months in some cases waits up to another nine months before being incorporated into the BR. It is however not possible – based on the current data – to say more on the actual impact of such practices on consistency of those SDs that use the BR as the frame. For instance, if the BR is used for sampling for annual SDs only, then it is completely possible that the BR still provides sufficiently up-to-date information to the sample created from it.

Frozen versions of the BR (step C) seem mostly to be created annually, but also monthly and at some other frequencies. For the surveys using them, this frozen frame is current mostly within either a year or a month. Timeliness of BR data is by the SDs generally perceived to be quite high, over 90% of the responses indicating sufficient timeliness, but with a relatively lesser proportion of such responses for Job Vacancy Statistics and Outward FATS.

Taking sample from the sampling frame (step D) occurs throughout the year, but with a pronounced preference for time around the new calendar year. It appears that this practice is motivated by when frozen sampling frames are created – that is, at this same time – enabling use of a recently frozen frame that provides both best timeliness of the sample and best coverage of the target population.

The results thus indicate that there are some inconsistencies between countries in the trade-off between horizontal consistency and timeliness. This may concern the entire system of business-related statistics, and also specifically particular statistical domains, from both users' and producers' perspectives.

5.1.4. Summary of findings on maintenance of the BR

The questions of this major theme covered practices of maintaining and updating the BR, touched upon also in the theme about relations between the BR and the SDs (Section 4.2).

The responses indicate a considerable variability in the BRs practices of maintenance of the BR and data update. This may not be surprising as there, as far as known, is no regulation on these practices. Summarising the two, Figure 2 depicts the processes taking place in this area.

The variation comes from both externally available sources of information in the participating countries and internal sources and practices of updating the BR. For instance, while in some countries there is an administrative body with whom any new business needs to be registered, in others there is not such a body. In these latter countries, a new business is observed only after the business has provided its tax data, after the end of the calendar year in which it was established (a minimum average delay of six months), or at the end of its first fiscal year (a minimum average delay of twelve months). Additionally, in some countries – a notable example being Germany, a large member state – the legislation disables sharing of data between governmental agencies, which discourages reuse of administrative data and adds to the burdens of businesses.

Methods to assign values of register variable to units vary considerably between the participating countries, including for instance whether a date stamp for record formation or change is saved or not.

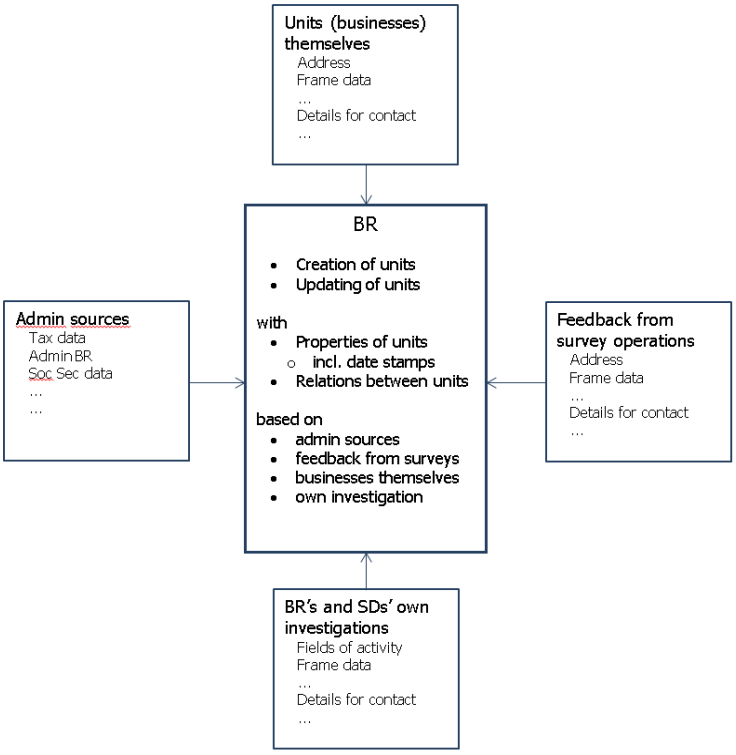

Figure 2 – Creation and update of information on units in the BR.

