

Adaptation of the SBR to changing external conditions

Context Wiesbaden Group, 2008 meeting in Paris, session 4

Author Søren Schiønning ANDERSEN, head of Division, Statistics Denmark

Abstract: This paper describes, firstly, Statistics Denmark's (SD's) Statistical Business Register (SBR) and its relation with the Central Administrative Business Register (CABR). Secondly, it describes recent external challenges to the SBR system and how SD has tried to cope with these challenges. Finally, it describes for each challenge how the ability to fulfil the basic objectives of the SBR has been affected.

The challenges dealt with include 1) the change of a key administrative source; 2) a change in the fundamental business model and redistribution of responsibilities in an integrated system; 3) how we needed to change the SBR in the light of these two challenges; 4) how to deal with ad-hoc challenges from politically strong users of the CABR in an integrated system; 5) challenges from new supra-national legal requirements and 6) the challenge to increase productivity.

Among the conclusions are that in a system where the SBR is highly integrated with ABRs it is necessary, but not sufficient, to have a thorough technical understanding of the sources. It is also necessary to be "politically" pro-active in order to maintain the SBR as a basis for quality statistics.

1. What is the purpose of this paper?

The purpose of this paper is, firstly, to describe the interface and integration between SD's SBR and its main sources, especially the CABR which is the main source for the SBR. Against that background, the purpose is, secondly, to describe recent external challenges to our SBR system; how SD has tried to cope with these challenges and, finally, how these changes have affected our ability to fulfil the basic objectives of the SBR (i.e. *full coverage, high quality, usage* in all business surveys) as well as the *costs* of running the SBR.

2. What are the main characteristics of Danish registers?

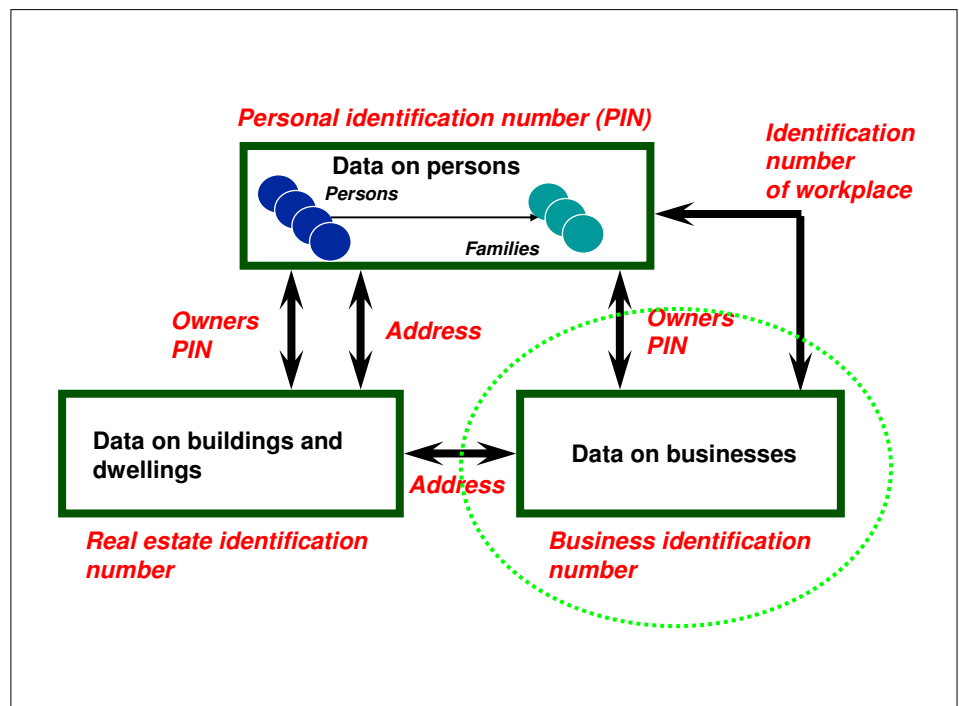
Essentially, SD's SBR is not *in itself* much different from SBRs in comparable countries. It is designed to comprise and deliver the units and variables, which the statisticians need in order to produce their statistics. For the most part, this is prescribed by the EU Regulation on Business Registers for statistical purposes [1], the EU Regulation on statistical units [2] and by the statistical EU regulations, UN recommendations and manuals etc. in the various statistical domains.

What makes the Danish SBR somewhat different from many countries is the *degree of coordination and technical integration with other register systems* – with the pros, cons and dependencies following from that.

The core elements of the Danish (Business) Register strategy are:

- The basis for establishing and running central *administrative* registers, including the CABR, is laid down in *separate* legislation. (The production of statistics on the basis of administrative records and maintenance of statistical satellite registers is done on the basis of the law on SD [3]).
- Unique identifiers of legal units (LU) and legal local units (LLU) used throughout the public administration and also used increasingly in the private sector.
- Strictly *coordinated definitions and classifications* of LUs and LLUs among the public authorities responsible for registration of business entities.
- Responsibility for *creating, updating and ceasing LUs* is unambiguously distributed among the public authorities involved: Each unit has one and only one *responsible data provider*. The responsible data provider is determined by the legal form (type of ownership). (A responsible data provider is not assigned to units at local level among those authorities having an interest in local units.)
- Responsibility for *creating, updating and ceasing LLUs* is allocated as close to the source as possible (ultimately through self-registration via internet by the entrepreneur).
- *Real-time on-line integration* between the CABR and the satellite registers of the other authorities involved as data providers, including SD's SBR.
- Maximum (re)use of data in the CABR, which is largely *financed via sales* of data and services. In order to minimise burden on businesses, usage by public authorities is required by law – usage by private business is encouraged via marketing.
- Coordination and integration with the Central Register on *physical/natural Persons* (who have unique identifiers) and *buildings and dwellings* (also having unique identifiers). The link is made via the *addresses* (also having unique identifiers), cf. figure 1 below.

