

COMBINING REGISTER-BASED AND TRADITIONAL CENSUS PROCESSES AS A PRE-DEFINED STRATEGY IN CENSUS PLANNING

Olivia Blum
Israel Central Bureau of Statistics

Abstract

Traditional and register-based censuses are two options in carry out censuses. The ability to maintain and update relevant registers leads some countries to conduct administrative censuses, while those who lack this ability have no option but to perpetrate traditional processes. Limited resources and declining inclination of the public to cooperate imposes a search for an efficient process that relies more on existing data. However, optimizing the use of resources does not mean striving for a pure register based census. Rational decision-making demands a decomposition of the main theoretical and practical components of a census, a cost/benefit analysis of each component WITHIN the established structure and with regard to the interdependence and interactions amongst its parts. The end result, in most cases, is a combination of both types of censuses rather than one or the other. Moreover, the established statistical system is not a mere census system but rather a census information system where the data are rich and detailed. But, unlike the conventional notion of a census as a snapshot of a stock, performed in defined time intervals, it is a dynamic and continuously 'breathing' statistical body, that monitors the flows and generates a sequence of frequent stock-snapshots.

1. Introduction

The process of population census-taking involves data collection from a defined population within defined geographic boundaries. These data, when processed, analyzed and transformed to statistical information, characterizes individuals and households, demographically and socio-economically, in small population groups and within detailed geographic units.

Censuses, by definition, rely on data collected from all members of a relevant population. This is done either by a direct interaction with the population, as in conventional censuses, or indirectly, by a secondary use of existing administrative files. Since there are no automatic processes of population data generation, an active operation of collection is always involved. Consequently, the question of direct or indirect data-collection is addressing in reality a broader issue of who is in control of or who is responsible for, the data. The issue is often translated to power relations when conflicting interests are introduced.

The objectives of statistical offices are formulated in terms of statistical information, while the carrying out of administrative roles leads to accumulated data produced in the form of reports and measurements as recorded during a specific, issue-oriented interaction. The derived gap between the required statistical information and the raw, register or register-like data, is the divergence point of a methodological fork that has numerous paths: one path ignores entirely the raw data and initiates target oriented projects (i.e. conventional census); another path is the option of statistical operations that rely solely on data collected by others (i.e. register-based census); while a third option includes hybrid operations of varying degrees of the two former options, dwelling on the continuum between these two extremes.

When to use what option, depends on the interface between interests and needs of the statistical and the administrative organizations, expressed in a set of attributes: population covered, type of data-collected, time reference, frequency and timing of updating, maintenance and reliability of the produced files, their accessibility and their content flexibility. Furthermore, the available technology as well as the technological horizon are major factors in defining interests that create or eradicate conflicts engendered by human entanglement in data handling. Playing a role in the overall 'data-game' is also the environmental background, i.e. legislation and public opinion.

When the needs of both, the statistical and administrative organizations, coincide perfectly, the administrative apparatus is superior (Johansson, 1991). When it is not the case, it is possible to accommodate data-collection to both purposes by an 'integrated data-collection', where the authorities collect data for statistical and administrative ends (Denmarks,

1995). However, this harmony is very fragile. It may lead to a bureaucratic constellation, where more people are needed to enable the integration, while the very same people cause the bureaucratic mechanism to be cumbersome and inefficient. Furthermore, administrative files cannot be maintained at the same quality level, with all the needed statistical parameters, unless the statistical office does it itself, not only because of different interests, but also because of differential change of interests (Laihonen&Thomsen,1998). A corporate-integrated data management (Priest,1996b), applied on the inter-organizational level may tackle this problem, but even here, it is not a solution for the long run, since it conveys blurry boundaries between different public institutions in a political power structure.

Moreover, no matter what solution is applied to enable the statistical use of administrative data, it always invokes the need to control data collection and data quality, by continuous evaluation, based to some extent, on fieldwork operation rather than on “pure” data mining. Hence, administrative data as the only input for the production of census statistical information is an unfeasible utopia.