Some of the responses can be deemed as indicating a potential to cause consistency issues, like for instance non-zero cut-offs mentioned in responses to the question BR 1.16, which might lead to units excluded from the BR although conducting some (sub-threshold) activity. Additional issues, mentioned especially in responses to questions BR 1.9 – 1.14, indicate the kind of variability in the procedures that may introduce an error in coding of the units on the variables in question (*NACE*, *employment*, *employees*, *turnover*, *institutional sector codes*), however with possibly lesser impact on consistency (except when an error is found and corrected, forcing re-publication of the statistics).

5.1.5. Summary of findings on reference periods

The main issue found in this major theme concerned reference periods other than a calendar year. The extent to which this was occurring in practice in the participating countries was reported as varying between a very low of a couple of per cents of the businesses to almost a quarter of the businesses.

Depending on whether some adjustments are performed in such cases or not, the latter case of no adjustment can lead to some estimation issues, especially so if recent change is to be estimated. Namely, the principle of using the fiscal year data as pertaining to the calendar year in which the fiscal year ends means that data from these businesses will be older than those from businesses where the fiscal and the calendar year coincide. A change occurring later in the reference period might thereby not be included in their data, and then the estimate of change becomes incorrect.

5.1.6. Summary of findings on sampling methods and coordination

While sampling method by itself does not represent an impediment to consistency, there are possible gains in coordinating the samples. Namely, samples taken from the same (frozen) sampling frame are improving consistency, as they refer to the same reference period (that of the frozen frame) and are characterised by same auxiliary information (this latter enabling consistency of breakdown variables as well as better consistency of the statistical estimates).

The analysis indicates that the strongest cooperation exists within the two groups of SDs already considered as related, namely the SBS and the STS annexes. Quite strong was also the cooperation between the two (the SBS and the STS), and also between these two SDs on the one hand and a number of other of the SDs.

In analysing sample coordination, one should distinguish between SDs of different periodicity, thus for instance monthly on the one hand and annual on the other, and also what kinds of estimates these SDs produce. Therefore, one may as one alternative prioritise integration of SDs of the same periodicity, as frame management, sampling and data collection can be performed jointly for all the surveys. However, even if two SDs are of different periodicity, their integration might be needed in order to improve consistency of their output. Having in mind that short term statistics (monthly) aims to produce forecasts of change while structural statistics (annual) to give reliable estimates of levels, there may be an interest that these two sources become more consistent). A way to achieve this consistency is then through integrating these two SDs of different periodicity. (See also Zwijnenburg 2012.)

Thus, for two surveys of the same periodicity, there are reasons – mentioned above – to strive for their sampling coordination, provided there are no major obstacles. These may include different temporal patterns of availability of relevant information, requirements to publish in vastly separate points in time, and similar.

For two surveys of different periodicity, the question is how the estimates produced by the more frequent of them (say, a monthly survey) relate to the estimates produced by the less frequent of them (say, an annual survey). Is the former estimate expected, when aggregated over the months, to yield the same or similar value as the estimate given by the annual survey? Or, are they simply two different – even if related – variables (say, a projection and an actual outcome), so that there actually is no real need for these estimates to be numerically the same? Such considerations and requirements need to be worked out when designing systems that are to produce such related statistics, which at the current state of knowledge, theory and practices of official statistics is an intricate process that cannot be fully guided by recommendations and regulations.

5.2. Integrated evaluation of findings

A starting point of this evaluation is a vision of an integrated system for production of statistics (see Section 2.3), as it is through such systems that consistency is assured "by design". While this vision is not yet fully implemented in any National Statistical Institute (NSI), the developments in a number of NSIs that have come some way towards these integrated systems are referenced in Section 2.3. They are all characterised by a long period of development, on the order of five or ten years and ongoing, indicating a considerable effort needed to put such a system in place even if the system is developed within a single country's/jurisdiction's frame of reference.

It is thus necessary to underscore that, in addition to resolving issues of horizontal consistency which the developments above – being on a single country level – were addressing, a European project that is to address also the issues of vertical consistency will have additional dimensions of complexity.

5.2.1. Relevance for the entire system of business-related statistics

A pronounced result of the analyses of questionnaire responses is a considerable degree of variation between the participating countries and between the statistical domains (SDs) in them with respect to many of the processes that are a part of producing the 19 SDs that were included in the questionnaire survey. Using the framework outlined in Figure 1, we will go through the components and indicate major deviations, each of which with the potential to be of relevance for the entire system of business-related statistics.

