

# Register Surveys – An Introduction

Three types of statistics based on microdata are published by national statistical offices – statistics based on *sample surveys*, statistics based on *censuses* and statistics based on *administrative registers*. This book deals with the third type, statistics based on administrative registers, where instead of collecting data through sample surveys and censuses, administrative registers from different sources are adapted and processed to make the data suitable for statistical purposes. This kind of survey is called a *register survey*.

We introduce a number of concepts and principles that are used when discussing register surveys. These concepts and principles form the basis for a theory of this type of survey. We primarily discuss register surveys at national statistical offices. There is growing interest in this area; many countries increasingly use administrative data for statistical purposes, and there is a growing demand for a theory of register surveys.

## 1.1 The purpose of the book

Our main purpose is to describe and explain the methods that should be used for register surveys. Conducting a register survey means that a new *statistical register* is created with existing sources. The statistical register is then used to produce estimates required for the survey. What methods should be used in creating such a statistical register? One or more administrative registers are used when a new statistical register is created and the statistical register can differ from the administrative sources in many ways.

A *system of statistical registers* consists of a number of registers that can be linked to each other. In the Nordic countries, the national statistical offices have developed systems of registers that are used in the production of statistics. When new statistical registers are created, this register system becomes an important source that can be used together with different administrative sources. Another purpose of the book is to explain how such register systems should be designed and used in the production of statistics.

When a national statistical office starts using more and more administrative sources, the *statistical production system* of that office will gradually change. From a system based on enumerators or interviewers, address lists or maps, the system will become increasingly register-based. Sample surveys will be based on the

Population Register or the Business Register instead of address lists or maps – variables in sample surveys can come from administrative registers as well as from telephone interviews or questionnaires. In addition to the change in methods used for sample surveys, new kinds of register-based statistics can also be produced. A third purpose of the book is to explain how administrative registers can be used to change the statistical production system of a national statistical office to improve cost efficiency and statistical quality.

### ***Preconditions in different countries***

The Nordic countries started to use administrative registers during the 1960s when paper-based administrative registers were transformed into computer-based flat files. The preconditions for using administrative registers for statistical purposes were good. This explains why the Nordic statistical offices now have access to large amounts of administrative data,<sup>1</sup> and that the quality of these data is high in comparison with most other countries. Consequently, it has been possible to create statistical register systems that have made statistics production efficient and even to conduct completely register-based population and housing censuses. Identifying variables as identity numbers for persons and enterprises have high quality and deterministic matching is therefore easy.

The preconditions for using administrative data in many countries are today not as good, and changing the production system into a register-based system will take many years. During that period, administrative systems will gradually be improved, so many other countries will be able to use administrative data efficiently in the future. Therefore, a clear understanding of the Nordic experiences from the beginning will facilitate development in new register countries.

However, we also discuss problems that arise in statistical offices in countries without the same preconditions. In North America, there is another tradition of working with administrative data. When identifying variables are of lower quality and coverage of administrative systems is poorer, methods have been developed for linking records and estimating population size that are important to use under these circumstances.

Our aim is to present statistical methods and principles of general interest, and we rely mostly on experiences and case studies from Statistics Sweden to illustrate these general methodological issues. As a complement to this aim, we also present some cases from new register countries that have recently started to develop register-based statistics.

We started writing books on register-based statistics during the 1990s, and during these years we have had access to registers and colleagues at Statistics Sweden. This access to a fully register-based production system has been vital for analysing and discussing register-based statistics.

Case studies are essential – in a book on register-based statistics we cannot present ideas with formulas as in books on sampling theory. We use case studies based on real data and charts with small miniature registers to illustrate register-statistical methods and quality issues.

---

<sup>1</sup> About 99% of the microdata stored in Statistics Sweden's databases come from administrative registers.

## 1.2 The need for a new theory and new methods

Sample surveys are based on methods that have been derived from an established theory – *sampling theory*. This theory has been developed within the academic world and statistical offices, and consists of terms and principles that are generally well known. Scientific literature and journals develop and spread the methodologies for sampling and estimation. Because the terms and principles are well known, people working with sample surveys can easily communicate and exchange their experiences.

Censuses with their own data collection are based on a long tradition of population censuses and the collection of data from local authorities, schools and enterprises. Measurement errors, design of questionnaires and nonresponse are methodological issues that also apply to sample surveys. Censuses and sample surveys are closely related in terms of methodology – censuses are often considered as special cases where the sample is the entire population.