On the other end of the continuum, when no available administrative data is useful for statistical purposes, the statistical office is entrusted with all tasks associated with the production of census information. However, this distinct situation rarely exist. The magnitude, complexity and costs of population and housing censuses make it worthwhile to use existing administrative files to support, supplement or substitute, a single element or sets of elements, in the different phases of a census. Simple count of units, their partial demographic profile or any other attribute that can be linked partially or fully, on a micro or macro level, can be used functionally for more efficient and parsimonious statistical processes. This assertion may seem to embody a contradiction: although a census applies to the whole population, a partial use of registers as well as a use of partial registers is suggested. The tendency to hold a “sterile intellectual position” in Sheuren’s (1992) terminology, makes it hard to see alternatives to conventional census taking, where partial use of registers is concerned. However, it is only a deceiving external cloak. Registers can be used for a single task, like imputation to supplement data not collected directly, or for completing under-covered population groups like young males, thus improving the overall census quality in terms of coverage and reliability. Furthermore, the registers themselves do not have to be comprehensive and full, since they can supplement census data with or without complementing each other.

While several European countries have already performed or tried to perform a register-based census, most countries over the world have usually performed conventional censuses with none or a minor use of administrative files:

Denmark and Finland have built statistical registers functioning as a census database, based on their administrative registers, by means of evaluation and modeling of estimators. Norway, Sweden, Austria, Belgium, Luxembourg and Switzerland are in a transition from conventional to register-based census, while the Netherlands is looking for another solution, other than register-based (Longva et al,1998; Laihonen&Thomsen,1998; Laihonen, 1999).

Germany and France although attempting the transition to a register-based census are still closer to a traditional one. A shift toward a register-based census as the main source of information is planned in Slovenia, while other countries, like Cyprus and the Baltic States, are trying to rationalize census taking by exploiting existing resources within a framework of traditional census.

Past censuses in Israel, although conventional, have used the population register intensively. The next census is expected to be planned under the rational assumptions of mixed processes, without deciding in advance the precedence of the source of information.

In the following sections I would like to discuss the alternatives and its main theoretical and practical components of a combined census taking, and elaborate on the decision making factors to be included in the utility function, of the sources of information to be used.

2. Census-System or Census Information-System

2.1. Changing Goals in a Changing Reality

Censuses generate detailed information of population and housing, and as such come to serve two types of users: those whose research and analysis are implemented in policy planning and making, and those whose research is an end in

itself. Both are conservative in their attitude to changes of census content, yet when expressing needs to be met, they would like to stretch the canopy and adapt to changing conditions. Although stability and consistency between censuses, as well as compliance to international recommendations, are kept for intra-national and inter-national comparison reasons (longitudinally and cross-sectional), census content is somewhat flexible and tuned to local culture and needs. However, while census content has varied, census goals have not been asked to accommodate to changing realities. Growing social differentiation and individualization implies growing complexity of the population, and a need for even more detailed information. Yet, this changing reality has resulted in decreasing willingness of the population to participate in common tasks (Germany Statistisches Bundesamt,1992). Expecting the census to be the same good old friend, solid and true, inspite of its 'ugly warts and wrinkles' (Farnsworth-Riche and Marx,1996), may lead to leaning on a broken reed. Some adaptations and alterations have to be considered.

Changes, in terms of census goals and objectives, can be either of overall strategy or of local tactics. In the supply and demand data-market, the need for change may rise from both sides. The demand as defined by the users is to be judged by their objectives and not by the data they would like to have in their offices. Sometimes they need less or different data than what they declare (Vliegen and Van de Stadt, 1989), and it is the role of the statistical office to identify the source, the type and the level of sensitivity and reliability of the data to be used for the declared purposes. Moreover, statistical offices are not demand-followers only, they may study past use and alter definitions of needs according to actual use. They may also anticipate future potential use. It means that the statistical offices have the ability and the obligation to shape the demand curve according to society needs that either have or have not been detected or articulated by the users beforehand. Changes from the demand side are usually addressed by local revision rather than by global transformation. It is more of changing tactics of investigation, by adding questions to the questionnaire or altering the configuration or substance of items that are already included in the questionnaire.

However, when the data supply is in problematic, a strategic change is called for. Limited resources and declining inclination of the public to cooperate call for a search for alternatives to data collection and processing.

Population data in all sources of information, derive directly or indirectly from the population. Indirect data are defined as such when they serve for a secondary use, meaning that they have been collected for different purposes. In the census arena, indirect data are administrative files, subject matter surveys, and censuses whose units are not individuals or households (agriculture census and such).