The discussion presented in Table 32 (next three pages) reflects the processes within an SD, that is, within a certain "stovepipe". There is a clear need and impulse (see Section 2.3) to integrate these into wider systems (or at least subsystems, to begin with) for production of components of economic statistics. Thus the central aspects of the processes presented in Figure 1 are likely to become centralised and standardised, thus reusable for a number of participating SDs. This especially puts the BR in a central position in such systems, and requires a particularly thought-thru design of the system.

How far this centralisation of the process might go is a matter for judgement and consideration, as it may include the BR as the central instance of frame management, but also the BR's production of

frozen frames for the SDs' use, and further a common sample for a set of SDs that are integrating their production process.

However, for a BR to play an increased role in common processes for statistics production, there ought to be an increased requirement of covering well the majority of SDs' target populations. Failure to do so might marginalise the BR away from its intended role.

With an increased BR coverage (tending to a 'total' coverage), issues like units passing in and out of target population restricted by certain thresholds ought to be easier to handle.

While over coverage of units actually no longer in the target population is statistically a fairly simple matter (once it is recognized), considerable inclusion of such units into operating samples generates additional data collection costs to the SDs during field operations, which can be more or less severe depending on the extent of the time lag, the amount of real-time changes in the business population unit and data collection costs. For the reason of effective use of resources, attempts ought to be made to reduce to a sufficiently low level the extent of over coverage.

5.2.2. Conclusion on effects for major statistical projects

The preceding subsection gave a general overview of the issues identified in the analysis of the questionnaire results. In addition to these, as reported earlier on in this deliverable, there are a number of issues that relate more or less directly to specific statistical domains (SDs), some of them major. The following may be among those to be addressed early when drawing action plans based on the project report.

FDI/FATS. In a number of participating countries, FDI or Outward FATS are done by central banks in a process separate from that for production by an NSI of most of the other official statistics, which makes this a typical example of a "stove pipe" model for production of statistics: it is difficult to integrate it with the production of other statistics. Additionally, either due to organisational issues (placement outside an NSI) or to legal obstacles, this SD in such cases may not to a sufficient extent rely on the BR. While the BRs tend not to have sufficient information to delineate the target population for FATS, use of the BR as a wider frame with which to coordinate activities for data collection for FATS would be likely to increase consistency between FATS and the other SDs, those that rely on the BR in their production. With the current state of affairs, there might be raised serious concerns about FATS's consistency, both horizontal (within the country) and vertical (to the European level statistics), with the other produced official business statistics.

5.2.3. Conclusion on consistency problems seen from users' perspective

The survey and its results did not specifically address concerns of the users. It is however known that inconsistencies between published statistics do cause user concern. In some cases it is a legitimate concern as a "same" variable is estimated in two SDs (Statistics Sweden 2003), while in some cases it stems from a comparison between a rapid estimate and a detailed estimate, where the latter has more data available than the former (Statistics Sweden 2008). Additionally, there are cases where the two variables, produced by two SDs, although similar, are not the same and thus complete consistency cannot be expected. For instance, short term statistics (e.g. monthly) may produce timely statistics based on the data available at that point in time; structural statistics is produced later on, when data for the whole period (e.g. a year) are available. The two statistics need not be consistent, for instance if one concerns predictions and the other achieved results; however, there is a request on the user side that the differences are minimised and causes of the potential inconsistencies are identified and their effects minimised.

Table 32 – Main issues and findings.