Although register-based statistics are a common form of statistics used for official statistics and business reports, no well-established theory in the field exists. There are no recognised terms or principles, which makes the development of register-based statistics and register-statistical methodology all the more difficult. As a consequence, *ad hoc methods are used instead of methods based on a generally accepted theory*.

One important reason for this shortfall is that the subject field of register surveys is not included in academic statistics. Statistical theory within statistical science is understood as consisting of probability theory and statistical inference. Sampling theory is included within this theoretical school of thought, but register surveys based on total enumeration are not.

Unfortunately, statistical science has so far not included any theory on statistical systems. Statistical offices, larger enterprises and organisations do not often carry out separate surveys. It is more common that *statistical information systems* are built, which constantly generate new data. A statistical theory is necessary to describe the general principles and to develop the conceptual apparatus for such statistical information systems. Register surveys should be included in this theory. We formulate four basic principles for using administrative registers (Chart 1.1).

**Chart 1.1 Four principles for using administrative registers for statistics**

### ***Transformation principle***

Administrative registers should be transformed into statistical registers.  
All relevant sources should be used and combined during this transformation.

### ***System principle***

All statistical registers should be included in a coordinated register system.  
This system will ensure that all data can be integrated and used effectively.

### ***Consistency principle***

Consistency regarding populations and variables is necessary for the coherence of estimates from different register surveys.

### ***Quality principle***

The register system should be used for quality assessment of statistical surveys based on microdata comparisons with other surveys in the production system.

We use these principles in the book and gradually introduce the register-statistical terms that are needed for the discussions.

Chart 1.2 illustrates the present situation. Estimates from four different surveys are compared, and these comparisons show clearly that the systems approach often is missing in the work with statistical surveys. People are fully occupied with their own surveys and different surveys are also published at different points in time. As a rule most estimates are unique for one survey, but in Chart 1.2 we have found one identical variable and created the table with corresponding estimates from each survey. If we look at one survey at a time, we do not see any errors except for the sample survey in (4) where we have margins for the sampling error. But when we look at the four surveys together, we understand that there must be more serious errors in these surveys. We thus need *a theory for systems of surveys and new methods for quality assessment*. We return to this example in later chapters.

**Chart 1.2 Employees by economic activity, November 2004, thousands**

| <b>Economic activity</b>         | <b>Business Register</b> |              | <b>Employment</b> | <b>Labour Force</b> | <b>Survey</b> |
|----------------------------------|--------------------------|--------------|-------------------|---------------------|---------------|
|                                  | Enterprises              | Local units  | Register          |                     | Error margin  |
|                                  | (1)                      | (2)          | (3)               | (4)                 | (5)           |
| Agriculture, forestry, fishing   | 35                       | 37           | 37                | 26                  | 5             |
| Mining, quarrying, manufacturing | 688                      | 636          | 717               | 640                 | 23            |
| Electricity, gas and water       | 21                       | 22           | 28                | 29                  | 5             |
| Construction                     | 197                      | 209          | 215               | 199                 | 14            |
| Wholesale and retail trade       | 456                      | 453          | 484               | 456                 | 20            |
| Hotels and restaurants           | 89                       | 93           | 99                | 106                 | 10            |
| Transport, communication         | 240                      | 242          | 243               | 236                 | 15            |
| Financial intermediation         | 83                       | 77           | 85                | 78                  | 9             |
| Real estate, business activities | 457                      | 524          | 457               | 470                 | 20            |
| Government                       | 139                      | 215          | 239               | 230                 | 15            |
| Education                        | 382                      | 408          | 431               | 462                 | 20            |
| Health and social work           | 836                      | 684          | 675               | 675                 | 24            |
| Other service activities         | 142                      | 163          | 175               | 168                 | 13            |
| Unknown activity                 | 0                        | 0            | 38                | 4                   |               |
| <b>Total</b>                     | <b>3 763</b>             | <b>3 763</b> | <b>3 924</b>      | <b>3 778</b>        | <b>43</b>     |

Why are there such large differences between the surveys? The estimates for mining, quarrying and manufacturing can be 636 or 717 thousands – the inconsistencies are more serious than the sampling error. The methodological work should consist of three steps: compare surveys and find errors and inconsistencies; find out why we have these inconsistencies; and finally, reduce the errors and inconsistencies.

Chart 1.2 also illustrates that we only have one established way of giving a numerical description of the quality of published estimates – margins for the sampling error. There is no commonly used way of describing the quality of register-based statistics. However, the non-sampling errors of sample surveys are as a rule not described in the same clear manner as the sampling errors; here we also lack methods for giving a numerical description of the quality of published estimates.

