25th Meeting of the Wiesbaden Group on Business Registers
- International Roundtable on Business Survey Frames

Tokyo, 8 – 11 November 2016

Irene Salemink
Statistics Netherlands
Session No. 4

Quality and Coverage

Statistical Business Registers as Backbone; issues regarding Business Register and outcome quality

Abstract

Statisticians nowadays face the increasing demand for integrated, consistent and coherent statistical information. At the same time they have to deal with the reality of the ever increasing complexity of organization of Global Enterprise Groups, difficulties to retrieve statistical data, the increased availability of administrative sources and the demand to use this type of data. And also small enterprises start to pay their contribution to the complexity issue concerning issues on outsourcing and the self-employed.

Major challenge for both statisticians and statistical business register staff is how to deal with this increased complexity, the pressure to combine all various data sources as efficiently and less burdensome as possible and still be able to release high quality frames and statistics in time.

This paper describes a possible approach in how to deal with this challenge and discusses various aspects of the organisation of related statistical production processes as a chain of statistical products. A crucial role is reserved for the Statistical Business Register (SBR) which at Statistics Netherlands (SN) is positioned as backbone for economic statistics and plays a central role in their production.

Related business economic statistics, including National Accounts, are designed and organized as a chain of statistical processes. This chain starts with a coordinated population of statistical units derived from the SBR and the operation is placed under central direction; the so called chain-management. The approach of chain management enables SN to orchestrate both the process and the statistical outcomes and governs quality management, including procedures to maintain SBR, coordination of adjustments of statistical estimates due to errors in the SBR populations and prevention of inconsistencies.

The SBR is the first place to correct errors and to repair them on a coordinated way in statistics. However, it's not only the quality of the SBR itself that counts, also the economic indicators derived from it are of importance for customers. In order to control the quality of the chain as a whole, the maintenance of the SBR is tuned with the statistical processes, e.g. by defining different maintenance groups of statistical units in the SBR or combining data of different sources (Administrative, National Central Bank etc.).

A major challenge lies in the existence and use of multiple sources and various modes that will further complicate the compilation of economic statistics. In this model of a "mixed-mode-multisource-approach" large and complex enterprises are dealt with in a custom fit approach, the role of administrative sources is enlarged and implementation of Standard Business Reporting & Reference Chart of Accounts is foreseen, with a limited collection of remainder data and variables.

1. Economic Business Statistics; A Chain of Statistical Processes

National Accounts are compiled using various data sources of which a substantial part originates from the business economics domain. The compilation of the underlying statistical data is most commonly organized like so called stovepipes, in the sense that the individual statistic is produced as a standalone process. The demand for integrated, consistent and coherent statistical information however increases and working in a standalone manner does not contribute to this demand, on the contrary. Due to the ever increasing complexity of the organization of global, large, complex enterprise groups, the difficulties to retrieve statistical data, the various entry points for various business statistics (International Trade, SBS, STS, FATS, FDI), increased availability and use of administrative sources, the complexity of small enterprises comprising phenomena like outsourcing, the increased economic role of the self-employed, the difficulties to describe everything correctly in the Statistical Business Register and the subsequent compilation of statistical data makes that none of the compilers of any business statistic alone can oversee this system of dependencies as a whole. It takes a helicopter view to oversee (at macro level) the increased links and dependencies between enterprises and economic sectors.

Subsequently a major disadvantage of the stovepipe approach is considered that this approach affects the quality of the statistical output; the accuracy of the individual business statistics outcome (for instance the discrepancy between the sum of the quarterly turnover and the annual turnover), the comparability within domains, timeliness and last but not least the consistency between statistics. In order to improve the quality of the system of business statistics and national accounts it's not sufficient to focus on the quality requirements of the individual local process (although that would be the first step [1]). Instead the focus of the primary production process should be on the quality requirements of the whole chain. The quality of which may be improved by coordination across statistics of definitions, of populations, of classifications, of methodology and the definition of a general revision strategy.

Another pitfall of a stovepipe approach is besides sub-optimisation of the local processes also the lack of re-use of data in and between processes. This may result in a large administrative burden, redundant data storage, and various "single points of truth" i.e. one for every separate process. This approach makes it also difficult to ensure that methodological issues and discontinuities (for example a change in the mode of data collection) are dealt with in the same manner by all statistical processes.

Besides the political pressure to reduce administrative burden the same holds for an increased use of (fiscal) administrative sources. When it comes to re-use of data and use of administrative data, usage of the same business register with central linkage of administrative units to statistical units is a prerequisite, often requiring a redesign of the whole chain.

So how to deal with this increased complexity, the pressure to combine all various data sources as efficiently and less burdensome as possible and still release high quality frames and statistics on time?

A possible solution is to integrate all stages of the production process of related business statistics and national accounts as part of a chain of statistical products, using chain management to orchestrate both the process and the outcome. The Statistical Business Register (SBR) is a crucial part of this chain that starts with a coordinated population derived from the SBR. Therefore in this approach the SBR is positioned as the backbone for all economic statistics and plays a central role in their production.