Component	Issues	Finding	Comment
Event -> Admin record	Under coverage (employment, turnover) Time lag (active/passive units; changes)	The participating countries vary considerably in the availability and timeliness of the different kinds of administrative records, as well as in the legal premises for sharing this information between governmental sectors (see next item). Ways of improving timeliness of these chains ought to be an active area for improvement.	There appears to be little chance to impact these factors. The consequence is that there will remain a considerable variation in the input data to Business Registers (BRs)/sampling frames, and thus in content and timeliness of BRs/sampling frames. However, as large amounts of unstructured data are increasingly becoming available, there might occur new sources of data, that in some cases can substitute for what is missing due to time lag and other reasons in admin data, and in some other provide relevant information that was not available before.
Admin record -> BR	Under coverage (employment, turnover, NACE, legal form) Delineation market/nonmarket Time lag (active/passive units; changes) In/out of target population (TP) Specially defined TPs (R&D, FDI, etc) Legal impediments	In addition to restrictions in the input to BRs/sampling frames, there may occur additional restrictions in the content and timeliness due to the way the BR (or another sampling frame) is established, maintained and updated. For instance, NACE sections P, Q, O, R, T, A and S; turnover below a threshold of up 80 K EUR; units with zero employees reported as having restrictions in coverage in a number of participating countries. Time lag between updates in admin data and updates in the BR varied from immediate to a year, implying differing practices and subsequent quality of the BR as a sampling frame. Target populations defined not by properties but by activity are an additional issue, as they present special challenges. Much variation in NSIs' practices was found. Large variations in participating countries regarding legal impediments to sharing admin data: from a restricted approach (like in DE), over fairly open (Nordic countries), to very open in a structured way (PT, EE).	To the extent that they are not caused by restrictions in input data, it should be feasible to increase to a certain degree uniformity between the national BRs with respect to data and procedures, and thus enable "by design" better vertical and horizontal consistency. This ought however to be a balanced decision, taking into account increased costs and potential burden, as well as the extent of benefits of making the change. A measure of BR quality – likely a multidimensional one – could be introduced in order to give a fast, standardised representation of a BR's usability w.r.t. content and timeliness. This may include results of own analysis of coverage, as well as existence of proposed methods for adjustment/compensation for known restrictions (if any).

Maintenance of the BR	What reference point in time or reference period does the BR represent? Is the BR updated to support in the best way carrying out of survey field operations?	Aim of a BR is to reflect as well as possible the actual state of affairs in the target population at a specific time or period. In some of the participating countries' BRs, information is lacking to determine which point in time they are reflecting (i.e. information is not saved about what reference period the data relate to, implying an unspecified "now"). In addition to being able to represent the state of affairs at a specific (past) period or point in time, the BR should also reflect in the best way the latest state of affairs, in order to support carrying out of survey operations by SDs (see below). Updates to the BR come from: – Admin data – Own BR surveys and other forms of seeking information (e.g. professional sources) – SDs, upon they finding new information while carrying out survey field operations – The businesses themselves (only occasionally, though) While a Frozen BR is the dominant form of the BR used for sampling in the participating countries, there are a number of instances where a 'live' BR is used. Practices vary strongly between the participating countries in ways of what is shared between the SDs and the BR, and how often the information is shared.	It should be feasible to increase to a certain degree uniformity between the national BRs with respect to maintenance of BRs and principles of relations between BRs and SDs, and thus enable "by design" better vertical and horizontal consistency. This ought however to be a balanced decision, taking into account increased costs and potential burden, as well as the extent of benefits of making the change. Specific topics to pay attention to: – Live vs. frozen frames – Changes in the sampling frame in the course of its use: pros and cons, specified for the different changes (inclusion/exclusion of units separately from contact details) – Relative priorities of the different source input to BR
BR -> Frozen BR/ Sampling frame	Time lag in admin data Balance between consistency and timeliness	Frequency of producing frozen frames varies between the participating countries' BRs. Mostly pronounced are annual and monthly frozen frames. This may naturally depend on periodicity of the involved SDs.	It appears difficult to set norms or requirements on this periodicity *without* respect to the context of use and other possibly related matter (e.g. survey operations). Consistent with what seem to be the experiences of Statistics Canada, Statistics Netherlands and several other NSIs, this should be developed taking into account specific user needs and requirements of the SDs that are to use the sampling frames. A prioritizing process is likely needed to occur, taking into account the somewhat conflicting aspects of horizontal consistency and timeliness, as well as vertical consistency, user relevance, costs, etc.