2.2. Indirect Data-Collection Supporting a Conventional Census

In most countries it would not be possible to stay in pure conventional census, based on direct data collection, because of the expanding needs for census information, increasing costs in absolute terms and per capita, and decreasing public cooperation (Schueren et al, 1992; Laihonon&Thomsen,1998). Furthermore, census operation is lengthy, and the data is provided in long intervals which is compounded by additional time lags between collection and dissemination. Thus, data collected for non census purposes are introduced to the census process, gradually.

Registers have already been used to improve coverage before, during and after the enumeration. Addresses known beforehand serve to prepare maps, enumeration routes and mail-lists, and to control coverage during data collection. Individual records may help just the same, to allocate reasonable enumeration portions to each enumerator, to pre-print information on the questionnaires or as a check-list during data-collection.

Registers can be used to reduce the data capture workload and to improve it in optical data entry system, by linking and comparing individual records in the register with the optically identified values of the census (Blum,1997).

Moreover, since in most censuses the socio-economic questions are addressed to a sample, and non-response is of homogeneous groups, registers are used to improve the quality of the data by serving as a sampling frame and for editing and imputation procedures. This type of use of registers reduces biases due to non-response and improves the estimates provided. It avoids the single-source output bias (Germany,1989; Harala,1996; Heihonen& Laihonon,1987; Huggins& Fay, 1988; Priest,1996a; Thompsen et al,1996). Registers are also used to increase the number of observations needed for small area and small population-groups estimates (Schaafsma-Harteveld, 1999; Slagter,1999; Leggieri, 1999).

In post-census activities, administrative files are used for evaluation purposes, as one among several sources of comparable data.

The above uses enable the reduction of direct data collection by addressing fewer questions to fewer people.

2.3. Direct Data Collection Supporting a Register based Census

In most countries, a full register-based census is not a feasible option either. This is because of legislation constraints or limited available sources of information, originating from and perpetuated by the lack of control over these sources.

Registers have to be evaluated on a constant basis, their content and coverage have to be adjusted to census purposes and their quality has to be kept.

Harald (1999) suggests that in the pursuit of high quality, the main role of Statistics Norway is to identify errors in the registers and to inform the authorities. Twenty percent of the total costs of the 2000 census in Norway, are allocated to the improvement of the registers.

Another aspect of the support needed for a register-based census, is the addition of variables that are not included in the existing registers. Several methods to collect the crucial variables are suggested by the Nordic and Benelux countries including:

integrated data collection in which data is collected for statistical reasons as well as administrative ones by the administrative authority,

random surveys with the possibility of attaching on ongoing surveys, designated surveys to target populations, and rolling sample surveys.

In addition, these countries suggest building new registers and conducting partial censuses, where limited issues are addressed to the whole population.

2.4. Toward a Census Information-System

In the second half of the 20th century, the evolving pattern of census alternatives that rely on indirect data collection, has been planned and perceived as a replacement of the traditional census processes. This limited perception ignores the wide spectrum of the new possibilities that the multiple source data enable. It should be considered as a bedrock for potential change in census objectives while developing and extending statistical options, and not as a mere replacement of the traditional census. Administrative censuses have not yet proven themselves to be a pure model of a secondary use of existing data. As a result, seeing the administrative census as a substitute is a source of new problems. Countries that have been trying to shift to a register-based census report problems of coverage and content deriving from the gap between interests: administrative files cover interest groups rather than the whole population, and variables of interest to the administrative authority are not necessarily variables of interest to the census information users.

The ideal situation is not to have a complete register-based census as a final objective, but rather to optimize the use of available sources in order to avoid the faults of each and to take advantage of the merits of each. In such a setting, the census becomes just one of several sources (Longva et al,1998), in an all-embracing statistical system, whose life span depends on continuous activities of data-collection, evaluation and processing. This within an environment of accelerated transformation of: society and social values, economy and economic capabilities, policy and its implementation in the political arena, and of present technology and technological horizon. It is a shift from a census system to a census information-system where the data are rich and detailed. But unlike the conventional notion of a census as a snapshot of a stock, performed in a defined time interval, it is a dynamic and continuously 'breathing' statistical body, that monitors the flows and generates a sequence of frequent stock-snapshots.