For Statistics Netherlands (SN) chain management is considered crucial to oversee nowadays the effects and impact of Business Register related events and to judge in coherence the relations between

statistical events. Therefore all related business economic statistics and the National Accounts are designed and organized as a chain of statistical processes.
In order to be able to execute decisions and effects of decisions in a coordinated manner chain management and a culture of shared responsibility between all partners in the chain are a prerequisite.

2. Chain Management

2.1 What is Chain Management?

Within our quality management system SN distinguishes between process management and chain management. Process management is confined to the direction and correction of the happy or unhappy flow within one process. Monitoring of the process flow is vital and makes feedback loops possible. Due to process management the owner of the process can adjust the process within manageable margins and be 'in control'.

Chain management is the coordination of the various statistical processes i.e. the processing and designing of various statistical products from the perspective of the whole chain. It comprises the whole set of management and operating activities which aim for improving the cooperation of all actors in the chain so that the result of this joint effort is optimal and transparent for all users.
The chain is designed as a set of links between processes. SN defined these links as so called "steady states" in the Business Architecture (BA). Chain management concerns the links of the whole statistical production process from observation to publication. It is however limited to steering those issues with dependencies between statistical processes and consecutive processing steps. Within the intermediate processes it is the responsibility of the process owner to control the process via process management.

However, chain management is not only about statistical processes and statistical outcomes. A rather crucial aspect of chain management is the acknowledgement that it is a way of working together in a non-hierarchical way. It is the art of working together, aiming for tuning activities. Therefore chain management cannot without agreements on management and operating activities that have one goal and that is to continuously improve the cooperation of all actors in the chain and even between chains.

2.2 Working under Business Architecture

An important role in the design of the chain of business statistics and a necessary condition to implement chain management has been fulfilled by the SN Business Architecture and the adoption of working under architecture. In the BA implemented at SN three layers are defined; Design, Chain management (concerning activities like planning, monitoring, data quality management) and the actual Statistics production from the start of data collection to publishing. Instead of stovepipes for each individual statistic, the architecture describes these statistical processes as a value added chain and defines several "steady states" in which the statistical data have a well-defined status. The first steady state is of course raw data, the final one the publication data. In the Business Architecture the statistical processes are split up in process steps that take the data from one steady state to the next one. The BA describes statistical processes as a value added chain in which each process step adds value by eliminating uncertainties about the data. At every step checks for possible errors are executed as well as checks for specific possible diversions from the metadata, which are corrected if necessary. This means that at every step the data have attained a specific higher level of quality.
This approach also facilitates the use of data from one process in another one; by making use of the steady states the possibilities to re-use data in other processes is largely increased. By working under architecture there exists no overlap between processes; the work is only done once and at the most efficient place in the overall process.

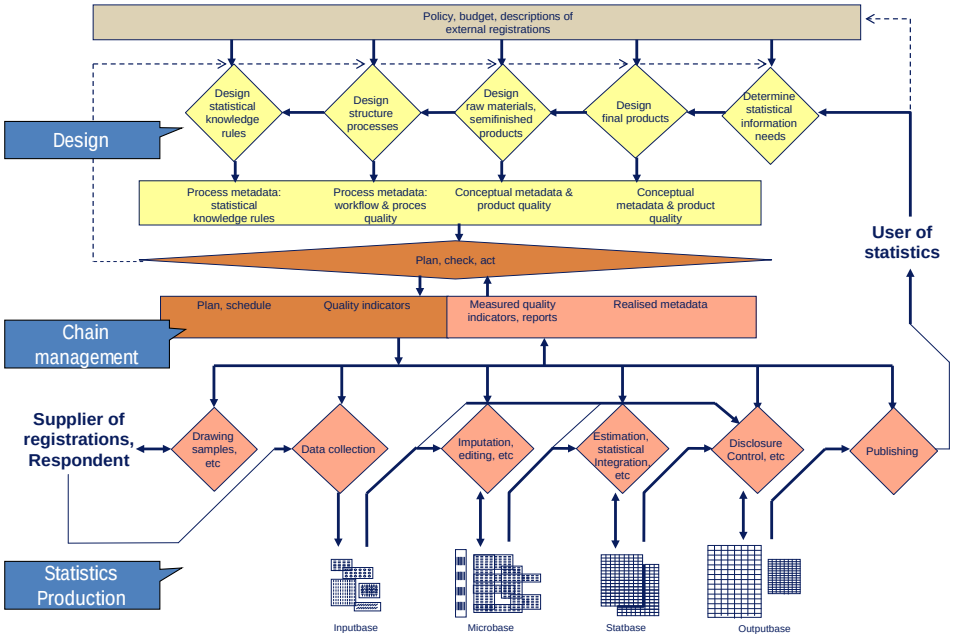

Figure 1: Business Architecture (GSBPM) implemented at SN. In this picture the steady states are represented by the "bases" depicted in the bottom layer

