

# **Different census systems in Europe: lessons for the transition to a register-based census system in Belgium**

Johan Surkyn

Interface Demography, Vakgroep Social Research  
Vrije Universiteit Brussel  
Pleinlaan 2  
1050 Brussel

## **1. Introduction**

Statistics Belgium is preparing for its transition to using administrative data for the population census. In spite of a well-equipped system of basic registers, Belgium was not one of the forerunners in this transition. Today however, we hope to benefit from the experience acquired in other European countries, and maybe avoid struggling with some of the challenges others were faced with before us.

In this presentation we shall try to draw some lessons from international experience, and evaluate the Belgian case against this background. Three examples were selected that offer a wide and informative range of approaches to the administrative census. Finland, because of its well-established and totally integrated system of administrative data-collection and exploitation; Switzerland because of its federal state structure comparable to the Belgian one (although, on the other hand, absence of a central population register in Switzerland creates an entirely different starting position); and the Netherlands because of the unique system of combining sample-based survey research and exhaustive administrative databanks into the “virtual census”.

Finally, given the fact that we are currently conducting a test run of the Belgian register-based census, a first glance will be offered at its contents and formats.

## **2. Administrative data as a source for social statistics**

Obviously, statistical use of administrative data is not limited to the population census. On the contrary, in most cases extensive use of these sources preceded the implementation of the administrative census, and continued to do so after it was established. Nowadays administrative data serve statistical purposes in many fields, not in the least fulfilling the E.C. (EUROSTAT) statistical requirements. Nevertheless, there is still a widespread consensus that both survey-data and administrative sources are needed to complement each other in a functional statistical system.

A number of major advantages have raised interest in the use of administrative data as a source of statistical information by the national statistics institutes. First of all, we must mention the fact that administrative data on a growing number of subjects have become available, and that the technical means for exploring them keep advancing. In most domains of administration paper documents have long been replaced with electronic ones, and information is transmitted (xml-format, RSA) over internal and external networks. Citizens, businesses and other administrations are offered the possibility of directly consulting or entering information to the administrations database. Sophisticated techniques of “risk-analysis” and identity authentication are designed to prevent fraud in this open system of data-exchange.

But more important substantial and organizational arguments are put forward:

The first and most often cited is cost reduction. Indeed, sample based survey research is extremely costly, labor intensive and time consuming. In several cases (Netherlands, Finland) the cost of an administrative census – discounting the one-time cost of implementing the database and its links with administrations - was estimated to be about one tenth of the questionnaire-based predecessor. Moreover, this lower price can be more evenly distributed over time. Contrary to the questionnaire-based census that caused a peak in expenditure every ten years due to the cost of data-collection and coding, data-collection in an administrative system is not a major cost at all.

This feature also allows the administrative censuses to be repeated more rapidly, reducing periodicity in census data. As a result of an integrated system of administrative functioning and statistical implementation, the full Finnish census dataset has been available for every single year since 1987.

Drawing further on the element of time and efficiency, the longitudinal character of most databanks is of great interest to many scientists: an advantage perhaps most appreciated by demographers. Classical panel studies are needed to disentangle age, period and cohort effects in demography and indeed in other disciplines as well. But the problems in terms of optimal design, costs and (selective) drop-out are well known to social scientists. Administrative databanks constitute an optimal basis for (retrospective) pseudo-panels, and their design can be adapted at any time in order to address issues requiring different selections of cases or different timeframes. As the demographic methodological toolkit of event-history analysis is used ever more widely in other fields, so does use of administrative pseudo panels stemming from registers on employment, businesses or educational careers, to name just a few applications.