Frozen BR/ Sampling frame -> Sample	Balance between consistency and timeliness	While taking samples has shown to be spread throughout the year, somewhat more prominent were the months around new year, that is, the period November – February. There is a trend to take the sample as soon as the most recent (and relevant) sampling frame is available, in which case the sampling takes place soon after availability of such a frame. Coordination of sampling between SDs occurs to some extent, in which case it is mostly through sampling from the same frame, then also at the same time, and to a lesser extent by positive or negative sample coordination. The Structural Business Survey cluster and the Short Term Statistics cluster are examples of the strongest sample coordination, but there are other examples too of coordinated sampling between SDs.	Like above, it appears difficult to set norms or requirements on procedures and times for drawing samples *without* respect to the context of use and other possibly related matter (e.g. survey operations). This should be developed taking into account specific user needs and requirements of the SDs that are to use the sampling frames. A prioritizing process is likely needed to occur. Experiences about integration of SDs into wider (sub)systems for economic statistics production need to be collected, studied and formalised.
Sample -> Survey field operations	Improvement of consistency by sample coordination Balance between consistency and timeliness	While no explicit questions on this area were posed in questionnaires to the SDs, open responses indicated their need of having as updated as possible contact information and information on unit characteristics of the sampled units in order to improve efficiency of field operations. This phase also provides information about recent changes in the target population and thus a possible input for updating the BR (see above)	
Reference periods	Consistency threat due to differing reference periods Consistency threat due to differences in approaches to adjustment for different reference periods	The units' reference periods (in e.g. accounting) not aligned with SDs reference periods (as defined by regulation). Varies from a very low proportion in some participating countries to considerable (quarter of businesses) in some other countries.	This is a property of businesses and their accounting practices, that is, it belongs to the businesses' side of data collection. However, this impacts consistency of produced statistics and thus needs to be addressed in some way.

5.2.4. Conclusion on consistency problems seen from producers' perspective

Practically all of the issues covered in subsection 5.2.1 are also concerns regarding consistency (and violations thereof) seen from the producer's perspective. The producer needs to be aware of threats to consistency outlined in the framework of section 5.2.1, and also initiate activities to adequately address these concerns.

Major stakeholders in the context of the present project include:

- NSIs and other producers of official economic statistics of the participating countries,
- internal (e.g. National Accounts) and external national users of economic statistics output of the SDs,
- Eurostat and other producers of official European economic statistics,
- internal and external users, supranational and national, of the official European economic statistics output.

While the stakeholders are many, two main levels can be discerned: the national level and the EU level, corresponding to the horizontal and vertical consistency perspectives. While this project is initiated from the EU level and takes such a perspective, the project did take note of developments towards integration of economic statistics developing on the national level, both within the EU (e.g. the Netherlands, Finland, France among others) and outside it (e.g. Canada, New Zealand among others).

Developments towards integration of economic statistics on both the national level and the supranational level are likely to continue, due to the expected benefits of integration of statistics (the transition from 'stove pipes' to integrated systems). As discussed in Section 2.3, these include increased efficiency, reduced costs, improved consistency by design, better ability to change the system and innovate the system, reduced administrative burden to businesses if larger use of administrative data is possible, to name the most prominent ones.

Parenthetically, there exist other European developments that partially cover same goals, like larger access to and sharing of public sector data, as for instance the Digital Agenda for Europe strategy and its component called Public Sector Information initiative.

It is further likely that developments will continue to occur on both the national level and the supranational level, which will aim at meeting partially same and partially different goals, from the contexts that are valid in each locale/jurisdiction. Which priorities will prevail is dependent on the context in which the development project takes place. It is therefore important that the different centres – on both the national and supranational levels – which develop integrated systems for economic statistics or contemplate development of such systems are well informed on what the other related systems are developing, which is needed in order to facilitate that these meet in an as optimal way as possible.

6. References

Bakker, Bart F. M, and Piet J. H. Daas (2012). "Methodological challenges of register-based research". *Stat Neerlandica* 66(1), pp.2-7. DOI: 10.1111/j.1467-9574.2011.00505.x

Biemer, Paul P. (2010). "Total Survey Error: Design, Implementation, and Evaluation". Public Opin Q (2010) 74 (5): 817-848 doi:10.1093/poq/nfq058.

Braaksma, Barteld, and Gert Buiten (2012). "Redesign of the Chain of Economic Statistics in the Netherlands". In on-line Proceedings of *The 4th International Conference on Establishment Surveys*, http://www.amstat.org/meetings/ices/2012/papers/302202.pdf (accessed 15 October, 2013).

Brinkley, Eden, and Annette Scott (2012). "Implementing an Integrated Program of Economic Statistics – an ABS Perspective". In on-line Proceedings of *The 4th International Conference on*

Establishment Surveys, http://www.amstat.org/meetings/ices/2012/papers/302206.pdf (accessed 15 October, 2013).