The concept of steady states is a key element in the BA. They can be accessed by generic tools and services, they play a crucial role in the redesign and production, because of the obligatory use of them, and they contain besides data also the relevant metadata. Working under BA may be demanding, the reorganization of statistical processes alone is time consuming and implies a temporarily break in time series and the development cost increases because new applications and software are needed. However working under BA outweighs these disadvantages because of the;

- Centralization of designing activities and documentation of statistical processes
- Control of statistical processes by joining similar statistical activities into one "base"
- Control of statistical processes and improvement of coherence in processes and activities by working with "steady states" and "steady stages".
- Unambiguous use of statistical concepts, the most important here to mention of course the use of the statistical units in the SBR
- The processing and maintenance costs will reduce as well as the administrative burden

2.3 The chain of Economic Statistics at SN and its main characteristics

The chain described in this paper comprises the compilation of the STS, SBS, QuarterlyNationalAccounts (QNA) and NA statistics. In order to move from the stovepipe approach for STS and SBS to this integrated chain approach SN had to meet some critical conditions and apply important changes. Additionally due to the influence of Dutch policy, SN has the obligation to continuously explore the use of available administrative data sources before using surveys. Because of the demand for lower administrative burden SN thus also had to redesign the production processes in a way that better linkage with administrative sources and the subsequent better use of this administrative data was achieved. Therefore the statistical Business Register has been placed in a central backbone role and is based on a new version of the statistical unit. In this way the SBR forms the basis for standardization and zips the information from administrative or other sources to the statistical units. Surveys are only used for the largest units; quarterly VAT data is used for the small and medium enterprises. Monthly estimates are based on a model that extrapolates quarterly data using monthly survey data for the largest enterprises.

By designing a process with well-defined intermediate data sets and in which every step ads value also redundancy in data processing (over processing) is prevented. The processes themselves were streamlined by enabling better use of reference data from other processes and more extensive use of automatic editing routines. Interactive editing is only performed after a top down analysis. This

development largely contributed to efficiency gains enabling SN to cope with the imposed budget cuts and at the same time to allocate resources for innovation.
When thinking about examples of problems that can be encountered in the chain of economic statistics the list may seem endlessly (for examples see Annex 1). A bothersome quality issue for SN has been in the area of 'coherence of variables within one series of statistics' resulting in too large differences between the GDP releases. In fact we suffered from a structural underestimation of quarterly GDP growth. This had mainly to do with quality issues regarding the outcomes of the underlying processes. Due to the reduction of the sample sizes (because of reduction of administrative burden) higher margins occurred. This quality issue could be overcome by using VAT data in addition to survey data but could not have been solved without applying more coherence in methodology and applying STS-turnover in the compilation of SBS statistics. Methodological coherence was reached by processing consistent economic data on the 350 largest enterprises (see §5.2), quarterly turnover data is based on practically complete coverage and used for growth estimates of STS, QuarterlyNationalAccounts and first annual NA as well as for level estimates for SBS and final annual growth estimates of NA. The monthly turnover is aligned with quarterly estimates:

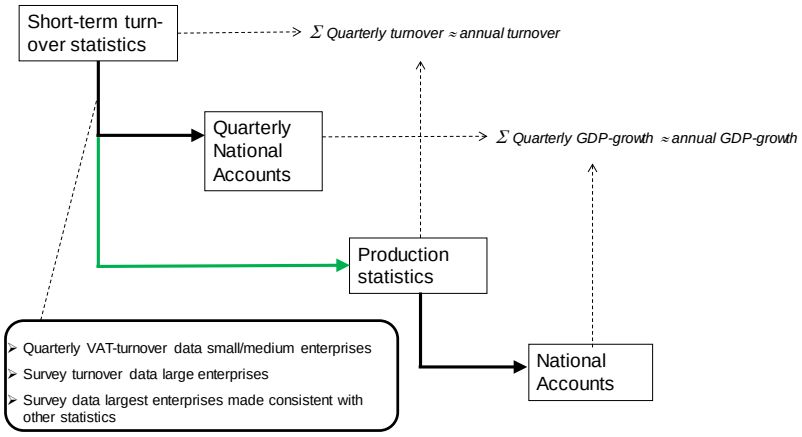

Figure 2: The chain of Economic Statistics in a nutshell

However, a redesigned process and methodological coherence alone are not sufficient to operate and maintain the chain of Economic Statistics. Therefore also chain management procedures were established; there exist detailed and binding agreements on deliveries from all chain members within in the chain. They are like service level agreements between all parties. Very crucial and subject of this paper is that the use of the SBR for all statistics involved is compulsory. Changes concerning statistical units and their characteristics are always done in a coordinated approach because they may affect other statistics. This also holds for handling changes in the VAT registers or when dealing with incidents. Detailed procedures (kept in a handbook that is evaluated and updated yearly) for handling these changes are used. And last but not least, the responsibility for monitoring the chain and executing decisions exceeding the individual statistics is delegated to the chain manager and the supporting chain management group, who have the lead.