Another argument for using register data is in lowering the response burden for the population. Indeed, citizens are flood with questionnaires in schools, the workplace, and at home. A variety of canals for interviews and questionnaires are frequently used: telephone, post, house calls, the internet, even popular magazines, and so on. For many citizens the distinction between administrative, scientific and commercial questionnaires is not obvious. Therefore a certain tiredness with interviews and questionnaires is said to be growing in the population. It seems however, in as far as efficient use is made of the

data (e.g. the same information is not collected over and over again, and the aims of research are clearly communicated), the willingness to co-operate with socially relevant survey research is still high. An interesting example in this respect is the subjective health question introduced for the first time in the Belgian 2001 (the last) census questionnaire. It was feared to yield very low response rates, especially since health is considered to be among the most privacy-sensitive subjects in the census. However, response rates to the census in general were quite high (about 96 %) and item non-response on the health questions was almost neglectable (less than 4 %). The information on subjective health has been used in several applications, and turns out to be among the most revealing ones in the census. Furthermore, this is a question that would have been almost impossible to replace with administrative data.

On the other hand, and largely outside of the information collected in the census on an individual level, it must be acknowledged that statistical as well as administrative collection of data weighs heavily on businesses and enterprises as well. Effectively, in a broader view on reducing response burden, administrative databanks can be very useful. They serve as a “collective memory” of information transmitted by and amongst citizens and enterprises, and permit efficient use of the data. At any rate, proportionality between the burden on individual, commercial and institutional respondents, and the intrinsic value of research results for society needs to be checked and respected at all times.

Exhaustiveness of administrative sources is also beneficial. In some cases the administrative universe even exceeds that of the classical census. School records include children stemming from refugee populations and other groups absent in the population register. The cadaster (or register of buildings and dwellings) covers dwellings occupied by those same groups, second residences and even unoccupied ones. Thus, switching to register data often helps broadening the research horizon.

However, in some cases the administrative universe may be narrower as well. Information on a part of the workforce working in another country than the country of residence is often missing. The same goes for the exact (highest) educational level of the growing number of students obtaining a foreign degree, or of foreigner entering the country. Just as non-response in classical survey research, incompleteness of registers is never unbiased. However, even if total exhaustiveness is sometimes out of reach, the sheer scale of administrative databanks has many advantages. Data do not have to be smoothed out for irregularities, and significance tests in statistical analysis become obsolete. This permits administratively collected information to be used as a basis for calibrating survey research data, in very much the same way the classic census was often used to redress sample distortions by weighting data according to over- or under-representation in respect to the total population.

But administrative data do also have a number of drawbacks:

First of all data collected (solely or primarily) for administrative purposes usually need to be re-treated extensively before being useful for statistics. For obvious reasons these data are not tailor-made for studying social questions, and definitions and typologies may

differ substantially from those used for statistical purposes. In a number of cases identifying the basic unit of analysis in register data can be a problem. The dwelling – the basic unit in housing statistics - is sometimes difficult to track in registers. In part this is due to the problem of defining what a dwelling exactly is. On the other hand the dwelling as described in the classic census (the constructional unit that is used as a living space, with all its extensions and facilities) is not a simple concept. One can wonder whether a detached garage, or even one rented around the corner should be considered part of it. If so, it is clear that register data will have great difficulty in capturing them.

But there are other reasons why administrative reality does not always reflect social reality. People may willingly interfere with the registration of their actual living situation because of financial reasons (tax evasion, unrightfully collecting social support). Or registration itself can be responsible. For instance, NACE-sectors of economic activity can get consciously miss-registered because administratively they only serve to determine the social security regime of the employees. In that case, any other NACE-code linked to the same regime can be used as a substitute. As a result, more “exotic” NACE-sectors can get underrepresented in statistics. Furthermore, variable detail that has no administrative use tends to be generally neglected.

Another point that has to be made is that continuous registration does not necessarily imply that data are up to date. For instance, many small-scale transformations to buildings and dwellings are not transmitted to the authorities. It may take a while until a sale or a larger scale transformation forces a re-evaluation of the property by the proper authorities.

A more substantive element is that certain types of information are not well-suited to be registered at all. Political opinions, subjective evaluations or any characteristics that tend to change at very fast rates, are generally not captured in registers. However, this is not an argument against use of registers. It merely indicates that survey-data and registers are complementary in offering good quality data at a fair cost in terms of budget as well as response burden.