Figure 1:
The three basic registers



Within this overall context, SD strives to have the statistical needs reflected as much as possible in the data and business processes of the administrative data suppliers. For the SBR this was initially facilitated by the fact that SD was in fact made responsible for establishing the CABR in the mid 1990's because of our experience, firstly, in integrating different data sources; secondly, in working with units and data at both enterprise and establishment level and, finally, in selling basic non-confidential data (tailor made address lists) on the market.

2.1. Data model for the Danish SBR

The SBR "embraces" the CABR

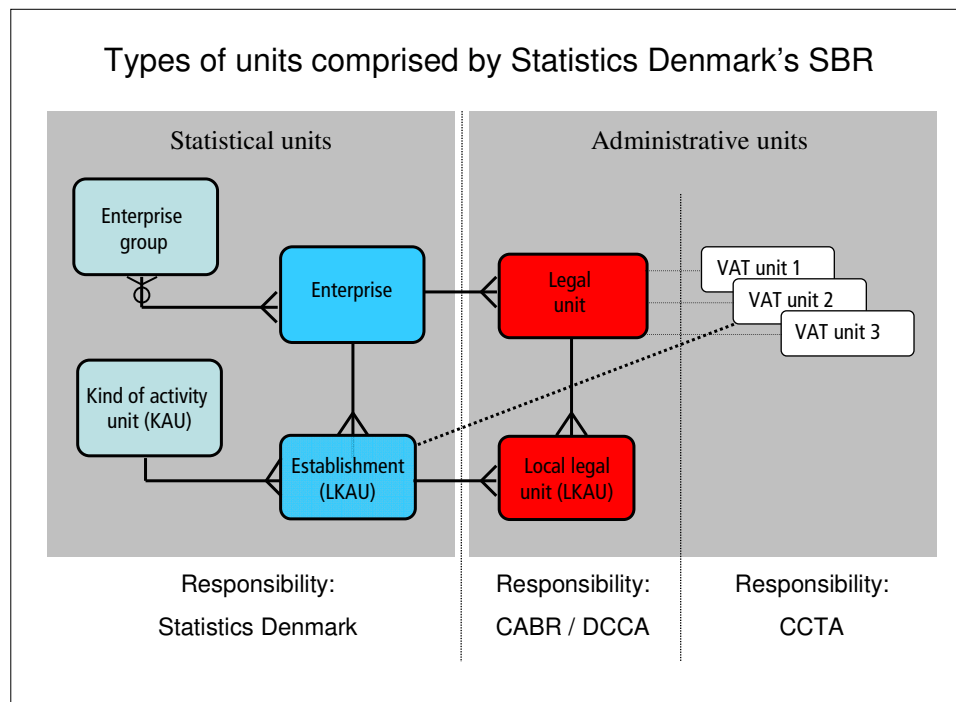
The CABR – see section 2.2 below – does not contain all the necessary units and data needed for statistical purposes. SD therefore operates a SBR. In practice the SBR "embraces" the entire CABR (all units and all data) and then adds and modifies units and data in order to have the information necessary for fulfilling its role in the production of business statistics. Also input/feedback from the statistical divisions is used in the process of maintaining the SBR.

Most statistical units are "transparent" i.e. identical to the administrative units

Having said this it is, however, important to note that the CABR is by far the most important source and that in almost all instances one LU is equal to one enterprise; one LLU is equal to one establishment (LKAU) – and data for the administrative units and statistical units are identical (the statistical unit is "transparent" – all data comes from the administrative unit).

The data model for the SBR is shown in figure 2 below, illustrating how our SBR comprises both LUs and LLUs from the CABR, the VAT units from the Tax authority's (CCTAs) register (the right side of the figure) as well as the statistical units used by our colleagues in the statistical divisions (the left side of the figure).

Figure 2:
Data model for the SBR



The main characteristics and objectives behind the business model for our SBR are:

- Unambiguous distinction between the different types of unit.
- Separation from the donor registers, including separate identifiers for statistical units (but links to unique identifiers, of course).
- Adherence to the identity over time for establishments.
- Registration of the dates of all events, e.g. commencement of activities, changes, corrections (“double history”) and sources of the updates.
- On-line updating from the main source (the CABR).
- Frequent (batch) data supplies from other administrative sources (most notably the CCTA (re e.g. activity status) and statistical divisions (re e.g. employment and turnover data)).
- Decentralised updating from (a few) statistical divisions.

*300.000 enterprises
of statistical interest*

The SBR has data on approx. 570.000 enterprises with a total of approx. 620.000 LKAUs. However, due to the very low threshold for VAT registration in Denmark – and thereby registration in the CABR – and the fact that many inactive enterprises are not ceased in the CABR, only about 300.000 enterprises has economic activity corresponding to more than one half-time employee, cf. the EU recommendations for registration in SBRs. Currently, the SBR is used – directly or indirectly – in 62 surveys in SD.

2.2. Sources for the Danish CABR – and thereby the SBR

One CABR

The new CABR (called CVR) started up in October 1999 where the data from the old registers in DCCA, CCTA and SD were merged (matched and prioritised). Since then Denmark has had only one CABR with a unique identification of LUs (8-digit CVR number) and associated LLUs (LKAUs with 10-digit P numbers). The set up and use of the register is regulated by the Danish Act on the Central Business Register [4].