2.4 Chain management; main benefits

To compile statistics using an integrated system is quite demanding; the whole chain of statistics, including all interdependencies should be transparent and for each of the relevant statistical products it must be known which population and statistical concepts it describes, how the data are compiled and how the various statistical products are used in the chain. Additionally for the whole system it must be known at what point in time which input is needed by whom, how much time each process takes, and how the various output can be combined. In an ideal world a 'happy flow process' could be designed and managers could oversee the compilation of statistics. However, external requirements change all the time, there is external pressure to reduce administrative burden, to improve quality and to increase efficiency and timeliness. Does this mean that achieving an integrated system is a mission impossible? No, it is our firm believe that working under architecture, integrating processes, assigning

the SBR a central (backbone) role and implementing chain management is beneficial and rewarding. In fact the main benefits are achieved on the same issues where we experience the most pressure; The structural underestimation of quarterly GDP was solved. Actually in general SN now suffers from less and also smaller revisions and thus an improved quality. Also there is more methodological coherence and consistency between statistics. Because of the increased re-use of data between processes and a better use of administrative sources in general led to a lower response burden. Also from the financial point of view benefits are achieved since redundancy in data processing is prevented and issues are solved in joint consideration and at that place in the chain where it is most cost-efficient for the whole chain.

3. SBR as Backbone in the Chain of Economic Statistics

3.1 SBR positioning and role

The place and role of the Statistical Business Register are very important to serve and support statistical processes. The Dutch SBR is part of the economic-statistics-system and has evolved to a business register system with subsystems. By the obligatory use of the SBR for the chain of economic statistics the SBR is *the* basic infrastructural backbone in this chain. The chain of business economic statistics as depicted in Fig. 2 comprises several main processes and flows between them. Besides the process Business Register, also the processes Consistency of the largest enterprises, Direct estimates turnover (STS), production statistics (SBS), Integration Quarterly National Accounts and Integration of National Accounts are distinguished. Figure 3 shows the main constituents as well as the main flows from each system towards the others.

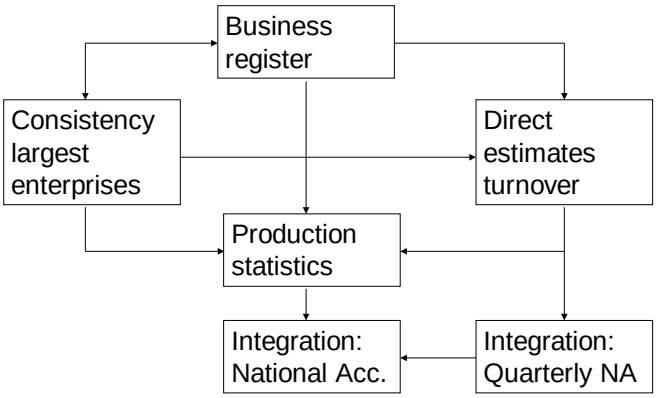

Figure 3: Main flows in the chain, central role for the SBR

In order to satisfy the requirement as structural backbone the SBR has to be the central place where statistical units and their main characteristics are stored and maintained. The Dutch SBR contains therefore the identifying and structural data on businesses at the level of three types of Statistical Units; the Enterprise Groups (EG), Enterprises (Ent) and Local Enterprise Units (LEU). The mandatory utilization of these Statistical Units and the frames derived from the backbone are in turn a precondition for coordination of the statistical output of the chain. And last but not least, the SBR has been expanded and should also facilitate the linkage of administrative units and their information to statistical units. The SBR thus fulfills the role of both:

- Population frame and;
- Statistical frame and as well as of;
- Coupling frame.

As a consequence the main uses of the SBR are also diversified as shown in;

Table 1: The main uses of the SBR in relation to its different roles.

1. Business survey frame	• Stratified sampling (selection of units and characteristics) • Survey dispatch (names/addresses) and tool for survey management and control • Instrument facilitating the use of data in administrative registers for statistics (incl. standardization and normalization). • Frame for control of administrative burden control and monitoring
2. Frame for coordination statistics	• Basis reference information • Comparability in and over time • Comparability between (sub)populations • Control of continuity
3. Statistical analysis	• Economic demography • Small area analysis
4. Internationally	• Comparability of statistics of different countries, contemporary and in time series • Analysis of globalization

The many uses of the SBR are indicative for the variety in the most important stakeholders of the SBR as well. Which in fact are all producers of economic statistics. Thus aside the constituents of the economic chain also;

- Statistic of Balance Sheets (Financial statistics)
- Foreign Affiliate Statistics (FATS)
- Foreign Direct Investment statistics (FDI)
- Statistics on (un)employment and wages
- Price statistics
- International trade

3.2 Tuning of statistical units and frames.

The backbone role of the SBR, its role and many uses in relation to the many stakeholders inside the economic chain as well as in parallel chains (see 3.1) demand a scrupulous tuning of the statistical units and frames. Therefore a substantial part of the agreements and procedures in chain management is comprised by a policy how to deal with real and non-real mutations concerning statistical units;

- Like a potential difference in the actual (i.e. real-time) versus the coordinated (as part of the statistical frame) NACE code.
- The coordinated NACE and Size class are fixed during year t on behalf of synchronizing year and monthly/quarterly figures.
- Corrections and mutations on coordinated NACE and size class are (only) executed at the end of the year/change of year i.e. upon validation of the coordinated characteristics for the upcoming year.
- For large and/or troublesome mistakes/faults that cannot wait till the end of the year a frame error procedure is in place for retroactive corrections.