One of the main issues a large scale register-based survey must deal with is the problem of consistency. Different registers may detect events with longer or shorter time-lags and this leads to inconsistency: people in the workforce may unduely still be found in school registers, retired persons may still occur in the registers of the active population. Such mistakes may be hard to detect and correct, as some of these situations may reflect reality (students can have a job, people receiving pensions can have a job). The basis of this problem is that a registered based census implies that a unit person record has to be constructed across different databanks. In a complete enumerated persons census, the record is received in a single throw, and consistency is enforced by the single respondent reporting all the information.

Finally, the privacy aspect of using registers instead of questionnaires must be mentioned. However, the balance is not an easy one to make. It is clear that register data themselves are more delicate than surveys. Respondents will not lie or refuse giving the information

in order to protect their own privacy. In addition, the content itself is more critical, as it touches on tangible consequences of situations relating to delicate subjects such as income, health, living conditions or belonging to minorities. In a register-based census, cross-relations between such subjects are established, revealing information that was often never combined in the same way before.

However, in some ways, privacy is also better protected. Without paper forms the information collected is rarely exposed to prying eyes, and fewer people will ever deal with raw data. On the contrary, most operations will be performed on coded data, and through automatic cleaning and linking software. Statistics Finland (2004) even boldly states that register data offer better protection of privacy.

### **3. Switzerland, the Netherlands and Finland: structural differences, alternative pathways, uniform goals?**

Finland, Switzerland, the Netherlands and Belgium present different stages in the transition to the register-based census. But, partly due to different institutional settings, they also offer a view on different pathways and strategies. First, the Swiss situation will be presented, as it is more or less in the same stage as Belgium.

#### **3.1. The Swiss statistical system**

From a Belgian point of reference, Switzerland is an extremely interesting case for international comparison. Just like Belgium it had a questionnaire-based census around 2001. It actually ran between December 2000 and July 2001, and covered all the classic subjects. For the future (2010), the Swiss statistical office (OFS) is preparing for a gradual transition to the use of register data. The Belgian approach is (announced to be) rather more radical.

Just like Belgium, Switzerland is a federal state (after the United States, the second-oldest in the world). However, the institutional setting is very different. The Swiss cantonal statistical offices have a high degree of autonomy, and differ widely in terms of their actual statistical capacity (Fellegi and Ryten, 2000). Several cantons do not even have a statistical office. However, statistical competences are distributed, not only between the cantons and the federal state, but the municipalities and the federal institutions all have some statistical functions.

The Swiss federal statistical offices main role is coordinating the activities in the local offices and in federal departments, agencies and offices. ONS does not carry out any primary data collection, but coordinates those activities and offers methodological guidelines. As an institution it under supervision of the Federal Statistics Commission (CSF).

It is widely recognized that a register based census, apart from substantial registers offering information on education, work status, living conditions, housing and so on, must rely on a set of basic registers that contain the clues for identifying all census subjects uniquely and univocally and for establishing certifiable links across different databanks. In an ideal setting, as the basic census units are mainly individuals, households, employed persons, dwellings, and businesses and enterprises, these fundamental registers should consist of the population register as a central point of reference, and a register of business and enterprises, a social security and labor status register and a register of buildings and dwellings.

The Swiss situation is marked by the absence of a central population register (it does have an exhaustive register of foreigners residing in the country). Population registers are a local matter, depending on municipalities and regulated through cantonal constitutions. This means that about 2900 separate communal population registers exist, which obviously poses great challenges to performing a census based on registers (and indeed any exhaustive census). At the same time a long-term process is in operation whereby municipalities merge into larger units.