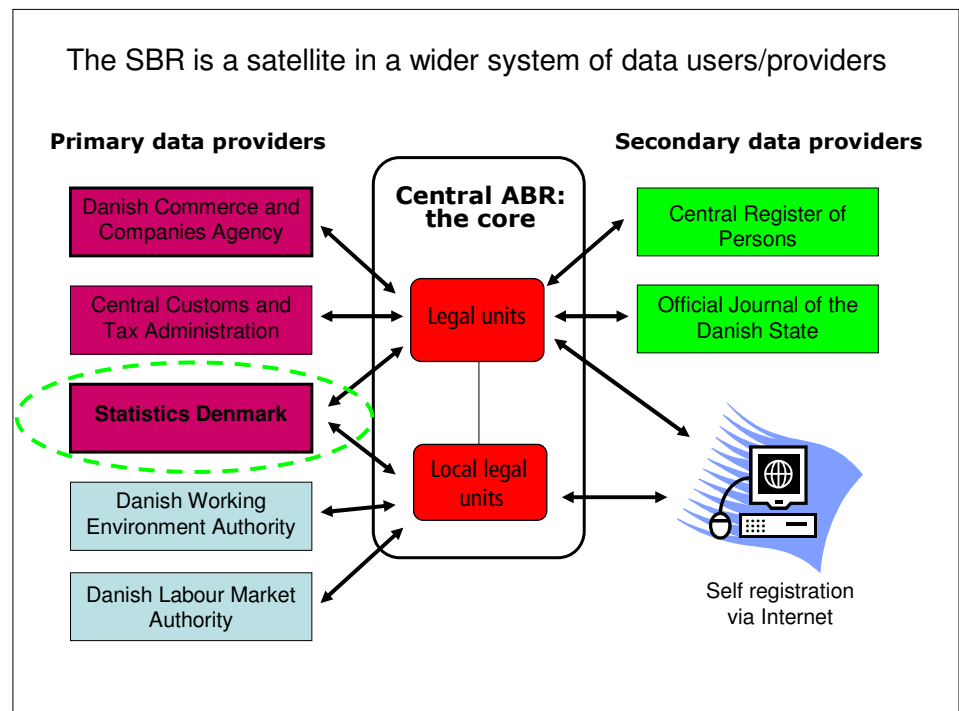
*Data are public – and
usage is mandatory*

The objective of the CABR is to store basic data on LUs and associated LKAUs, to undertake an unambiguous numbering of these units and to provide basic data for public authorities, institutions and private users. All public authorities are *obliged to use the CVR number* when corresponding with LUs, and the LUs are *expected to use the CVR number* when corresponding with public authorities and with other LUs.

*Primary data suppliers
from different ministries*

According to the act the Ministry of Economic and Business Affairs is responsible for the operation of the register in co-operation with the Ministry of Labour and the Ministry of Taxation. Within these three ministries, the Danish Commerce and Companies Agency (DCCA), the CCTA and SD are the primary data suppliers to the CABR, as it appears in figure 3 below. The units in the core (in red) are identical to the units (also in red) on figure 2.

Figure 3:
Main sources for the SBR –
via the CABR



- The DCCA is currently responsible for the running of the CABR – the core of the system. Furthermore, they do not “only” register public limited companies, but also the initial registration of sole proprietors – in cooperation with the CCTA. The DCCA is responsible data provider for LUs established according to the laws on limited companies.
- The CCTA is responsible data provider for the remaining LUs in the private sector.
- SD is the responsible data provider for all public sector LUs. This role is conducted in close cooperation with the Danish Agency for Governmental Management (Ministry of Finance), the 5 regions and the 99 municipalities. SD is also an important provider of LKAUs, and SD provides employment figures for all units. Currently, we provide quarterly figures for LUs and annual figures for local units. Finally, SD can assume responsibility for activity codes, i.e. “lock” activity codes and thereby preventing other data providers from changing the code(s) in a given period of time.
- The Danish Working Environment Authority (DWEA) is mainly a *user* of data in relation to their planning and conduction of inspections of the occupational health and safety at the work places. However, they also create a limited number of LLUs and provide updates to others.
- The Danish Labour Market Authority participates mainly as data user. Data are used in their work related to the allocation of labour to the actual work places with vacant jobs.

Secondary data providers Another group of sources are the so-called *secondary data providers*. They include:

- The Central Register of (physical/natural) Persons (CPR). The CPR provides data on sole proprietors (owners).
- The Official Journal of the Danish State, which provides data on LUs subject to bankruptcy or liquidation.

- Finally, the enterprises themselves are becoming more and more important as data providers. An increasing number of registration processes (both creations, updates and cessations) can be conducted via the internet by the entrepreneur or manager using a digital signature. Some of these processes are subject to manual checking by clerks or lawyers in the above-mentioned government bodies, but increasingly such processes are automated, and the ambition of the Danish Government is to have all communication with enterprises digitalised from 2012. This poses huge challenges to the complex systems we have built up.

3. What are the current challenges to this system?

The framework conditions for the Danish SBR were relatively stable from 1999, when the CABR was put in place, and until 2006. Since then a number of challenges of different nature have necessitated substantial changes to the SBR system.

3.1. "e-Income" – a new source for employment data

<i>Background of e-Income</i>	As an integral part of the maintenance of the SBR, SD is linking every employee in Denmark to a LKAU. Since 1981 this has been done on an annual basis. The source for this process has up to now been the annual income register from the CCTA. However, in 2008 a new administrative income register ("e-Income") has been introduced by the CCTA, which also will be widely used by several other public authorities including SD.
<i>Data will have financial implication for tax payers</i>	Previously, another identifier of the LKAU where the employee works was collected annually for statistical purposes "only", but the new system will use data on LKAUs (including the identifiers from the CABR) to automatically calculate the eligible tax deduction of employees expenses for commuting between home address and workplace address (via the x,y coordinates) in addition to the administration of a number of labour market initiatives.
<i>Increase re-use of data in the public sector</i>	The wider aim is to increase the re-use of data in the public sector, to improve the accuracy and actuality of data and to enable the enterprises to report the data via the internet as soon as the data are generated by the enterprises (i.e. at least monthly - when salaries are paid). A whole series of monthly and annual reporting obligations from employers to the CCTA will be integrated and streamlined in the e-Income register.
<i>Reporting "one time and in one place"</i>	The e-Income register will contain information about wages and salaries, hours worked, pensions, social contributions etc. for each employed person and – for LUs with more than one associated LLU – the number from the CABR of the LLU where the employee works. The register shall ensure that information in this field, which are needed by public authorities, is reported only "one time and in one place". Monthly reporting became obligatory for all employers from 1 January 2008, although 2008 is considered an "introductory year" – for the CCTA (as well as for SD). The first test data from the e-Income register is expected to arrive in SD in the near future.
<i>Quarterly employment data for LKAUs</i>	The new register will enable us to link every employee to a LKAU on a quarterly – and perhaps ultimately monthly – basis. The information is used to calculate the number of employees for each LKAU in the SBR. Currently (with

the annual source), this involves a comprehensive validation process (about six man-years annually) conducted by the SBR team.