The idea of a supported census, by either conventional or a register-based one, means that although each process and sub-process of both options has its own merits and liabilities, decision making is usually based on choosing one way census. The recruitment of supporting elements of the alternative census process, is introduced only when a problem is

detected. This decision making process cannot be defined as a pure rational one, but rather as a bounded rationality that was pre-selected as such. The ideal-type of a rational decision-making relies on the decomposition of the main theoretical and practical components of a census, and on a cost/benefit analysis of each, WITHIN the established structure and with regard to the interdependence and interactions amongst its parts.

3. Building Blocks of the Decision Making Process

The main building blocks to be taken into account of in the decision making process are presented in the diagram, followed by a discussion of selective components.

The idea is that rational decision making means weighing the pros and cons of the use of different sources of data, on a micro as well as macro level. This should be done while taking into account the different interests and the differentiation of interests between data collectors, along time, and considering the alternatives not as enumeration of people vs. enumeration of files, but rather as a combination of both.

Building Blocks of the Decision Making Process ([See page 13](#))

3.1. Legislation

Statistical offices draw their legitimacy and power from laws specifically legislated for their functioning. However, these laws also seek to protect people from the intrusion to their privacy and from the violation of their basic human rights, by the very same agencies who are endowed by law. This tension exists throughout the census activity and beyond; in data collection, processing, dissemination and actual use. When the use of multiple source data for census purposes is introduced, a new set of legal questions and derived legislation, follow suit. They involve the statistical bureau's right to:

1. get and use, for statistical objectives, data collected for other purposes;
2. influence the data collected by other agencies;
3. build up and maintain registers;
4. add the same unique identification number to each record in all registers;
5. initiate designated statistical operations in the field (surveys or census-like);
6. inter-link different sources of information;
7. produce integrated statistical information;
8. pass on integrated information;
9. allow each individual the access to his/her personal information;
10. and address security and storage issues.

Answering the set of questions pertaining to legislation issues is a prerequisite for multiple source census. However, positive answer is not required en-bloc but can be solved selectively. For example, aggregate statistical linkage is an option if identification numbers are missing, complementary fieldwork operation is a valid alternative when registers are missing and are not allowed to be built, and so on.

Yet, this logic, where most components are intertwined with each other, possibly serving as partial alternatives to the same end, implies a complex inter-dependency in which changes in one component affect others, causing the need for contingency plans.

Regulations have a stabilizing effect in a seeming wobbly situation, however, the complexity of the system makes it sensitive to changes in the regulations themselves. Laws and regulations are needed not only to exploit administrative data but also to negotiate standards and influence content of the registers (Priest,1996b; Thomsen et al,1996). Although all the agencies are under the government umbrella, the overall pyramid is an hierarchy of prestige rather than of legitimate authoritarian relations. The statistical system's laws, structure the power relations in the absence of hierarchy.

Furthermore, different sources of data implies a need for harmonization among registers and therefore a new statistics act for better coordination (Longva et al, 1998). However, changes of laws that affect the data sources and the

accessibility to them, are not rare and “there are too many last minute surprises” (Spieker,1999). Consequently, the frequency and timing of possible changes of laws and regulations are parameters to be taken into account in a contingency plan.

Changes in government regulations and policies, not directly related to censuses, can also have an impact on the usefulness of the administrative data for statistical purposes. For example, taxation may reduce coverage while benefits may cause over coverage. The susceptibility to regulators’ whims implies the need for not only flexible plan under uncertainty, but also for shortening the census planning phase to the minimal time required and for an attempt to include time horizon in the ordained laws. This is more feasible in census related legislation, where the duration of the validity of the act can be specified. It is not the case with laws that have indirect impact on the quality of the registers used. Knowledge and follow up of these regulations becomes an issue in itself.

3.2. *Public Opinion and Public Behavior*

Conventional censuses face a growing objection, by the population, to answer the census questionnaire. At times it is a refusal to selected questions, at times to the whole questionnaire, and face to face interaction is rejected altogether. Explicit explanations given to this behavior are social *alienation* derivatives (‘who cares’, ‘it’s none of your business’ etc.). They also express *skepticism and mistrust* about census goals, that leads to disbelief in the statistical office motives, disbelief in information disclosure to other government authorities, like the IRS, and to speculations that decrees, like per-capita taxes, are to be expected.