However, Switzerland has both a Federal Business Register (BUR) and a Buildings and Accommodation Register (GWR). The first one (BUR) does not contain any individual data on the workers and employees, but can be useful for studying size of enterprise, type of economic activity or location. The latter one (GWR) was, until recently “under construction”. The Swiss 2000 population census was used to supply the Buildings and Accommodation Register with a number of stable characteristics of buildings and dwellings (number of floors, building period, number of rooms in case of dwellings, etc.). Updating the GWR is based on the yearly OFS construction statistics.

The GWR offers some compensation for the absence of a central population register by providing a register of dwelling addresses and even GIS coordinates (geocoding). In order to facilitate future censuses, OFS recommends entering building codes and unique dwelling codes into the communal population registers. In other words, since Swiss constitution does not allow the federal statistical authorities to directly exploit and model the communal registers (even less to join them into a total persons register), OFS recommends injecting an identification system that is consistent and complete over the whole of the federation.

But, regardless of the motives for developing the GWR and its system of registering and identifying buildings and dwellings, it is very clear that GWR will be instrumental to the future register-based census in many ways. This point will become clear from the Belgian situation, where, caused by the absence of such a register, existing good quality administrative data on housing cannot be used to their full potential.

Therefore the Swiss GWR, and especially its system for identifying separate dwellings and linking dwellings to their occupants (as well as buildings) deserves some more attention. GWR is based on a few basic principles. The first one is defining the objects in the database. Dwellings are defined as: “the total of rooms constituting a constructional

unit destined as living space for one or more private households, and including minimally the following infrastructural elements:

- Installation for cooking with permanent water supply or energy supply
- Central heating or another permanently installed heating system

A Dwelling has its own entry, either from outside or from a common space inside the building (staircase)”<sup>1</sup>.

By definition, dwellings are found inside buildings. Therefore, special types of non-standard facilities used for habitation are registered, even if not complying to the definition of buildings as below: barracks, caravans, mobile dwellings, commercial vending stands and so on:

“Buildings are permanent or long term constructions attached to the ground, serving for dwelling or professional purposes, education, culture or sports ... ”

Buildings are identified using a unique 9-digit number that remains in existence until the building is demolished. The number is on national level, administered by the central statistical authorities. Buildings are also identified by an address and one or more possible entrances. Inside buildings a further distinction is made between separate dwellings. Each dwelling receives a (random) dwelling number. Separately, a dwelling entrance number is recorded. This is the usual address extension as used in postal addresses: 23/B or 23/7. To complete the identification part of the record, a floor number is registered as well, and space is left for some additional (alphanumerical) information helpful in situating dwellings (left/right/center/owners or occupants name etc) located on the same floor.

In order to keep track of the nature of dwellings and evaluate the necessity of including or excluding dwellings from the sample according to research demands, the normal occupance status is included. Distinction is made between permanently inhabited (domicile dwellings), partly inhabited (tourism, second residence, ...), re-destined (turned into commercial use, ...) and uninhabitable dwellings.

Establishing the link between (identified) dwellings and their inhabitants is achieved when communal authorities record citizens addresses in such a way that a unique match between dwelling (extended) address and domicile address is possible. Obviously, if based on such a match, dwelling numbers are included in the population system, the database is ready for statistical implementation.

### 3.1.1. The Swiss census: between registers and questionnaires

The Swiss 2000 census covered all classic major subjects:

- Date of birth, gender, marital status, citizenship, place of residence, place of

---

<sup>1</sup> Authors translation

birth, position in household, number of children, religion, languages, education, learnt and practiced profession, current employment situation, job title, place of work/study, distance traveled to place of work/study, means of transport used and commuting time were collected on individual level.

- For households: the number of individuals living in the household and the type of household was asked.
- For dwellings the apartment floor number, number of rooms, surface area, kitchen, type of occupant, and amount of rent paid have been enquired.
- Finally, on the level of buildings the geo-coordinates, type of building, time of construction, renovation, number of floors, heating system, hot water supply, building floor and type of owner were asked

We can be very brief on the subject of the Swiss register-based census. The last (2000) edition was almost entirely based on questionnaires, the only exception being imputation of a number of fields of missing data based on registers. Missing data for variables date of birth, gender, marital status, place of residence, citizenship, main language, position in household, current employment situation were completed by register data.