High quality is needed due to the wide usage The rationale for this relatively high input of resources is that it is critical that the quality of this information is as high as possible due to its importance for correct stratification, extraction of populations and samples and grossing-up the results for all statistics on LKAU and KAU level which are based on the SBR. In addition, it has considerable impact on business demography statistics and various labour market statistics.

Error checking is still needed – and it has positive side effects Although e-Income is expected to improve the accuracy compared to the old source it is still envisaged that comprehensive validation of the information will be needed by SD. Also because the error checking leads to many other corrections in the SBR, especially with regard to following establishments over time.

With the new frequency, format, content etc. we are required to completely re-design our IT application and working process. In return SD will be able to publish short term employment statistics on a much more detailed level. Additionally, this new source will enable the development or improvement of other statistical information in high demand.

Consequences for the SBR with regard to ... As mentioned above, we have yet to see the first test data, so the expected consequences are based on our contacts with the CCTA during the project.

... coverage With regard to *coverage* the e-Income register is expected to slightly improve the coverage of local units. However, the current coverage of LKAUs in the SBR is also considered to be close to 100 percent for all those with employed persons, so in that respect the improvement is rather limited.

... quality However, with regard to the *quality* of data, e-Income is expected to improve:

- The *timeliness* with which new LLUs are registered;
- The *frequency* and *timeliness* of our employment data at LKAU level - from annual data available at t+15 months to quarterly data at t+3 months;
- The *accuracy* of the dates of creations/cessations/take-overs, which could (hopefully) improve the accuracy of our business demography statistics;
- The *relevance* of the data, since they are expected to enable new statistical products (e.g. in the labour market domain) and improve existing products (e.g. the timeliness of the “threshold of significance” applied in the SBS, which again could improve the *coherence* with other statistics).

... usage Against that background, the *usage* of SBR is expected to increase. (At the same time, the SBR team will pursue a more coherent usage of the SBR in the production of statistics – via the use of “frozen” versions/extracts, meaning that all statistics using the SBR for a given reference period (except maybe some agricultural surveys) will base itself on the same version of the SBR.)

... initial costs The costs of *establishing* the new validation system is estimated to approximately one man-year, with half of it being used on the business side in analysing the new data source, designing the new process, specifying needs and requirements for the IT system, and project management. The other half is used for the IT development. The new system shall be operational by the end of the first quarter of 2009.

... running costs Taking into account the expected resources available to SD in the coming years, the costs of *running* the new system must be lower than in the current situation. Considering that essentially the same amount of information must be validated four times instead of one time during the year – and with fewer resources – means that the validation process must be fundamentally re-engineered and automated to a much larger degree than now.

... the change-over This is a huge challenge for us – also because the transition inevitably involves a “double year” – i.e. validation of annual data for 2008 from the (last year of the) old source and validation of monthly data for 2009 from the new source.

3.2. “Self-registration”

As we mentioned in section 2.2 above, creations, updates and cessations of the units in the SBR are to a sharply increasing extent made by the data subjects themselves via the internet.

Drivers for the change towards self-registration On 1 December 2006 the DCCA put the so-called “self-registration-solution” (“web-reg”) into operation. The introduction of “web-reg” was caused by a number of new legislative initiatives whereby data on local units in the CABR were given a key role – essentially making them into *legal* local units.

The most important of these initiatives was the e-Income register, cf. above, but also new *food safety* legislation about controls on local units selling fresh food and legislation about *financial transfers between government and enterprises* (the so-called “easy account”) were drivers in this process. Finally, the process was also driven by the need of the DCCA to cut costs in the CABR. One of the means to do this was to hand-over the responsibility for correct data to the data subjects themselves, i.e. making them “responsible data providers for their own local units” and thereby limiting the possibilities of e.g. SD to update data which we consider erroneous.

One could say that the underlying perception of maintaining basic data about businesses changed from something being relatively difficult (requiring expert skills and complex IT systems) to something relatively easy (which “everybody” can do via a simple internet application).

Overall aims fulfilled at the expense of SD’s needs In order for the “web-reg” system to be understandable and user friendly for the enterprises, some of the key characteristics in the CABR had to be changed - it needed to be much simpler. This simplification to fulfil the overall objectives was made at the expense of some the specific needs of SD – needs and functionality that require expert skills which “laymen” do not have.

Local units could no longer be followed over time Specifically, the most important change was that local units could no longer change its relation to a LU, meaning that local units could no longer be followed over time in the CABR. When a local unit is taken over by another LU, the old local unit is closed down and a new local unit is being created – irrespective of whether the location and activity of the local unit remains unchanged.

Consequences for the SBR with regard to ... Although SD acknowledges the necessity of this change in the CABR from an overall societal point of view, there is no doubt that this change to the perception and business model behind the CABR was a substantial – and very re-

source demanding – challenge to SD, where e.g. the need to follow establishments over time remains.

... coverage So far there have been significant technical challenges for the DCCA in making the “web-reg” application function correctly. This has often lead users to think their registrations of new local units have not been accepted by the system because they have not received an identification number back from the system. Consequently, they have tried again resulting in duplicates (i.e. over-coverage) in the CABR and subsequently in the SBR.

... quality Potentially, the quality of the data should improve when the number of data providers increase from five governmental authorities to approximately 300.000 employers having data at first hand – especially if/when the correctness of data can have legal and/or financial implications for the enterprises. With regard to the *timeliness* of the data on creations of new local units this holds true, cf. above regarding the e-Income register.