In 1995, Statistics Denmark published *Statistics on Persons in Denmark – A Register-based Statistical System*. The Danish book presents a systematic review of register-statistical work and describes how to design a well-prepared register system. The book was the first attempt to create a theory for register-based statistics and to describe the methods that are used. We build on and add to that work in this book.

### 1.3 Four ways of using administrative registers

When a statistical office plans to use administrative registers for statistical purposes, the office faces a survey design issue. How should the new sources be used? How should the existing surveys be modified or reduced? To answer these questions the administrative sources should be analysed by experienced subject-matter specialists and methodologists with a good overview of the production system.

An administrative register or source can be used in four different ways:

#### 1. *Completely alone.*

If the source has good coverage and the variables in the source are of good quality, then the source can be used alone for producing statistics. Trade statistics based on only administrative registers with monthly data from Customs are an example of a source that many countries use alone for statistics production.

#### 2. *Alone, but combined with a base register.*

The Population Register and the Business Register are two important base registers that are used for all surveys regarding persons or enterprises in the Nordic countries. Base registers are discussed in Chapter 5. If an administrative register or source is combined with a base register, the quality can be improved and controlled. It will then be possible to produce consistent register-based statistics.

The base register contains important classification variables that can be used together with the administrative source. The Annual Pay Register in Section 1.5.4 is an example of using a source in this way.

#### 3. *In combination with a base register and other administrative registers.*

In many cases an administrative register does not have sufficient coverage and the variable content is too limited. Then it is not advisable to use the source alone for statistics production. But if many sources are combined, it may often be possible to use the combined data set to produce register-based statistics. We mention two examples of this kind.

*Example:* In the Swedish Income and Taxation Register of persons, about 30 different sources are used regarding different kinds of income. If all these different kinds of income are combined, it is possible to create *disposable income* of good quality for all persons.

*Example:* A business register at a national statistical office is based on administrative sources. With five sources we created a Business Register for Sweden containing all enterprises active during a specific year. Each source consists of the legal units in one taxation system. In the table below, undercoverage and overcoverage of the sources are compared with our final Business Register. The

administrative object sets in each source are adequate for each of the five taxation systems. Taken alone, each source is of low statistical quality; however, if all sources are combined, the coverage is good.

**Over- and undercoverage in five administrative sources, per cent of all legal units**

|               | Source 1 | Source 2 | Source 3 | Source 4 | Source 5 |
|---------------|----------|----------|----------|----------|----------|
| Overcoverage  | 41%      | 0%       | 0%       | 0%       | 0%       |
| Undercoverage | 21%      | 74%      | 74%      | 30%      | 9%       |

*4. To improve other surveys, i.e. to improve the production system.*

*Example:* There was no information on economic activity for some small enterprises in the Business Register. In the yearly income tax returns from small enterprises, there is text information from the enterprise that describes economic activity. This text was automatically coded into economic activity. In this way the yearly income tax returns were used to improve the Business Register.

In the Nordic countries, most register surveys use a base register as in 2 and 3 above. New register countries that have not yet developed good base registers will start with register surveys of the simple kind as in 1 above. When base registers have been developed, it will be possible to create register surveys according to 2 and 3.

## 1.4 Preconditions for register-based statistics

Preconditions differ between countries for sample surveys, censuses and register surveys; hence, the preconditions for statistical methods are different. The choice between cluster sampling and one-stage sampling depends on whether you have a Population Register or if you must use address lists. Regression estimation and calibration are methods that depend on the number and quality of available register variables. This means that an *increased use of administrative registers will change the preconditions for all kinds of surveys*.

For register surveys, the differences between countries are even more significant. Legislation on national registration and the taxation of persons and enterprises determine the character of the administrative systems that are used in each country. The legislation regarding statistical production and protection of statistical data also differs, and as a consequence certain methodological issues are important in some countries but not in others. The two main preconditions for using administrative registers for statistical purposes are stated in Chart 1.3.

**Chart 1.3 Two preconditions for using administrative registers for statistics**

***Identity number principle***

Unified systems of identity numbers are used in all administrative systems. The same identity number should follow an object over its lifetime.

***Legal principle***

A statistical office should have access to administrative registers kept by public authorities. This right should be supported by law and the protection of privacy must also be protected by law.

### 1.4.1 Reliable administrative systems

Reliable administrative systems will generate data of good administrative quality. Good administrative quality is a necessary but not sufficient condition for good statistical quality. The systems for tax administration and welfare programmes will gradually develop and change, and these changes will determine what administrative data can be used for statistical purposes in the future. It is therefore important that national statistical offices maintain close and long-term relations with administrative authorities and politicians.