The business and enterprise register was not used, except perhaps for refining location of the workplace for mobility studies (commuting). As mentioned before, the relationship between the census and the GWR (buildings and dwellings) register goes in the opposite direction. GWR will use the information stemming from the 2000 census as a starting point, to be updated later using the OFS construction statistics.

### **3.1. The Dutch statistical system**

The Netherlands present an entirely different case. Contrary to its Swiss counterpart, the Dutch Statistical Office (CBS) is quite centralized and independent. CBS depends on the Dutch government for funding, but its policy and research agenda is based on a mixed independent external body: the Central Commission for Statistics (comparable to the Swiss “Commission de la Statistique Fédérale”).

The Dutch situation relating to basic registers (registers containing identifiers relating to individuals, workers, dwellings and businesses and enterprises) compares favorably to the Swiss one. The Netherlands have a population register that, even if kept by the municipalities, is centralized. Secondly, a jobs file register exists that covers employees and self-employed persons (the latter actually through incomes tax registers). Furthermore, the social security administrations offer additional information. Work on integrating these registers is currently being done in the WALVIS-project. A Basic Register of Dwellings also exists in the Netherlands, but it is unclear to what extent it allows physically describing dwellings and their equipment. Finally, a central register of businesses, enterprises and institutions is in preparation: a project of law ordering the installation of such a register has been entered in 2005.

In short, the Netherlands have a configuration of basic registers that is surely near-optimal to serve as the backbone for a register-based statistical system, and therefore also for the further development of the (virtual) census.

### 3.1.1. Transition to the Dutch virtual census

The last Dutch questionnaire-based census dates back to 1971. By that time, response rates had dropped and the budgetary arguments against a complete survey gained importance. After that date, the Dutch Statistical Office (CBS) gradually moved towards use of administrative data. However, and this is the unique element in the Dutch census and broader statistical system, from the start efforts were made to integrate surveys into the system.

In the eighties large scale surveys still figured alongside register sources. This meant that tables were produced using different sources, but table totals and margins would differ depending on the source of the data. Lately, using the “repeated weighting”-method, both types of data are linked more closely and integrated in such a way that, at least in table-format results, the distinction between register and survey data has become almost invisible. However, on the level of micro-data it is clear that individual level linking of information would be partial and probabilistic. No complete “flat” data file underpins the results. This is why CBS introduced the term “virtual census” indicating that a more or less complete census was constructed from data rather than physically undertaken on enumerated individual level. Or, put in another way, the Dutch statistical authorities chose imputing (estimating) table cells, using sample data, instead of mass-imputing micro-data.

A number of arguments can be brought against this strategy. Imputation of data relies on statistical techniques and hypotheses that are not beyond criticism, even if the imputation is carried out on table level instead of on the level of individual records of data. However, in view of the objectives of the present study, it is unfair to compare the Dutch system to a complete questionnaire-based method. When a registered based approach is used as a point of reference, a rather positive and balanced picture emerges. This is linked to the fact that the “virtual census” technology actually combines three aspects: first, linking data from different sources between which no unique and certifiable (micro-)link can be established; second, scaling up sample partial data (in this case surveys) up to the level of the total population; and third, compensating for the selectivity that always accompanies sampling, regardless of whether the sampling stems from non-response, survey design or incompleteness of registers.

Most of those problems would have to be dealt with in a registered-based approach as well. Internal and external (cross-)inconsistencies, differences in typologies and definitions, and lack of identifiers that cover all register sources force a probabilistic element in linking different (register-) sources. Secondly, be it for different reasons, register data are often just as incomplete as surveys, though it may help that the source of incompleteness of registers is often better documented or understood. Thirdly, a range of

subjects is not suitable to be recorded in registers, and will only be found in sample surveys (e.g. “Subjective health”). Therefore, integrating survey-data in the census helps to make good use of available data and relief response burden for citizens and enterprises.