However, objections to direct data collection are supported by a non ideological reasoning too. In the 1995 census in Israel, one of the most frequent arguments was that the Bureau asks questions that the government authorities already have the answers to; all the demographic variables are in the population register, number of cars are in the files of the motor vehicles authority, income is reported to the IRS and the to Social Security Institute, etc.. This *functional* argument that the data are already in the Bureau’s possession has to serve as a beacon; if a rational use of all data sources is not the main census apparatus, the public will force it on the office. Waiting for the public to initiate the use of existing sources, will likely result in a dent in the public relations mantle of the statistical office. It may also create a timing disadvantage, by forcing the adjustment when the office is less prepared for it.

In spite of the above arguments, a use of available data sources is not problem proof when public opinion is concerned. The very same public does not accept easily the idea of record linkage between registers to replace the conventional census. Placing all data in the hands of one agency, on a continuous basis, generates fear of a Big Brother syndrome. The ability to reveal phenomena never intended to be revealed, by adding up the person’s characteristics, given for mutually exclusive reasons, to separate agencies, stirs emotions in itself. In countries like Sweden, the Netherlands and Germany, an objection to this idea restricts or even bans administrative census operations. When balancing problems of public acceptance on one side and problems of response burden on the other, the verdict in different national settings is not clear cut. An Orwellian situation may be evaluated as a dead end, as far as the use of registers for census purposes is concerned, but it can also serve as a starting point for creative solutions.

Another aspect of public behavior is the quality of the registers. Cultural differences between societies and their economic development hint both at the inclination of the public to report information and the quality of this reporting. In a dynamic society, where data in registers have to be updated often, these tendencies play a major role in determining the registers’ quality and its relative advantage as a source of statistical information. When economic or political interests depend on the register content, it should be expected to contain passive and active errors: changes are not reported on time or not reported at all, and false reports are introduced. Since only rapid and reliable reporting routines lead to updated files, registers are of different quality levels, and related to interests involved.

It is not clear if a direct data collection is a more reliable source of information. The very same considerations that led to problematic administrative files may lead to biased answers to questionnaires or enumerators, when the respondents do cooperate. Direct interaction with the data collector may produce additional unidentified bias. It means that even if the public is persuaded that the census bureau has the public interests as its first priority, and that the social and financial costs generated by the data collection process is minimal, quality of the data is still not guaranteed. Having absolute true values is not a feasible option because of the subjective human involvement, expressed by potential and actual

behavior.

The social contract between the citizens and the government authorities, anchored in a web of laws and regulations, is the bedrock on which the statistical system is positioned and from which it derives its potential abilities. This social contract is a function of pre-defined basic rules and of regulated changing norms. As such, it is negotiable and allowing for setting up a new statistical system, where the direct and indirect interaction between the census bureau and the public is multifaceted.

3.3. Adjusting Needs

Statistical offices, when planning a census, try to figure out what are the needs of the main users. These declared needs are sorted to those to be answered in the census, those that can be answered in surveys, when the induction unit is rather large or when in-depth investigation is required, and those that may hamper the census by evoking hostility. This logic of acceptance and rejection of requests is not extremely violated when multiple-source data are used. The additional decision determinant is 'what's available', and at times it provides the users with data that would not have been collected, either because of limited interest or because of the problems their collection could have caused.

Nevertheless, the issue of marginal cost of collecting data is exerted in a multiple source census. It may be relatively high when a unique process is introduced. In Norway, where censuses rely on administrative files, there is still a need for direct data collection. These direct data, while only 7% of the total data, costs 30% of the total expenditure (Thomsen et al,1996; Laihonen& Thomsen,1998). As a result, a better screening of needs has to be incorporated. Determining needs is done on the basis of users' objectives, the ability of the statistical office to supply the relevant data, by past use and, in the multiple source setting, by the availability of close alternatives. The differentiation between basic primary data, and secondary, usually specific data, becomes more crucial in the decision making process.

Furthermore, multiple source data often draw along with requested variables, a tail that provides additional variables, enabling an analysis of unusual profiles of the population. This is the contribution of the registers for better and wide range statistics in censuses. Redfern (1989) sees in it administrative advantages that lead to a fairer society. More information means a more equally divided social benefits and social burden.

Another aspect of creating such a system is the stimulus to develop new needs. Using as a metaphor the idea that the boundaries of the language are the boundaries of the thinking, one may say that the accumulated and linked data, originated from multiple sources, serve as an enriched language that expand the world of possible needs. These new needs are not expected to be answered by existing data for the long run, and may initiate a new cycle of adding information sources for identified needs and merely by doing so, creating new needs.