Additionally a decision tree for coding units has been developed and also tuning processes for complex individual cases are in place.

Up to now the focus has been on describing the redesigned system of business economics and NA to an integrated chain model supported by a BA. Also the necessary conditions are met, like a central role of the SBR, with statistical units facilitating maximal use of administrative sources, and there is agreement on maintenance procedures for changes in statistical units as well as other chain management procedures that are in place.

But how about quality issues and what can be the distinct role of the SBR?

4. Chain management and the governance of quality management

The approach of chain management governs quality management by working according to the principle to act locally but always from the perspective of the chain. This means that the actions to govern or improve quality are taking place where they are the most effective. In the next paragraphs this is illustrated by examples concerning the quality of the SBR itself (§4.1) and SBR related events considered in relation to the statistical processes (§4.2).

4.1 Governance of SBR and data source quality management

As shown in Fig. 3 all processes are linked together and subsequently improving the quality of one process and outcome will affect the other processes and outcomes as well. Thereby it is also clear that improving the quality at the beginning of the chain is more rewarding and beneficial than improving it at the end of the chain for example by correcting only a statistical outcome without correcting the source of the error. When a correction is or cannot be used by other processes then these processes will not be able to benefit from the quality improvement. With the SBR at the beginning of the chain of economic statistics it is the usual suspect to govern quality. This is in turn is also true for data sources feeding into the next process.

4.1.1 Procedures to maintain SBR

The SBR maintenance strategy [Ref 2] comprises of four pillars. First there is the use of the Trade Register (TR) to daily update the statistical units. The TR is the administrative source for the Dutch SBR which is kept at the Chamber of Commerce (CoC). Based on algorithms information on the legal units of the TR is used to delineate Statistical Units. On a daily basis SN receives updates of all the new information that the CoC has processed in their TR. At SN we only keep the most actual situation of this TR which is stored in a so called cache register, cache because all information is cached-refreshed- on a daily basis. The very first quality check is on a daily basis on the information present in the cache to check if it meets requirements agreed upon in advance. Anomalies in the source data present in the cache are shown in a reporting service. If a detected deviation is severe then the content of the cache of that day is not processed further into the SBR. Instead, the CoC is contacted and requested to correct the error first in their source i.e. the TR. Illustrating that chain management can even extent to external partners and that correcting data may go as far back as the source of the SBR even.

Obviously the characteristics of the legal units used to derive the Statistical Units need to be complete, recent and up to date. It is also true that the TR does not always hold recent and properly updated proof of any economic activity of a legal unit. Therefore to derive the correct active legal units also

other sources like the number of employees according to the Social Security Register is used in combination with the turnover from the VAT register. Similarly information from the Dutch Central Bank is used in order to exclude so called Special Purpose Entities when the Enterprises are delineated.

Secondly, the Dutch SBR itself is a live register that is daily updated. However in order to retrieve for our users statistical population frames that are topical, reproducible in time and with all changes caused in the source being traceable, so called snapshots are taken (monthly). A snapshot contains all statistical units and their characteristics and relations that are unchanged, those that are changed and need to be deleted and units, characteristics and relations that need to be added. This snapshot is therefore a preliminary frame. The macro validator checks the changes compared to the previous final frame on a macro level for enterprise groups with more than 20 employees. Anomalies may be corrected by a profiler (see §4.2). The macro validator also consults with the statisticians and informs them about the changes that are about to happen. After consultation and correction the macro validator approves the information in the preliminary frame which becomes then frozen and a final frame that is uploaded in the output area. The macro validator is therefore the guardian for the quality of the final frame.

Thirdly there is the feedback information from statistical surveys. In the same way as described for additional sources information from surveys (information on financing large enterprise groups) is processed in the SBR.

The fourth pillar is the process of profiling. Profiling is undertaken for the top 2100 most important Enterprise groups, where the statistical units are delineated manually from an Enterprise group. It is a tailor made method to analyse the legal, operational and accounting structure of large businesses (Enterprise group) in order to establish the statistical units within that group, their links, and the most efficient structures for the collection of statistical data. This profiling itself is also divided in three approaches, see § 4.2.1 for a more detailed explanation.

4.1.2 Coordination of adjustments of statistical estimates due to errors in SBR populations

Since coordination and cooperation are key in the régime of chain management frame errors are corrected according to an agreed upon procedure applying to *all* members of the chain. A frame error can be detected at various places in the statistical production process as well as at the SBR, it is however the statistician of the affected economic variable who needs to approve the correction. For coordinated variables like NACE code and size class the corrections are always corrected in the SBR.