For all these reasons, the Dutch experience is highly instructive for any statistics institute undertaking (gradually or in a single step) the transition to new sources of data for the purpose of the population census, and may even find useful implementations far beyond the census.

### 3.1.2. Content of the Dutch census

The Netherlands adhere to the EUROSTAT census programme (“Gentlemens Agreement” on census) and therefore the census covers all the classic subjects, following the guidelines of the European statistical authorities on variables and levels of regional detail (NUTS). Sources of the census are a mix of register and survey data, integrated through the technology of “repeated weighting”.

The 2001 virtual census does not derive directly from its register and survey-based sources. First a statistical application (database) was created as an interface, integrating data from the registers summed up above and surveys such as the Labour Force Survey: this constitutes the SSB: Social Statistical Database. Obviously, SSB is used much more widely than just for the ten-yearly census. It is the basis for reporting statistics to EUROSTAT and is widely used for thematic social research. However, the SSB was not the only source for the census. The Survey on Employment and Earnings (SEE) was used for additional information on jobs and on commuting, and the Survey on Housing Conditions was used for the housing tables.

Results from the 2001 census were presented in the form of 40 extensive tables, eight of which concern housing, two tables concern commuting. The other thirty tables are demographic tables relating to occupation, level of education, and economic activity (Nordholt et al., 2004). Consistency between these tables was ensured by repeated weighting. This technique ensures that tables sharing a common margin (for instance a partitioning of income levels) can be made mutually consistent by assigning weights, as if they were samples within the same population. Each additional table is then re-weighted to be consistent with earlier ones (hence “repeated” weighting). The sequence depends on the scale of the population in the table: large or exhaustive tables come first, so their original margins have priority over later – smaller- tables. Regression-techniques are used for estimation of table weights.

### 3.3. The Finnish statistical system

The list of countries in this comparison follows the logic of increasing dependence on register-data for the census. The Finnish census relies exclusively on register data, and presents the “nec plus ultra” in integration of administrative and statistical functions of registers. Not surprisingly, registers on all major components and units in the population

census are elaborated and mutually linked. By 1970, Finland already had a central population register; a register of real estates, a social security and pensions register, and – contrary to all examples above – a register of degrees and completed education. Ten years later a business register (later completed with the public sector); a register on job-seekers, a work pensions register and a buildings and dwellings register were added. Finally, in the late 1990's a students register was set up.

Statistics Finland itself keeps the students register, the register on degrees and the business register. The other registers are managed by the respective administrative authorities and the Central Population Register Centre. Even though it has been operational for quite a long time, the Finnish statistical system is still improving. As an example, the addition of information on Finnish students abroad kept by the Ministry of Education can be cited. Thus, on their return to Finland, the register of degrees is updated with foreign degrees, ensuring completeness of the educational data.

As a result, Finnish social statistics are among the best in the world, allowing analyses on a broad series of subjects in stocks as well as in flows, and available in single year intervals since 1987. Flows can be studied within as well as between registers (e.g. transitions from the student population to the workforce, or death rates by professional background, etc.).

All this would not have been possible without a strong legislative framework, imposing primary use of existing data sources for statistical purposes on Statistics Finland; imposing supply of unit-level data on register holders, including the identification data; and stipulating strict data protection procedures on all parties.

### 3.3.1 Transition to the Finnish register-based census

Finland's censuses go back as far as 1749: one year after the statistical bureau was established. Questionnaire-based census editions took place until 1985. Starting 1987 all census subjects and many other domains of social statistics could be covered using register-data. In 1990 the first full-pledged register-based census took place. After that, censuses became yearly. It must be underlined that the transition, now 25 years ago, was gradual and improvements keep getting made.

However, Statistics Finland also continues to conduct survey-research, even if it does not depend on it for the census or for most employment and population statistics. On the contrary, even in the sample surveys maximum use is made of available (register) data on education, demographic characteristics and so on, so that the questionnaire can focus on the questions most essential for the survey, such as labour force participation, hours worked, consumption, etcetera.