Colledge, M.J. (1995). "Frames and business registers: an overview". In *Business Survey Methods*, Eds. B. Cox, D. Binder, N. Chinnappa, A. Christianson, M. Colledge and P. Kott. Wiley.

Creswell, John (1994). *Research Design: Qualitative & Quantitative Approaches*. Thousand Oaks, CA: Sage Publications, Inc.

European Central Bank (2008). "ECB Statistics Quality Framework (SQF)". Available at http://unstats.un.org/unsd/dnss/docs-nqaf/ecbstatisticsqualityframework200804en.pdf (accessed 15 October, 2013).

Eurostat (2011). European Statistics Code of Practice - revised edition 2011. ISBN: 978-92-79-21679-4, DOI: 10.2785/18474, http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-32-11-955/EN/KS-32-11-955-EN.PDF (accessed 15 October, 2013).

Hedlin, D. and Wang, S. (2004). "Feeding Back Information on Ineligibility from Sample Surveys to the Frame". *Survey Methodology*, 30, 167-174.

Ravindra, Daniela (2012). "The Integrated Business Statistics Program: An Overview". In on-line Proceedings of *The 4th International Conference on Establishment Surveys*, http://www.amstat.org/meetings/ices/2012/papers/302175.pdf (accessed 15 October, 2013).

Saint-Pierre, Étienne (2013). "Rolling Estimates and the Dynamic Assessment of Quality: a Generic Approach and a Source of Efficiencies". Paper prepared for *The 2013 European Establishment Statistics Workshop* (EESW13), Nuremberg, Germany, September 9-11, 2013, http://enbes.wikispaces.com/file/view/Saint-Pierre%202013.pdf/456105480/Saint-Pierre%202013.pdf (accessed 15 October, 2013).

Seyb, Allyson, Ron McKenzie and Andrew Skerrett (2013). "Innovative Production Systems at Statistics New Zealand: Overcoming the Design and Build Bottleneck". Journal of Official Statistics, Vol.29, No.1, 2013. pp. 73–97, http://www.jos.nu/Articles/abstract.asp?article=291073 (accessed 15 October, 2013).

Statistics Sweden (2003). Individ- och företagsbaserad sysselsättningsstatistik – en fortsatt jämförelse mellan AKU och KS (in Swedish; "Employment statistics from a social survey and an establishment survey – a continued comparison between Labour Force Survey and Short-Term Employment Statistics"). ISSN 1103-7458.

Statistics Sweden (2008). Improvement of the Swedish rapid estimates in the foreign trade on commodities. ISBN 978-91-618-1467-1.
http://www.scb.se/statistik/_publikationer/HA9999_2008A01_BR_X100BR0812.pdf (accessed 15 October, 2013).
United Nations Statistics Division (not dated). "Fundamental Principles of Official Statistics". Available at: http://unstats.un.org/unsd/methods/statorg/FP-english.htm (accessed 15 October, 2013).

Wallgren, Anders, and Britt Wallgren (2007). Register-based statistics: administrative data for statistical purposes, John Wiley & Sons, Chichester, UK.

Zwijnenburg, Jorrit (2012). Redesign of the system of economic statistics in the Netherlands. On-line proceedings of Q2012 European Conference on Quality in Official Statistics. Available at http://www.q2012.gr/articlefiles/sessions/14.2_Jorrit%20Zwijnenburg_Redesign%20of%20economic%20statistics.pdf (accessed 15 October, 2013)

7. Appendices

Appendix 1. Questions grouped into six major themes, with question content

Appendix 2. Broad Structure of NACE Rev. 2 - levels 1 (Sections) and 2 (Divisions)