4.1.3 Prevention of inconsistencies

The motto “better to prevent than to cure” is also true for inconsistencies between statistics, especially when these sources are feeding the NA process. Preventing the occurrence of inconsistencies is very rewarding. AT SN this is done for around 350 of the largest and most complex enterprise groups who together comprise almost 75 % of the total turnover. These enterprises are dealt with in a separate process where the relations with the EG are maintained, the structures are profiled in the SBR and where also micro data processing takes place with the aim to deliver consistent micro data for direct use in statistics including National Accounts (ref [3]). Also this approach for the large enterprises is an example of achieving high quality as soon as possible in the production process. In this case by

looking at 14[1] variables across 10[2] different statistics (incl. Tax data) inconsistencies are detected and corrected early in the process and at a micro level. In the classical approach the consistency work is done at the end of the process at the integration of NA, making NA responsible for the consistency work. With this process we see the consistency work moving en being concentrated on the first part of the statistical process. See the Annexes Fig A for an animated explanation.

4.2 Quality management of the whole chain; tuning SBR maintenance with statistical processes

Data quality management on SBR or sources is not a standalone process. It needs to be tuned with the statistical production processes in the chain. This tuning is perhaps the most difficult and complex task to achieve and comprises many diversified activities. For example the focus on those large and complex enterprise groups which affect economic statistics the most (§4.2.1). The derivation of subgroups with distinct maintenance strategies is one example of dealing with this challenge. Other examples are highlighted in (§4.2.2).

4.2.1 Defining different maintenance groups of SU in the SBR

One of the key features in the maintenance strategy is to focus more on those large and complex enterprise groups which affect economic statistics the most and at the same time taking into account restricted personnel costs. Under the orchestration of chain management and in close cooperation between BR specialists, profilers and statisticians SN distinguishes between three different subpopulations of the SBR all with a different maintenance strategy.

The first is the maintenance of the so called ConGO Enterprise Groups (350 large and complex EG's) by active profiling and a full manual update on a regular basis including compilation of statistics in a consistent way (see §4.1.3). The next 1750 large enterprise groups (Top-X) are also treated by active profiling however without the statistical consistency checks. In the method of active profiling the profilers are keen on announcements in media about mergers, bankruptcies, restructurings, acquisition and selling of legal units within the Enterprise Groups in their panel. The profilers will keep track with the changes in those Enterprise during the entire operation. The intensity of profiling depends on the size/number of changes within in an Enterprise Group. Finally there are manual updates by BR staff on a non-regular basis, the so called reactive "profiling", based on signals from the system (fully automated procedures using source data, coming in on a daily basis) or because of special reason These comprise all other units (non-Top-X, 1.5 mln EG's, mainly small or medium sized).

The Enterprise groups relevant enough to treat by pro-active manual profiling are selected by an algorithm. This algorithm is called the CSI factor where CSI stands for Complexity and Statistical Impact. The CSI algorithm ranks the Dutch EGs according to their relevance for statistics and their size and complexity. It uses information on: the Number of legal units belonging to the EG, Control relationships, both total number and the number of layers in cluster of control, Locations, Number of Nace sections, Number of fiscal units, the Number of persons employed (size class), Balance Sheet Total, contribution of EG to the NACE (in terms of employment), contribution of LeU to the NACE (in terms of employment).

The CSI-algorithm is executed 14 times a year. One day before the finalization of the monthly frozen frame, the CSI-score for a (small) set of new EGs is calculated. This set consists of new EGs that are

[1] Depreciation, Earnings, Employee benefit costs, Wages, Export, Import, Intermediate consump-tion, Investments, Net financial result, Net sales and other operating revenues, Number of employ-yees, Production, R&D expenses, Turnover.

[2] Finances of enterprise groups, SBS, STS, Prodcom, Investments, International trade (Goods and Services), R&D, Statistics on employment and Wages, Corporate Tax and Value Added Tax

the result of demographic changes in the current Top-X population. The new EGs with CSI-scores high enough to be included in the profiling population (Top-X) will be included from the next month on. Once a year in December the Top-X population is being refreshed by taking into account also CSI-scores of the EGs in the non-TOP-X. Groups which were too small and/or not important enough last year, may have become larger or more important next year. The algorithm is also executed in October in advance serving as a first indication or even to start with profiling to get a good delineation from January next year on.

Additionally other criteria are taken into account defining the final list of Top-X EGs: Significant issues with iintegration of an EG in National Accounts, The EG being critical for quality of the foreign trade statistics, there are significant issues in legal unit structure of groups and stability issues with CSI-score leading to frequent changes from Top-X to non-Top-X and vice versa.

4.2.2 Other examples of control SBR quality tuned with statistical processes.
Somehow communication always seems to fall short; therefore the consultation with statistician's needs extra attention at all times. The most important moment however is just before the final frame is derived. The process can be further improved when both statistician as well as the macro-validator uses the same analytics tool on the same data.
Because of the reproducibility in time of frames and the traceability of changes within the frame the changes that affect the structure or characteristics of an SU (i.e. events) are stored. For the same reason of reproducibility and traceability the frozen frames are derived from the SBR live environment. During the approval process by the macro-validator also application dates are assigned to the units in that frame. SN uses 'application dates' instead of 'event dates'. The application date implies the validity of the SU in the frozen frame for the reference period and for all business statistics involved. The benefits of assessing the quality of the data-source have been addressed, however to be able to do so shared reporting facilities should be in place. With the growing extent of the use of administrative and other data this becomes even more pregnant. The complexity of governing quality management in relation to new developments is illustrated in the next paragraph.