## 3.4. The Belgian statistical system

Belgium is a federal state. The Federal Statistical Agency is, among other missions, responsible for conducting the census, organizing large-scale surveys, collecting data from enterprises and institutions and reporting to EUROSTAT. The agency is situated within the Federal Public Service Economy (Directorate-General Statistics Belgium), and is supervised by an external body: the High Council of Statistics.

The Brussels, Flemish and Walloon regions have independent statistical institutions. Contrary to the Swiss case, Statistics Belgium did not depend on the regions for conducting the census. However, as some of the registers that will be used in the future administrative census resort under regional competences, this is bound to change. Cooperation with competent regional authorities in preparation of the register-based census has already started.

The Belgian system of basic registers is well-equipped: only a register of Buildings and Dwellings is missing among the classic four central basic registers. The population register is operational since 1985. The social security register was established in 1990. It is an integrated network of about 20 registers concerning child allowances, job-seekers, health insurance, social contributions, pensions administrations and others. Since 1998, this network (Crossroads Bank for Social Security) has its own statistical application that is widely used for labour market and social security statistics (Datawarehouse Labour Market and Social Protection). Finally, the Ministry of Economic Affairs also houses a Register of Enterprises (KBO) that covers businesses, enterprises and institutions and all their local branches (2003). Expanding

In principle, registers of students and degrees exist. However, education is a matter of linguistic communities, and therefore the situation somewhat differs between them. Moreover, none of these registers go back before about 2000 and foreign degrees are not captured, regardless of whether they concern immigrants with prior education or Belgians leaving the country (temporarily) for studying. However, efforts of creating more formal and complete registers on education are currently taking form.

The absence of a register of buildings and dwellings is certainly a minus. However, the Belgian cadastre (Patrimonium Documentation Administration) has started introducing buildings as separate units in their data-models. On the level of dwellings, some more progress needs to be made before a dwellings register can be implemented.

#### 3.4.1 The Belgian census: a few milestones

The Belgian census has a long tradition. Adolphe Quetelet organized the first census in 1846, and the next edition already contained a large set of questions of social-economic relevance. Over 160 years 16 censuses were organized: roughly every 10 years. Originally the census was set up to provide (corrected) municipal population figures, and to determine regional distribution of political representatives. After the central population register was established in 1985, the following census lost these functions. The census no longer had a direct administrative significance.

Starting the 1991 edition, the population register was used for pre-printing personal identification information on census forms. The 2001 questionnaire-based population census almost didn't take place, as it was planned that a register-based approach would be implemented in the future. However, it was clear that preparations for this transition had not advanced enough to undertake such a step. The 2001 census was then explicitly organized as the last in the old format. Marking the transition and the long way the census had run since the early years with Adolphe Quetelet, it was renamed as the "Social-Economic Survey". For the 2001 census, further use of the registers was made, and the census contents were designed to be the starting point for longitudinal databanks on housing, education and other subjects.

Belgium takes part in the European Census program. The 2001 Survey covered demographics and household position, employment status, sector of activity, occupation, education, health, giving and receiving informal care, mobility (commuting), housing and living conditions, energy consumption, the living environment and subjective appreciation of neighborhood quality: totaling to about 150 questions. Response to the Belgian census was obligatory. It always had very high response rates (both record and item response). The latest edition only lost about 4% of empty or non-returned forms.

#### 3.4.2. A preview of the register-based census

Compared to the 1991 edition, in 2001 further use was made of register data: 15 questions (mostly demographics) were now pre-coded using the population register. In the process of cleaning and coding the results the population register and the register of businesses and enterprises were called upon. It had become clear that fertility was underreported in the questionnaires, and the population register was consulted in order to determine births that had not been reported by the mothers. In another application business addresses were consulted to complement information of commuting trajectories.