In the world of census information systems, uses as well as the number and types of users are expanding. Census data reflects a stock of population and housing characteristics once a decade, while a census information system is a breathing system, and as such, may provide data of flows and accumulated stocks on a continuous basis. Users within the statistics system can shift from partial, narrow, designated, local systems to the census information one, and produce demographic and other continuous estimates in short intervals.

3.4. Quality

3.4.1 . Coverage Quality

The issue of coverage in censuses is addressed on several levels: the geographic boundaries of the census area, the definitions of the census population, individuals and families or households as the basic analysis units, buildings (addresses) and dwelling units to enable a scrupulous spatial analysis, and the definitions of geographic and administrative divisions. These definitions of the coverage-units can be altered involuntarily, when sources of data are added or changed. Flexibility of definitions is a prerogative of traditional census takers, whose information is obtained directly from the population. Administrative records are not as flexible nor amenable.

When multiple data sources are introduced and enumeration of people is combined with enumeration of files, compromises with regard to the census units and their definitions, are only pragmatic. In addition, there is an

interdependency between the coverage units, where a decision with regard to one influences the mere definitions of the others. For example, if coverage is expressed in 'individual' terms and not households, housing may refer to buildings and not to dwelling units. The definition of the target unit dictates the substantive issues in mixed census taking, and may be one of the factors that causes a fallback from the idea of having a full register-based census, as Germany experienced (Vliegen&Van de Stadt,1989).

In conventional censuses, under-coverage originated in the office or in the field, is problematic on most unit levels. In rare occasions, where political or geographic boundaries are not well defined or are not agreed upon, perceived interests may result in local over-coverage, as happened in East Jerusalem in the 1995 Census in Israel. The use of administrative files is a practiced option to control coverage of addresses and individuals, and to evaluate it. When enumeration of files is the main mechanism, surveys and partial censuses supplement or substitute registers whose coverage quality is low, and the costs of these registers use are high. In both directions, the quality of the available files and the dosage of the concoction of registers and fieldwork operations, stipulate the flexibility of alternative definitions of the coverage units and the quality of the coverage.

Quality of coverage in administrative files is a function of their administrative role, interests involved, scope and frequency of actual changes, frequency and timing of reporting, and the quality of their maintenance. Incomplete coverage of administrative files is immanent because of their different administrative functions. For example, the files of the social security institute covers tax paying residents and those who are entitled to social benefits, while police registers cover alleged and convicted criminals. Interests may create bias of coverage in the direction of the interests, but it may also improve coverage of specific groups. When it is not rewarding to be registered, under-coverage might be the dominant phenomenon.

The ability to maintain and update a register is conditioned by prompt reporting of the public as well as a smooth and timely data capture. However, a time lag is to be expected since the process is ex-post. In a dynamic society, where changes in the population and its attributes are frequent or in large volumes, these difficulties are accentuated.

Beyond that, registers present a formal/legal picture, while the demographic 'black market' includes special population groups that are missing from the register but are of interest to the census takers (mainly foreigners and other groups at the fringes of society). Hence, direct data collection is required in some time intervals, while statistical registers, on which statistical estimates are based, provide a solution for the interim.

3.4.2. *Item Content*

Item content considerations in conventional censuses concerning the inclusion or exclusion of an item, are determined by needs, the quality of the answers that can be generated from the public and by the different constraints, like questionnaire scope, effectiveness and costs. Adding administrative files as a data source opens the floor to a new set of considerations to be made. The quality of the variables in the administrative file is, as in the coverage issue, a function of its administrative role, interests involved, scope and frequency of actual changes, frequency and timing of reporting and the quality of its maintenance. However, when content is concerned, the extent to which registers are used in conjunction with fieldwork operation, determines not only the character but also the scope of the difficulties, as more registers implies less flexibility and less control over census item content. Variables found in registers and administrative files are not always the exact census ones, and even if the concepts are the same, their definitions may be varied. Multiple source data sustain problems of terminology and of theoretical definitions that are actually problems of content and comparability, between sources of information and between censuses. In the long-run, even if concepts remain the same, their definitions may not (Borchsenium,1996).

In most cases, the combination of administrative sources is not a full substitute to direct data collection; some variables are hard to get and some are impossible to get (Laihonen&Thomsen,1998). Adding hard to get variables to surveys, causes an increase in the response burden (Slagter,1999), and may have an overall negative effect on the survey. Yet, the use of administrative files provide the users with rich and detailed data that cannot be ignored when item content is concerned.