However, with regard to the *accuracy* of dates on demographic events, the expectations have yet to be fulfilled. This is among others because the registered data is – by default – the date of registration when using “web-reg”. Since the enterprises themselves have no interest (or awareness) of the need for correct historical data in SD’s SBR, they very rarely use the technical possibility to let the registration take effect from the real date (if it differs from the date of registration).

Secondly, the technical possibility to make updates via “web-reg” also goes for activity coding. This can potentially enable a significant improvement to quality, since SD by no means can keep track of activity codes for all units. On the other hand, it also enables enterprises – with or without willing – to register an incorrect code and thereby deteriorating the quality. We will get back to this in section 3.4.

... usage The usage of the SBR is not directly affected by self-registration – it has not meant that additional surveys could use the SBR, nor has it meant that surveys have stopped using the SBR. However, the statisticians are aware that the recent changes to the integrated system of registers favour administrative usage at the expense of the statistical needs, and obviously they see this as a loss, which does not strengthen the role of the SBR in the NSI. This is most directly experienced in relation to activity codes.

... costs The costs of fundamentally changing the SBR application itself due to “self registration” is dealt with in section 3.3. However, in addition to this, “self registration” in the CABR has cost implications for SD’s work in the SBR.

Eventually, it will save costs ... One the one hand, there is no doubt that transferring the responsibility for correct data to the enterprises themselves shall enable cost savings – also in SD’s SBR. Otherwise there would not be a cost-effectiveness gain for DCCA, CCTA, SD and DWEA by “adding” 300.000 extra data providers. Again, taking into account the expected resources available to SD in the coming years, the costs must – and will – be reduced.

... but so far we struggle with “teething troubles” On the other hand, like any other system “web-reg” has had its “teething troubles”. Among examples of the effects this has had on SD are:

- Firstly, the dating of events in “web-reg” has caused SD to handle this separately on the statistical units in the SBR i.e. to do “double work”.
- Secondly, that the many duplicates created in the system, cf. above, had to be identified and closed/deleted. This involved a lot of coordination with the enterprises (in order to find out which identifier(s) they were using for different administrative purposes) and with the DCCA.
- Thirdly, a lot of enterprises keep using SD as their point of contact – instead of the DCCA – because they used to be contacted by SD in relation to the error-checking of LKAUs. This has caused a lot of extra work for us.

On top of that we had a structural reform

Finally, a particular cost-implication occurred because of a concurrent comprehensive reform of regions and municipalities in Denmark taking effect on 1 January 2007, where the number of regions and municipalities was drastically reduced. As mentioned above, the registration of public sector units in the CABR is the responsibility of SD. Thus we had to change/check the registrations for all those units at the same time as the new regions and municipalities were starting to self-register local units in great numbers causing a lot of confusion and extra work.

3.3. A new interface between CABR and SBR

Our response to the above-mentioned challenges

As a consequence of the changes to the CABR explained in the two sections above, SD had to change the functionalities and work processes in the SBR rather fundamentally. Essentially, we had to re-build the “statistical functionalities” that had been removed from the CABR (the common core) in the SBR system, and to switch the working processes from mainly working in the SBR via the administrative units to working directly with the statistical units. In brief these changes include:

- That the LKAUs in the SBR are maintained/followed over time instead of the LLUs in the CABR.
- That the “take-over-check” (regarding local units) is removed from LLUs in the CABR to LKAUs in the SBR.
- That the 1:1 relation between LLUs in the CABR and LKAUs in the SBR will gradually disappear.
- That data will be registered on the LKAUs (i.e. the statistical units will no longer “only” be “transparent copies” of the LLUs).
- That demographic events are registered at LKAUs in the SBR instead of LLUs in the CABR.
- That extracts from the SBR to the statistical divisions are always based on the statistical units (plus information about the LU’s identification number for mailing statistical questionnaires and reminders).
- That the actual data registered for LKAUs in the SBR to a gradually increasing extent will differ from information registered for LLUs in the CABR.

Consequences for the SBR in relation to ...

Basically, these changes imply a less “tight” relation between our SBR and the CABR, where we aim at “keeping the advantages and avoiding the disadvantages”. One could say we end up having a “real” *statistical* business register – with more freedom for SD to pursue statistical priorities – and not “just” a copy of the CABR.

- ... *coverage* The coverage will as point of departure remain unchanged and the CABR will remain as *the* source for the SBR. However, the new model enables us to cover units not currently covered by the SBR, because they are not liable to registration in the CABR, but still subject to (a few) statistical surveys, especially in the field of agriculture. We get back to this in section 3.6.
- ... *quality* The quality of data could increase – from a statistical point of view – because we are better able to deviate from the data in the CABR in those instances where we consider having better information via the statistical surveys, but are not allowed to update the CABR. In these cases we will update the statistical units only.
- ... *usage* The usage of the SBR – from the point of view of surveys already using the SBR – is not directly affected by these changes, at least not in the short term. However, the possibility of adding units to the SBR currently not covered will enable additional surveys to base itself on the SBR. This will be pursued in the agricultural domain, where a separate survey register is currently being maintained. Also, we expect the statistical divisions to welcome the possibility to let data deviate from the CABR in cases where “the administrative point of view” does not seem correct from a statistical point of view.

From the point of view of the SBR staff, the usage will be substantially changed because most working processes will now be directed towards the statistical units and not the administrative units.

- ... *initial costs* The implementation of these changes to the SBR has been extremely costly compared to the resources available to our department. The total direct costs of analysis, specification, development and testing of the system are approx. 6-7 man-years of work. In addition, the mandatory nature of this project – triggered by external changes – meant that a number of other development projects of high priority had to be postponed.
- ... *running costs* We do not expect that the running costs will be significantly affected by these changes.

3.4. Ad-hoc challenges from politically strong users of CABR data

- Reuse of public data is vital for SD’s productivity* ... As mentioned in section 2, a key characteristic of the Danish “info-structure” strategy is to maximise the use – and reuse – of public register data across all government bodies. Together with the general strive for digitalisation, the aims are to reduce the (measurable as well as the perceived) burden on private enterprises and to increase productivity within the public sector. Within this context there is a strong focus on basic data (unique identifiers, name(s), address(es), activity code(s), legal form, size).