At first glance, judging by the availability of basic registers in 2001, Belgium was already in a favorable position to implement a register-based census. However, in order to make it successful, setting up specific longitudinal databases (alongside with other preparations) was judged necessary. For instance, follow-up on the educational level of the population by keeping track of degrees as they were issued year by year would constitute such a database. Therefore, the 2001 census was decided to be the final one in its old format, and preparations for the transition to using register-data started shortly after 2001.

After a few years of preparatory study by a group of experts, the "Microcensus 2006" project was started up, funded by the Science Policy Public Service. This project is a test-run of the 2011 register-based census, exactly half-way between the two censuses, using a 20% sample of the Belgian Population. Within Microcensus 2006, all major registers are consulted.

Work with the actual register data has just started. Databases on education stemming from the Flemish, French-speaking and German-speaking communities are consulted in order to determine students and degrees. Data from the Datawarehouse Labour Market and Social Security will yield information on many aspects of professional life, retirement, and social benefits. The Central Population Register is consulted for studying fertility and nuptiality, age structure, household type and size. Patrimonial data on built ground parcels and their owners will be looked into regarding housing characteristics. The Register of Businesses and Enterprises yields NACE- activity sectors and additional information on institutional, commercial and private owners of real estate. Finally, information in the 2001 census will be used as a basis for updating certain variables, notably the educational level.

### 3.4.3 Outcomes and results

The registers described above contain most of the information of the classic census. However, a number of subjects are not available in any of the registers: health status (objective and subjective), receiving and giving informal care, professional training, occupation title, actual number of hours worked, some variables concerning commuting, house rent, energy consumption (heating system, types of energy, ...), connection to sewage system, physical condition of the dwelling-unit, subjective appreciation of the neighborhood, vehicle types owned are the main variables for which no suitable replacement was found in the registers.

It is far too early to judge the quality of results from this pilot-census. However, challenges in completing the project can already be identified. Two elements can be named as examples. First, reconstructing the educational level of the population is a challenge, because it requires starting from the 2001 census, and checking whether new degrees were obtained. Therefore, the continuous registration data of at least five different education systems in three language regimes must be compiled and investigated. These systems differ substantially, in terms of content as well as in terms of identification variables used.

Second, the register-based housing census is a challenge of another order. In the land-register the housing unit is not defined, and the registration system is based on legal rights (ownership or others) rather than on the practical use as a dwelling, a small business or a shop. While the classic census was focused on practical use of space and facilities in and around the dwellings, the land-register refers to the original destination of the property-unit, regardless of whether it is used as such today. This means that, although many relevant data are available, partitioning them in terms of dwelling units that can be uniquely attributed to their inhabitants is an almost unmanageable task.

To sum up, the Belgian case clearly demonstrates why a basic register of buildings and dwellings is a condition for effective use of the content of patrimonial databases. Furthermore, Finland and the other Nordic examples (in a completely different setting Switzerland makes the same case) have shown that such a register is even more valuable for administrations than it is for the register-based census and for statistics in general.



## References

Fellegi, I. a,d Ryten, J. (200). *A peer Review of the Swiss Statistical Office*, Office Fédérale de la Statistique, Neuchâtel.

Hoge Raad voor de Statistiek, 2005. *Rapport van de wetenschappelijke adviseurs over de uitbouw van een statistisch systeem ter vervanging van de Algemene Socio-economische Enquête 2001*, Brussel.

Schulte Nordholt, E., Hartgers, M., Gircour, R., (2004). *The Dutch Virtual Census of 2001. Analysis and Methodology*, Statistics Netherlands, Voorburg/Heerlen, 279p.

Statistik Schweiz, 2000. *Merkmalskatalog. Eidgenössisches Gebäude- und Wohnungsregister*. Version 2.3, Neuchâtel, 56 p.

Statistics Finland, 2004. *Use of Register and Administrative Data Sources for Statistical Purposes. Best Practices of Statistics Finland*, Helsinki-Helsingfors, 65p.