5. Challenge; A mixed mode multisource approach for compiling economic statistics

5.1 The Challenge explained
The mixed mode multisource approach is depicted in Fig.4. In the area of business statistics SN is aiming at extending the use of the subgroups, not only for maintenance reasons but also for data compilation. The large and complex enterprises will be dealt with separately as is the case already (ConGo and Top-X profiling population). Also, their observation will remain in the mode of direct observation using questionnaires. For small and medium size enterprises registers on the requested variables are increasingly available and their use will grow to a bigger extent. Also other external sources and Big Data sources belong to the possibilities.
For the financial variables these sources are less suitable. However their importance for making statistics is large and therefore new developments are undertaken in order to gain direct access to business administrations of medium sized enterprises. For the remainder subgroup of enterprises limited observation of missing data probably will be needed. This approach implies that various sources (multisource) and various modes (mixed mode) altogether are being the input for making economic statistics.

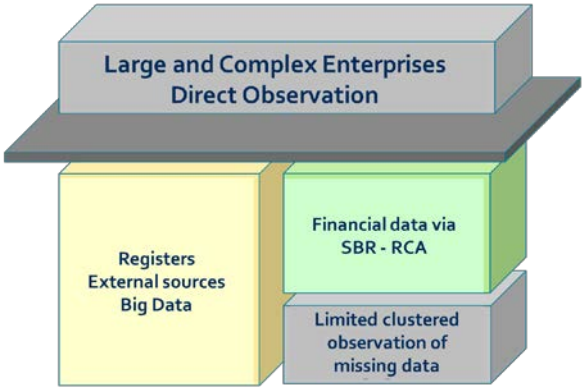

Figure 4: mixed mode multisource approach for the compilation of economic statistics.

5.2 Enlarged role of administrative sources

An advantage of administrative data is that these data are tailored to the purpose of the register. However, the main purpose of an administrative register is not designed for statistical purposes. Therefore statisticians always have to deal with what is available. Also registers are not the answer to the main sources of non-sampling errors like item non response, measurement- and coverage errors. Administrative data needs to be transformed and linking register units to statistical units stays a tricky business with many n:m relations that have to be dealt with. Another threat is the dependency on these sources and the impact on SBR and statistical outcome because of serious changes that may occur in these datasets. How does SN manage this dependence? First and foremost, there is the statutory regulation (2004) for free access to registers covering all registers of the national government, government agencies and municipalities. Secondly SN keeps good and sustainable relationships with the management and employees of registers as well as with the departments that are political responsible for the registers. Thus the good practice of chain management is extended to outside the office and includes account management on strategic, tactical and operational levels.

5.3 Standard Business Reporting and Standard Chart of Accounts

A new way of obtaining data from small and medium enterprises without causing administrative burden is by gaining direct access to their business administration. More and more of these administrative processes are supported by or even performed with software implying that more and more financial data are digitally available.

This is however not an easy task. There are many partners involved, the first bottleneck being the primary users of the data of the business administration like the Chamber of Commerce, Banks (credit supply), shareholders and investors (capital supply) and fiscal authorities (tax collections). They are the key players when it comes to determining the information to be held in the business administration. The biggest challenge will be the standardization which means the development of a reference chart of accounts in order to simplify the data translations within the whole financial administrative domain. In order to realize this SN is involved in the network of contacts with software providers, accountants firms and large users around XBRL/Standard Business Reporting. Also there is an initiative with government and market partners for the development of a Standardised Reference Chart of Accounts which has proven to be successful. More and more software packages support now the Reference Chart of Accounts. Standardisation is the key – but it is not only in our hands.

References

[1] van Delden, A., Lammertsma, A., van de Ven, P. (2009): Chain management in statistics: best practices, SN discussion paper 090403.

[2] Konen, R. (2014): Maintenance strategy of the Dutch SBR, 24th Meeting of the Wiesbaden Group on Business Registers, Vienna, September 15th.

[3] Vennix, K. (2012): The treatment of large enterprise groups within Statistics Netherlands, ICES IV, Montreal, June 13th 2012.

Annexes

Table 1: Examples of problems in the chain of economic statistics

No	Areas of attention	Example of a main problem
		Statistical Output
1	Relevance	Output has not enough detail
2	Accuracy	Large differences between releases
3	Coherence of variables within one series of statistics	Use of different variable definitions use of different methods
4	Comparability over time and between domains	No harmonisation of the classifications
5	Consistency between statistics	Uncoordinated corrections in structures of businesses
6	Timeliness	Processing time is too long
7	Punctuality	No production time schedule
8	Clarity	No information metadata is given to the users of the output
9	Completeness	Input sources are not reliable
10	Confidentiality	No coordinated confidentiality policy
		Process
11	Efficiency	Too many steps are done manually
12	Complexity	Actors cannot judge the output because they don't understand the methodology
13	Flexibility	Changes in input sources lead to a full redesign of the processing steps
14	Transparency	Changes in one statistical series have unexpected effects for other series
		Actors within the Chain
15	Openness	Mistakes are reported too late

Figure A.