Appendix 1: Questions grouped into six major themes, with question content

	Major theme	BUSINESS REGISTERS	STATISTICAL DOMAINS
1	Frame coverage	BR 1.1) Are there any restrictions to coverage for your register i.e. thresholds BR 1.2) Are there any cut-off thresholds for the administrative data sources on your business register? BR 1.3) Has an estimate been made of the coverage of your Business Register in relation to the target population? BR 1.4) Do you have any issues of under coverage in relation to your target population? BR 1.5) Do you have any issues of over coverage in relation to your target population?	SD 1.1) Do you have any specific difficulties in covering the target population in your country? SD 3.7) Are you aware of any discrepancies between your target population and the frame available from BR SD 3.8) Do you adjust for discrepancies between your target population and frame …
2	Use of BR by SDs including Relations between BR and SD	BR 2.1) Which surveys / domains use the Business Register as a source for sampling? BR 2.2) For those surveys / domains not using the Business Register, are steps being taken … BR 2.5) Do you provide information on changes to the register for those surveys using frozen … BR 2.6) What information is fed back to the survey areas and how frequently? BR 2.7) Do you receive feedback from the following surveys on units in your register and how frequently? BR 2.8) What information is fed back to the Business Register from surveys? BR 2.9) Is the information received from surveys updated immediately or periodically? BR 2.10) Do you allow any of the following survey areas to directly update the business register? BR 2.11) How do you resolve conflicts of information between survey area feedback and administrative sources? BR 2.12 Do you give priority to a specific survey / domain when resolving conflicts in NACE classification? BR 2.13) Do you give priority to a specific survey / domain when resolving conflicts in employment values? BR 2.14) Do you give priority to a specific survey / domain when resolving conflicts in turnover values? BR 2.15) How are changes in units' NACE codes dealt with on your register?	SD 3.1) What frame do you take your main sample from? SD 3.2) If you do not sample from your Business Register ... why not? SD 3.10) Is information on changes to the Business Register since ... sample available ... SD 3.11) What information is available from Business Registers relating to your frame and how frequently SD 3.12) Do you update your sample frame for the following changes and if so from which sources SD 3.13) Are you allowed to directly update the Business Register for the following changes? SD 3.14) For the following list of variables, please indicate if you feedback information to BR SD 3.15) If you do not feed back information to BR ... why not

	Major theme	BUSINESS REGISTERS	STATISTICAL DOMAINS
3	Temporal aspect Including BR management	BR 1.6) Has an estimate been made for the lag between the administrative BR 1.7) For new units how frequently do you receive and update ...from administrative data BR 1.8) For existing units how frequently do you receive and update ...from administrative data BR 2.3) How frequently are frozen extracts created from your Business Register? BR 2.4) For those surveys using a frozen extract from Your Business Register, how current is the extract?	SD 3.3) When do you take your main sample? SD 3.4) Why do you take your sample at this time? SD 3.9) Is the data ... from your Business Register of sufficient timeliness …
4	BR maintenance	BR 1.9) What information is available to allow NACE coding of units on your business register? BR 1.10) How is employment derived on your register for new units? BR 1.11) How are employees derived on your register for new units? BR 1.12) How is turnover derived on your register for new units? BR 1.13) Do you record institutional sector codes for units on your statistical business register? BR 1.14) How do you calculate institutional sector codes for units on your statistical business register? BR 1.15) Do you record inactive units on your statistical business register? BR 1.16) How do you identify inactive units on your statistical business register? BR 1.17) Does your Business Register record a date stamp against the following variables [unclear response alternative]	
5	Reference periods		SD 4.1) Do you experience ... difficulties in meeting the reference period requirement … SD 4.2) What period do you collect survey data for? SD 4.3) Do you adjust for units returning data for different periods to calendar year?
6	Sampling methods and sample coordination		SD 3.5) What method is used to derive the sample from the frame population? SD 3.6) Is your sample coordinated with other surveys/domains and if so how?

Appendix 2: Broad Structure of NACE Rev. 2 - levels 1 (Sections) and 2 (Divisions)

Section	Title	Divisions
A	Agriculture, forestry and fishing	01 – 03
B	Mining and quarrying	05 – 09
C	Manufacturing	10 – 33
D	Electricity, gas, steam and air conditioning supply	35
E	Water supply; sewerage, waste management and remediation activities	36 – 39
F	Construction	41 – 43
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	45 – 47
H	Transportation and storage	49 – 53
I	Accommodation and food service activities	55 – 56
J	Information and communication	58 – 63
K	Financial and insurance activities	64 – 66
L	Real estate activities	68
M	Professional, scientific and technical activities	69 – 75
N	Administrative and support service activities	77 – 82
O	Public administration and defence; compulsory social security	84
P	Education	85
Q	Human health and social work activities	86 – 88
R	Arts, entertainment and recreation	90 – 93
S	Other service activities	94 – 96
T	Activities of households as employers; u0ndifferentiated goods- and services-producing activities of households for own use	97 – 98
U	Activities of extraterritorial organisations and bodies	99