This is strongly supported by SD where we for more than three decades have benefited tremendously from reusing administrative records and registers in the production of statistics.

- ... *but it is a double edged sword* However, it is well-known that this strategy is a double edged sword. With gradually more users of the data, gradually more needs and wishes are articulated towards the administrative registers – they must encompass more purposes. And some of these originate from strong political players or emanate

from issues that – for longer or shorter periods of time – capture the political agenda.

*A recent example:
Food safety* A recent example in the Danish context – as well as in many other countries – is that of food safety, where new laws aim to strengthen the control of hygiene in “premises” where different kinds of fresh food is produced, processed, sold and/or eaten. The data on the “premises” needs to be registered somewhere and the CABR is this obvious choice, of course. The entire infrastructure is in place and most units are readily available with addresses and – more or less – usable activity codes.

However, different control measures are pursued for different kinds of food - be it fish, pork, beef, poultry, fruit, vegetables, eggs etc. Consequently, in order to administer the food safety legislation efficiently, the authorities need different (types of) units and claim, firstly, that e.g. supermarkets must be subdivided into separate local units representing each type of food (or each type of control measure), that unmanned vending machines are registered as local units, or that non-stationary and non-permanent vending facilities are registered as local units. Clearly something we as statisticians would disagree with and would not like to see in our SBR.

How important is statistics? Secondly, if it establishes a legal right for the data subject (the enterprise) to be registered with a given set of data in the CABR (e.g. a given activity code or address) – and if it could even have legal/financial implications for the enterprises if these data are changed – then the “relative weight” of statistical purposes and considerations in relation to the data could be significantly reduced: What is most important – food safety or coherent use of a public database? Public health or statistics?

Could the NSI be liable for costs or losses? Furthermore, if there is a two-way cooperation and the NSI is also a data provider due to its expertise in a given area (e.g. business registers), what will then happen if a member of the SBR staff – in good faith – makes creations, updates or cessations which will cause costs or losses for enterprises? Will the NSI be liable for such costs? What level of documentation is necessary for updating a simple activity code or address?

Administrative use of activity codes Another well-known example – from many countries where the business register is publicly accessible – is the increasing use of activity codes for administrative purposes. In Denmark, activity codes are used to determine insurance premiums, to impose indirect taxations (on e.g. waste), to issue parking licenses in cities, or to determine the frequency of inspections of businesses handling fresh food etc.

Clearly, this creates an incentive to “activity code shopping” in order to reduce/avoid costs or to obtain an advantage. This brings risks to data quality, increases costs for SD (because of checking, correcting and answering inquiries from enterprises – as well as from the administrative users). Also, it can create reluctance or even bad-will towards the NSI, because enterprises feel that we inflict tangible costs on them – without giving any tangible benefits. Finally, it can create confusion among enterprises, if e.g. activity codes are repeatedly changed due to different views of what is statistically most correct vs. what gives lower costs. This is something we see from time to time in relation to enterprises in the Prodcom survey.

Consequences for the SBR in relation to ... Common to such ad-hoc challenges is that they are difficult to anticipate – and so are their consequences for the SBR. Furthermore, if they are driven by a strong political agenda it can be quite difficult to manage them in a way that allows an overall sensible weighing of “short term solutions to high-profile/tangible issues” (like food safety) vs. “long term solutions to low-profile/intangible issues” (like quality of statistics).

... coverage The consequences we have seen so far with regard to coverage relates to over-coverage, meaning that some units not belonging in the CABR according to its legally defined definitions have nonetheless been registered in order to fulfil a control purpose with an administrative user of the CABR. Most often the units had been self-registered with data that would not be rejected by the automated validation routines (e.g. ferry boats to Norway had been registered as local units with an official address in the main street of Copenhagen and an activity code in retail trade because of an administrative need to register the ships’ authorisation to sell certain products in their duty free shops).

... quality Potentially, quality effects can be both positive and negative. Positive, if the ad-hoc need draws attention to areas not previously investigated or checked – *and* if subsequent decisions are made according to statistically sound criteria. Negative, if the situation is the other way around.

Effectively, we assume that we have been subject to both. Positive effects – of which we are not really aware – where administrative usage (incl. control purposes) has actually lead enterprises to correct their data (e.g. location and activity), and it just went into the current stream of updates. And negative effects – of which we are aware in those cases we have discovered – where the criteria and business rules in the CABR have been obviously twisted. The net effect could very well be positive, although we have no firm basis for assessing this.

... usage So far the usage of the SBR has not been directly affected by such ad-hoc challenges. However, as we also saw in relation to self-registration, when statisticians become aware of cases where the business register has been “abused” – and even if the effects of such isolated instances may easily be overestimated – it does not strengthen the role of the SBR within the NSI.

... costs So far the costs of managing such ad-hoc challenges has mainly related to analysing and discussing the expected effects of the suggested solutions to new or ad-hoc needs with the CABR and sometimes directly with the particular user. In cases where the suggested solutions challenge the interests of SD, or the underlying rules of the CABR, this can be rather time consuming. Certainly, such issues are also on the agenda of our current meetings with the CABR, which are held on a quarterly basis or as needed.

Monitoring of new initiatives at an early stage In order not to be caught by surprise we have established an electronic monitoring of parliamentary and governmental dossiers, which shall identify – at the earliest possible stage – new legislative initiatives that involve use of data from the CABR. On the basis of the proposals we try to asses, whether the initiative could have negative effects on the register from a statistical point of view, or whether it could perhaps create incentives for the enterprises to prefer incorrect data. In such cases we will contact the owner of the proposal with a view to explain the consequences and discuss potential alternatives.