The long-term strategy requires high-level contacts to promote strategic changes that will improve statistics production. The statistical office must explain to the administrative authorities how their data are used for statistical purposes. The statistical office also needs detailed information on how the administrative systems are organised and what changes are planned. Close and long-term contacts at all levels are required for these purposes.

What aspects of national administrative systems are important for statistical offices? We note two such aspects here, coverage and identity codes.

#### *Coverage – the systems should cover all*

The Nordic systems for child benefits are good examples. All children in defined age groups are entitled to a sum of money. All parents want the entitlement – but to receive the money, the parents must be registered as parents to the child in question and national identity numbers are required for the parents and child. This system covers all children and all parents. As the information in the system's registers is maintained and updated, all persons in the country will gradually be covered and the register will contain administrative, but also statistically important, links between all parents and children.

It is important for good coverage that the administrative systems cover both urban and rural populations, rich and poor citizens, and small and big enterprises. The ideal is that there is no selectivity. If suitable methods are not developed, selectivity will result in biased statistical estimates. For instance, in the Nordic countries all seriously ill persons will see a doctor, and all doctors know that cancer patients should be reported to the National Cancer Register. In this way we can be almost absolutely sure that all patients with a cancer diagnosis are in the Cancer Register. If rural or poor persons are underrepresented, estimated cancer incidence and mortality figures would be of low quality.

#### *Unified systems of identity codes*

Identities are important in administrative systems. Legally important relations between persons, such as husband and wife, or parents and children, are registered with the identities of the persons in question. In many registers the legally important relations between owners and different kinds of property are recorded with both the identities of owners and identity of property. For taxpayers, it is important that the tax paid is recorded together with the identity of the taxpayer. It is therefore in the interest of each taxpayer to use a correct identity in each transaction. The legal importance of identities explains why identity data as a rule are of high quality in many administrative sources.

The best way to handle identities in administrative systems is to use *national identity numbers*. Persons, enterprises and property should be given unique identity numbers that are used in all administrative systems in the country, and the same number should follow each person, enterprise or property over its lifetime.

Not only will administration become efficient; the statistical production system will become efficient when administrative data are used for statistical purposes, as it will be possible to link records and create important statistical comparisons. With unique national identity numbers, record linkage will be easy and the risk of false matches and false non-matches will be low. The statistical possibilities that national identity numbers create will be explained in the following chapters.

It is advantageous if the identity numbers have no relation to any attributes of the objects that are to be identified. For example, identity numbers for persons should not depend on name, sex, or address of the persons, because such attributes can change over time. Throughout the book we will use the abbreviation PIN for national identity numbers for persons and BIN for national identity numbers for legal units representing enterprises.

### **1.4.2 Legal base and public approval**

There are preconditions concerning legal base and public approval that make possible the efficient use of administrative registers for statistics. These preconditions are discussed in UN/ECE (2007) and we build on that discussion here.

#### ***Legislation determines what data are generated***

The national administrative systems for taxation and welfare are based on legislation that determines the kind of administrative data that are generated within these systems. If, for example, citizens pay income tax to municipalities, then the authorities must know where each citizen lives. The municipal taxation and welfare systems are the legal base for the Nordic administrative population registers. They are used not only for taxation and municipal welfare, but also for elections where the population register defines where each voter votes. For statistical purposes, this creates very good links between persons and geography that facilitate regional statistics. The administrative registers are updated every day, which makes possible timely monthly demographic statistics.

#### ***Legislation to improve the national statistical system***

Politicians want to reduce the response burden of persons and enterprises as well as the direct costs for the production of community statistics.

- Legislation should provide the national statistical offices access to administrative microdata including identities, and the right to use the data for official statistics and research.
- Legislation should provide statistical offices the authority to match data from different sources and use data that were not originally generated for statistical purposes.

- Legislation could also instruct statistical offices to first use data from administrative registers and to conduct sample surveys or censuses only if available administrative data are insufficient.
- Some laws have the sole purpose of making register-based housing and population censuses possible. For example, the Nordic parliaments have decided that all employers must provide information on where all employees work – the local unit address for all. This information is given with income statements with data on employer identity, local unit identity, employee identity and wages and preliminary tax paid. These income statements play an important role in the Nordic statistical systems, as we obtain important links between three different object types. The parliaments have also decided that all persons should be registered at the dwelling where they live. It will then be possible to create statistics for households defined by the common dwelling in the register-based census.