Hence, the combination of direct and indirect data collection is preferable on the use of each separately, because of the limitations of each and because of the relative advantages of each. The exact portions in this combination is a derivative

of the availability, quality and costs of the components involved, and it differs from one state to the other.

3.5. Time

The use of multiple data sources like administrative files, partial censuses and surveys, sets up a different viewpoint of the time reference of a census, time intervals between censuses and the character of the data with regard to measuring a stock vs. a flow.

The conventional census measures a stock of characteristics of the population in a certain point of time, usually referred to as a snapshot of the population. Partial censuses and designated surveys do not differ from the conventional one in this respect.

Indirect data collection relate to the reference day as the date up to which the file is updated. The practical implication is that process of data collection from administrative files may be a long one, since different administrative files have a different updating pace.

Ongoing surveys and rolling sample surveys are challenging collection methods with respect to time reference. Ongoing surveys usually relate to the week the sampled unit is encountered, while rolling sample surveys aim to cover a portion of the population each year. If it is a 10% sample, the whole population is covered along a decade. In a census information system, because of the dynamic nature of the data, two solutions are possible: Changing the idea of time reference, or changing the scope of information generated at a defined time period.

When data are kept updated on a continuous basis, one can generate information of a flow, and when accumulating changes, also of a stock. In a system that relies more on direct data collection or when the administrative registers are not accessible on a continuous basis, a widening gap between two types of stocks can be expected; a stock that is a result of accumulated changes and an actual stock measured in time intervals. In such a case, a census action is required. It can encompass all census subjects, or a selective part of them in short time intervals. For example, the statistics bureau can produce annual estimates based on the flows, a demographic census, once in two years, and a full census, once a decade. Maintaining and updating this kind of system is a function of the timing, the type (and scope) of the census and the intervals defined. The idea of rolling samples can slide easily to such a system, to the partial census part. Scheuren et al (1992) suggest that measuring along a decade might be found better than measuring once a decade, since the phenomena measured are changing all the time. However, censuses with rolling samples are partial because of the population included and not because of the selective subjects investigated. It means that a representing national picture is obtained once a decade and not once in two years.

3.6. Technology

A census information system with multiple data sources is technological tools intensive. The integration of data sources as well as data warehousing, data-mining and data-retrieval abilities, are the core of the system. Data collected by different methods (questionnaire, fax, voice etc.) have to be captured and linked in different methods. The accessibility to the data is a challenge in itself; it has to be easy, flexible and fast in large and diverse databases. However, technological improvement and development are in accelerated path and what seems to be pretentious today may seem as a simple task in a very short period of time. Therefore, not only existing enabling technology is a decision factor in planning a census, but also the

technological horizon.

Furthermore, the interaction between census goals and available or possible technological tools, affects both ends. The idea of a breathing census information-system as a goal, instead of a census system, is a cause and a result of technological abilities.

On a micro level, the technological working environment allows for a creative thinking as far as census tasks are concerned. For example, when the conventional census, as a provider of sampling frame, is missing, the geographic information system provides a substitute, based on concepts and tools of a different dimension; The land becomes the sampling frame and an area, defined by grid-coordinates, turns to be the sampling unit.

All in all, the role of technology as a major decision determinant in generating census information, is growing and becoming more important with time.

4. Concluding Remark

In different national settings, census sources of information are not similar in their quality, accessibility and the costs of their use. Countries find themselves performing mixed censuses that rely on a direct data collection while indirect data collection is the main data source of others. However, this should not be a result of a predisposition to have a traditional or a register-based census, but rather a result of a rational decision making process in which the use of each data source and its subsequent structure are weighted.

Bibliography

Belgium, National Statistical Institute 1999a

“Possibilities of using administrative registers to perform the population and housing census in Belgium”. Working paper No. 21. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Blum, Olivia. Israel Central Bureau of Statistics 1997

“Keying Module” in Euro-Med New Technologies for the 2000 Census Round: Euro-Mediterranean Workshop. Israel.

Borchsenium, Lars Statistics Denmark. 1996

“From a conventional to a register-based census of population”. Working paper No.9. SCECE Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Denmarks, Statistik 1995

Statistics on Persons in Denmark: a Register-based Statistical System. Eurostat and Denmark Statistik. Luxembourg.