3.5. New supra-national requirements

<i>Enterprise groups must be maintained in the SBR</i>	The new EU Regulation from 2008 on business registers for statistical purposes is a well known challenge to many countries, but – since we do not currently have Enterprise Groups (EG) integrated as units in our SBR – it differs from the above mentioned in the sense that it does not so much challenge “the way we do things now”, nor the quality of current statistics. The challenge is to extend our current SBR system and build an efficient process which will comply with the new requirements and fulfil the needs regarding (national and multi-national) EGs in spite of the fact that the necessary data are not available in the main source for our SBR.
<i>Currently, there is no official source in Denmark</i>	Because actually, although there is a multitude of official and high quality registers in Denmark, there is currently no official register of company owners (share holders). The creation of such a register is likely to be proposed in connection with an upcoming proposal for a comprehensive modernisation of the Danish legislation for limited companies. However, even if it should be decided to create such a register, and it would be usable, it will most likely take quite a while before it is operational. Until then we need to find a different solution, and thus we have decided to acquire monthly data on ownership from a commercial provider and to use those data together with data from the CCTA on joint taxation of companies, which we already have in the SBR, cf. figure 2. This will largely replace the current (mainly manual) annual profiling of the largest (national) EGs in Denmark.
<i>Consequences for the SBR in relation to ...</i>	In relation to the growing importance of statistics related to “globalisation” and to the EG level there is no doubt that this change will have a high impact on our ability to fulfil our objectives in the future.
<i>... coverage</i>	Based on discussions with the statistical division responsible for the FATS statistics about their needs in relation to both EU requirements <i>and</i> national needs for statistics, we will aim for a solution with a higher coverage than stipulated by the EU Regulation, namely in relation to Danish companies’ affiliates outside EU.
<i>... quality</i>	Obviously, the implementation of Enterprise Groups in our SBR – and the data exchange via the Euro Groups Register – will greatly improve its <i>relevance</i> for the FATS and FDI statistics (compiled by the Central Bank who gets a full copy of the SBR on a monthly basis). However, it is also expected to gradually improve the <i>coherence</i> in the register as a whole.
<i>... usage</i>	Currently, EGs are not implemented in our SBR, but it is on the way. Our ability to meet the needs and requirements very soon will be significant for the credibility of the SBR, since the surveys that need these data are already being implemented.
<i>... initial costs</i>	The costs of <i>establishing</i> the EG part of our SBR application system is estimated to approximately one and a half man-year, with a bit more than half of it being used on the business side in analysing the new data sources, designing the new process, specifying needs and requirements for the IT system, and project management. The rest is used for the IT development. The new system shall be operational by mid 2009.

... running costs The costs of *running* the EG part of the SBR will be a bit lower than what we previously have spent on profiling (approx. two and a half man-years). However, the tasks will be changed rather substantially.

3.6. Integration of the last separate registers of statistical units

Not all surveys are yet based on the SBR The ideal situation for the NSI is to have only one SBR, which fulfils the needs of all surveys comprising business units. Unfortunately, in SD we are not there yet. The domains, which are not yet fully based on the SBR and the standardised statistical units, fall in two main categories. The first category comprises a few short term statistics relying on administrative records based on VAT units, cf. figure 2. However these are not really problematic from a statistical point of view.

Agriculture has its own register ... but for how long? The other category comprises the majority of our surveys in agricultural statistics, which to some extent must cover also units which are not currently in the SBR because they are not liable for registration in the CABR. These are mainly small farms, but also structural and conceptual differences hampers the usage of only the SBR in agricultural surveys. Thus, SD's division for agricultural statistics runs a separate register, where approx. 90 percent of the units are also in the SBR, but do not benefit from the on-line updating which takes place in the SBR.

Cost savings is the driver However, a change to this situation will be pursued. The current reorganisation of SD's Business Statistics Directorate (which includes agricultural statistics) has as one of its objectives for 2009 that the Register of Agricultural Units shall be integrated with the SBR. A specific project will be created for that purpose. The underlying driver is *cost savings*, so the challenge we face here is the standing criterion of ever increasing productivity in the NSI. (Of course, the *internal* SBR processes *themselves* must also increase productivity, but that is not dealt with in this paper.)

Consequences for the SBR in relation to ... As mentioned in section 3.3 the changes recently implemented to the SBR system enable the inclusion of units in the SBR which are not in the CABR, and like it was the case for EGs, cf. section 3.5 new sources are also needed in this area. So far, data from a series of administrative registers – from different authorities – on land, dwellings, domestic animals and crops have been combined in order to maintain the register. However, from a preliminary assessment it seems, firstly, as if the data model could relatively easy be adapted to the SBR and, secondly, that it would probably be more efficient to use only these sources for units that are *not* in the CABR and in the SBR. The units already in the SBR should benefit from the on-line updating just like the rest of the SBR.

... coverage The coverage would increase, enabling the SBR to deliver the units needed in an area we currently do not cover completely. The new units will mainly comprise small units – too small to be liable for VAT, but big enough to be subject to EU requirements for statistical reporting.

... quality Obviously, the *relevance* of the SBR will increase for the agricultural statistics division. Also the *timeliness* and *accuracy* of data for the new units will improve due to current updating via the CABR, and finally, data for those units already in the SBR will become more *coherent*.

- ... usage To have the agricultural surveys using the SBR instead of their own satellite system obviously increases the usage. Also, having the agricultural statistics experts updating the SBR directly instead of their own satellite system will be an important gain.
- ... costs So far only an initial assessment has been made, but it seems as if the initial costs will be limited. With regard to the running costs the integration of the agricultural statistics register must enable cost savings from the outset, cf. above. Maybe the biggest barrier in this project is one of tradition, or different “corporate cultures” in agricultural statistics vs. business statistics?

4. What can we conclude – and perhaps learn?

The discussion of recent challenges in relation to SD’s SBR confirms a number of well known advantages as well as disadvantages of having a SBR which is strongly linked to ABRs. These pros and cons can be summarised as in figure 4 below.