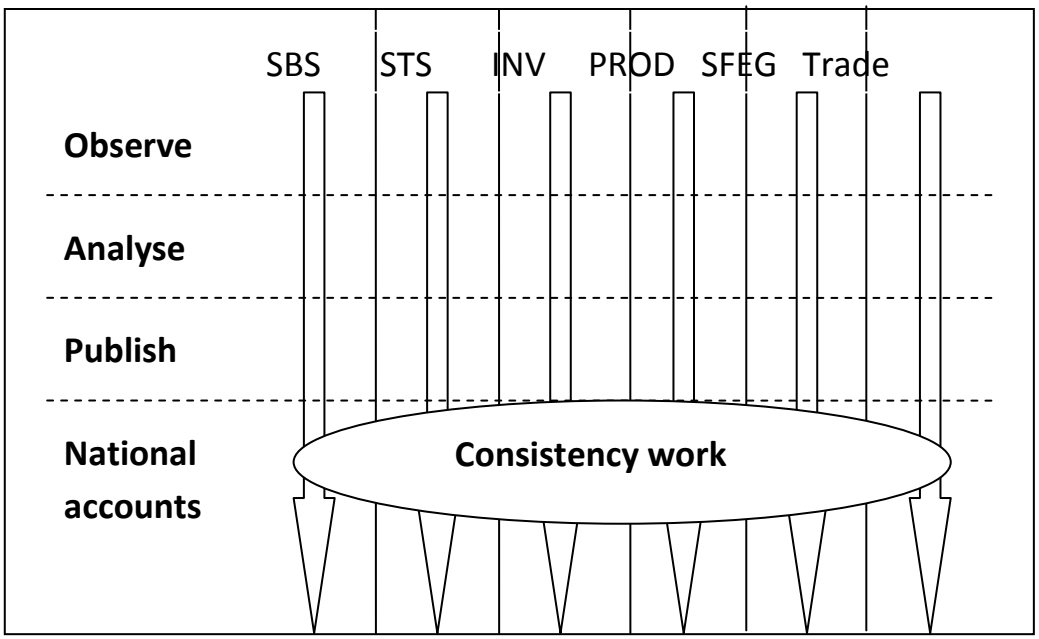

Figure A1: the place of consistency work in the statistical process until 2010. Vertically the statistical process from data collection to delivery to the national accounts is represented. Each "stove-pipe" represents an isolated individual statistics such as SBS, STS, etc.

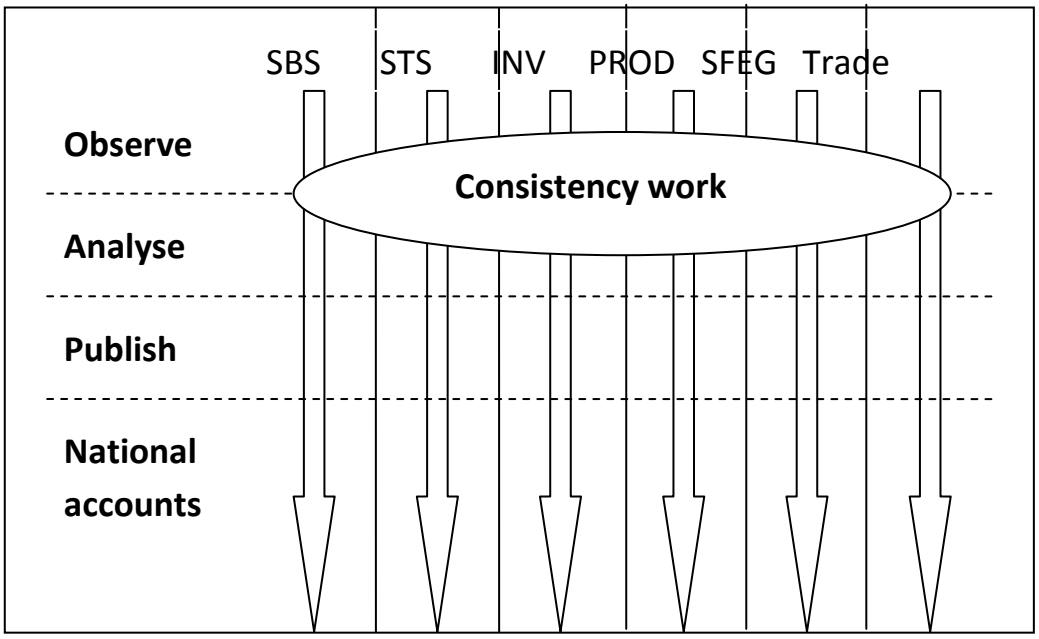

Figure A2: the current place of the unit for large enterprise groups in the statistical process