Farnsworth Riche, Martha and Robert W. Marx US Bureau of the Census 1996

“Census 2000: will you recognize an old friend?”. Working Paper No. 38. SCECE Work Session on Geographic Information Systems. Washington DC.

Germany, Statistisches Bundesamt 1992

“Considerations of alternatives to censuses and census-type statistics: the case of Germany”. Working Paper No. 13. SCECE Work Session on Population and Housing Censuses”. Geneva.

Germany, Federal Statistical Office 1989

“Prospects for replacing population and housing censuses, either totally or partially by surveys and administrative registers”. Working Paper No. 14. SCECE Seminar on the relevance and importance of population and housing census data. Wiesbaden, Germany.

Harala, Riitta Statistics Finland. 1996

“Continuous quality assessment of the register based census and regional employment statistics”. Working paper No. 6. SCECE Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Harald, Utne Statistics Norway. 1999

“Population and Housing Censuses in Norway toward a register based solution”. Working paper No. 3. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Heinonen, R. and A. Laihonon Central Statistical Office of Finland 1987

“New approaches to the production of census data: Finish experiences from the 1985 census”. Working Paper No. 12. SCECE Seminar on computer-related aspects of population and housing censuses. Belgrade, Yugoslavia.

Huggins, Vicki and Robert Fay 1988

Administrative data in SIPP Longitudinal Estimation”, American Statistical Association Proceedings.

Johansson, Sten Statistics Sweden. 1991

“Statistics based on administrative records as a substitute or a valid alternative to a population census”. Invited paper 11.2. 46th Session of the ISI.

Leggieri, Charlene US Bureau of the Census. 1999

“Uses of administrative records in United States Census 2000”. Working paper No. 5. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Laihonen, Aarno Eurostat. 1999

“Development of the use of administrative data in population and housing censuses in Europe”. Working paper No. 6. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Laihonen, Aarno & Ib Thomsen . Statistics Finland and Statistics Norway. 1998

“Interim report of the project on reducing costs of censuses through use of administrative records” in the “Final Report From the Development Project in the EEA: on Reducing Costs of Censuses Through Use of Administrative Records”. Laihonen, Aarno, Ib Thomsen and Britt Laberg. Statistics Norway and Statistics Finland.

Longva, Svein, Ib Thomsen, and Paul Inge Severeide 1998

“Reducing costs of censuses in Norway through use of administrative registers”. International Statistical Review (1998) 66, 2. pp. 223-234.

Priest, G. Statistics Canada. 1996a

“Issues of meta information and integration”. Working paper No. 2. SCECE Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Priest, G. Statistics Canada. 1996b

“Challenges and opportunities in administrative records”. Working paper No. 4. SCECE Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Redfern, Philip 1989

“Population registers: some administrative and statistical pros and cons” . The Journal of the Royal Statistical Society A (1989) vol. 152, part 1 pp. 1-41.

Schaafsma-Harteveld, Berna Statistics Netherlands. 1999

“Disablement benefits: combining survey data with register records”. Working Paper No. 20. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Scheuren, Fritz, Wendy Alvey and Beth Kilss 1992

“Paradigm shifts: administrative records and census taking”. Selected papers given in 1990 at the Annual Meetings of the American Statistical Association. Department of Treasury, IRS. Publication 1299.

Slagter, Herman C.A. Statistics Netherlands. 1999

“Compiling structure of earning statistics using existing survey data and register data”. Working Paper No. 7. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Spieker, Finn Statistics Denmark 1999

“Formation of central variables in a decentralized statistical system”. Working Paper No. 25. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Thomsen, Ib, Elisabetta Vassenden and Britt Laberg Statistics Norway. 1996 (1997)

“Availability and use of administrative record systems in the ECE region”. Working paper No. 7. SCECE Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.

Van de Stadt, Huib and Mathieu Vliegen Netherlands Central Bureau of Statistics 1992

“An alternative for the census? the case of the Netherlands”. Working Paper No. 12. SCECE Work Session on Population and Housing Censuses. Geneva.

Vliegen, Mathie and Huib Van de Stadt Netherlands Central Bureau of Statistics 1989

“Is a census still necessary? Experiences in the Netherlands”. Working Paper No. 5. SCECE Seminar on the relevance and importance of population and housing census data. Wiesbaden, Germany.

Building Blocks of the Decision Making Process