Figure 4:
Pros and cons of close
integration of SBR and
ABR

Benefits / advantages	Costs / disadvantages
<ul style="list-style-type: none"> • High coverage (depending on requirements for legal registration) • High quality (if data are validated by the administrative body and updates are well coordinated) • Reduces costs • Reduces administrative burden on business • Supports frequent statistics (if updated on a current basis) • Potentially more timely data (depending on technical set-up and administrative processes) 	<ul style="list-style-type: none"> • Administrative definitions can deviate from statistical needs and definitions • Problems regarding match and consistency • Use of different classifications (and different use of the same classifications) • Limited possibilities for collecting supplementary data (because of burden considerations and division of responsibility/competence) • Reluctance among enterprises regarding data exchange (necessitates assurance of confidentiality and “documentation” of positive cost/benefit ratio) • Dependency on providers • Vulnerable to political and/or administrative changes • Potentially less timely data

These pros and cons – and the lessons learned from the above-mentioned challenges – reflect some inherent differences in the nature and raison d’être of the two types of registers. Some of the key differences are summarised in figure 5 below.

Figure 5:
Key differences between
SBRs and ABRs

Aspect	SBR	ABR
Purpose	Supports production and dissemination of (mostly) aggregate statistics.	Supports binding legal or administrative decisions about individual entities.
Content	Focus on economic phenomena – at institutional and productive level, respectively. Complex approach to units.	Focus on legality and responsibility of individual operations. (More) simple approach to units.
Consequences	Neutral - no direct consequences for the data subjects.	Defines rights and obligations – direct legal and/or financial consequences.
Relation to policies	Indirect	Direct
Political weight	Relatively “light” – benefits are less tangible. The SBR is seen a bit as a “bi-product”.	Relatively “heavy” – implications are more tangible. The ABR is seen as the “main” product.
Time perspective	Basis for coherence and comparability over time – complex approach.	Basis for decisions taking effect as from the date of registration – simple approach.
Quality philosophy	Multi-dimensional approach. Data are accurate when equal to a real value – which is often unknown. The level of “sufficient quality” is unclear.	More simple approach. Data are accurate when equal to the self-perception of the unit. The level of “sufficient quality” is (more) clear.

Against that background, we could ask ourselves how we – under these framework conditions – could become better at meeting different kinds of external challenges to our SBR in the future.

The first challenge was the replacement of a key (sole) administrative source for a business process in the SBR. The important factors seem to be:

- To communicate needs and positions pro-actively to the owner at an early stage where the content can still be influenced – i.e. in the early legislative phase (conception phase). Networking is a key factor here.
- To make sure that the administrative data collected comply – as much as possible - with statistical concepts and definitions.
- To pursue a solution where the administrative authority has an interest itself in the coverage and quality of the data – and perhaps in using data more “statistically”.
- To follow the project all the way to the practical implementation.

The second challenge was a change to the underlying business model and perception of business registers. The important factors seem to be:

- To try to set (or keep) the agenda. Official statistics has a lot to offer ABRs – well founded concepts, well defined units and classifications etc. In this context we could perhaps even ask ourselves whether we could contribute more *from* the SBR *to* the ABR – i.e. making some of our non-confidential data publicly available, e.g. data on EG structures?
- To keep as much of the SBR business logic in the ABR as possible, but at the same time minimise the SBRs vulnerability in case the owner of the ABR changes its strategy. In other words: “Exploit the advantages and avoid the disadvantages”.

The third challenge was to adapt our systems and processes in the light of the first two external challenges. A lot of factors related to technology and project management could be mentioned here, but in this context we will emphasise:

- To keep the SBR system simple. The more complex and sophisticated (i.e. ambitious) it gets – the more expensive is it to change.
- This also includes that the NSI is modest when defining requirements towards the ABR.

The fourth challenge related to political pressure on the business model of the CABR which could deteriorate quality. The important factors seem to be:

- To identify them – and early enough to still be able to influence them.
- To take a pro-active role in relation to the owner of the proposals/”threats”, preferably with a view to suggest a better alternative, e.g. in relation to administrative use of activity data, alternative aggregations etc.
- To emphasise the need to keep the overall system sustainable and not “let the ABR chase the latest policy need” if it is not in line with the overall aim and strategy for the register system *as a whole*.

The fifth challenge related to new supra-national requirements, which could not be met by the existing administrative sources. In this context the important factor seem to be:

- To try to have new needs and requirements incorporated in administrative sources in order to avoid parallel systems. (At least not to build permanent parallel systems).

The last issue related to the NSI’s permanent challenge to “do more with less”: This could relate to many aspects, but the important factor in this example is:

- To exploit the main source/system to the maximum extent and not accept redundant data, systems and updating processes.

Summing up

To sum up we could say that, fortunately, SD already has a strong platform in the “register-sphere” in Denmark. Among others this is ensured by our legal basis that 1) obliges other public authorities to consult SD when establishing or modifying registers and 2) ensures SD’s access to public registers for statistical purposes. However, in a system where the SBR is highly integrated with ABRs, it is necessary – but not sufficient – to have a thorough *technical* understanding of the sources. It is also necessary to be “*politically*” *pro-active* in order to maintain the SBR as a basis for quality statistics. Thus:

In order to better fulfil our role and objectives we need to:

- **Manage our partnerships better – we are not strong enough alone**
- **Communicate proactively – we cannot wait for others to contact us**
- **Always be part of the solution – not part of the problem**
- **Manage our risks better – otherwise they seem to manage us**
- **Keep things simple – balance ambitions with abilities**

5. References

- [1] Regulation (EC) No 177/2008 of the European Parliament and of the Council of 20 February 2008 establishing a common framework for business registers for statistical purposes and repealing Council Regulation (EEC) No 2186/93.
- [2] Council Regulation (EEC) No 696/93 of 15 March 1993 on the statistical units for the observation and analysis of the production system in the Community.
- [3] Act on Statistics Denmark - Consolidated act No 599 of June 22, 2000.
- [4] Lovbekendtgørelse nr. 653 af 15. juni 2006, and Bekendtgørelse nr. 1403 af 6. december 2007.