### ***Legislation on data protection***

According to the second precondition in Chart 1.3, a national statistical office should have access to administrative registers kept by public authorities. This right should be supported by law and the protection of privacy must also be protected by law. Legislation that gives a statistical office access to administrative data is discussed above, and the protection of privacy and integrity are discussed below.

The principle of *one-way traffic* is important for data protection. Microdata can go from administrative authorities to the statistical office but never in the reverse direction.

The legislation on data protection should rest on a reasonable balance between protection of integrity on the one hand and increased costs and difficulties for statistics production on the other. An important task for top management at a national statistical office is to explain the consequences generated by proposed legislation to lawyers and politicians.

### ***Public approval***

The cooperation between register authorities and national statistical offices should be open and transparent. The fact that administrative data are used for statistical purposes should not be kept quiet; instead, the benefits and the efforts to protect integrity should be explained in open discussion and public debate.

It is important to explain that individual records regarding persons are anonymous in statistics production, in contrast to how administrative authorities handle the same data.

If the national statistical office has a good reputation as trustworthy, it will be easier to gain access to administrative data for statistics production. However, one mistake in the protection of integrity can immediately destroy this reputation.

Persons and enterprises do not want to be required to report to both an administrative authority and the national statistical office. Not having to do so will make public opinion more favourable to the use of administrative data for statistical purposes. It will become more difficult to motivate the double provision of data – why respond to a questionnaire on the enterprise's turnover when you also submit a value-added tax return to the Tax Agency which includes the same information?

Evidence that double provision of data to Statistics Sweden and to another authority is regarded as unreasonable can be seen in this newspaper clipping:

*Translated from a newspaper article:*

### **Refuse to send statistics to Statistics Sweden!**

Mr R from the B-farm thinks that the authorities should be able to find the information from their own registers. Mr R refuses to send in statistics to Statistics Sweden. Because he already sends in information every other week to the Swedish Board of Agriculture, he thinks that the authorities should cooperate with each other instead. ...

## **1.5 Basic concepts and terms**

Two principles form the basis of this book – the *survey approach* to administrative data and the *systems approach*. The survey approach means that we discuss estimates, estimators and quality as in a book on sample surveys. The systems approach builds on the *register system* concept that is introduced in Chapter 4 and is used throughout the book. We also discuss the *production system* at a national statistical office and the role of administrative registers in the design and development of that system.

We discuss three concepts in this section: what is a *statistical survey*, what is a *register* and what is a *register survey*? We also give examples of register surveys that illustrate some important principles discussed in later chapters: The Income and Taxation Register is a survey of persons and households and the Quarterly and Annual Pay Registers are business surveys.

### **1.5.1 What is a statistical survey?**

This term is a central term used by statisticians at all national statistical offices. For many statisticians, however, the term is synonymous with *sample survey*. This will cause confusion when we discuss statistics based on administrative registers.

To avoid this confusion, we follow the distinction between different kinds of surveys that Statistics Canada (2009) use in their Quality Guidelines. The guidelines are written with censuses and sample surveys as the main focus. In this book, we focus on register surveys (3 below), but also discuss and compare other survey methodologies.

Statistics Canada, Quality Guidelines:

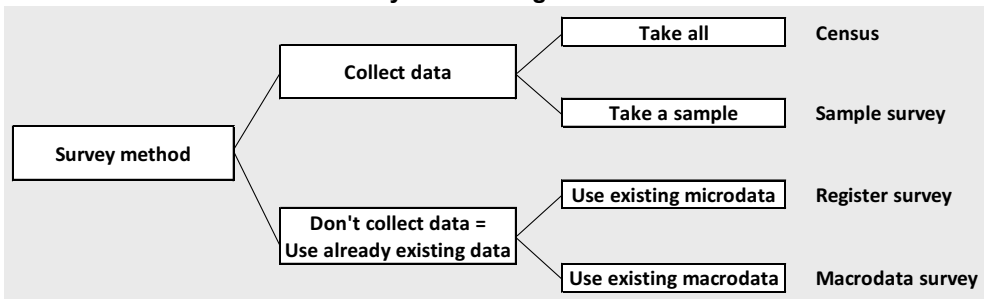
The term survey is used generically to cover any activity that collects or acquires statistical data. Included are:

1. a *census*, which attempts to collect data from all members of a population;
2. a *sample survey*, in which data are collected from a (usually random) sample of population members;
3. collection of data from *administrative records*, in which data are derived from records originally kept for non-statistical purposes;
4. a *derived statistical activity*, in which data are estimated, modelled, or otherwise derived from existing statistical data sources.

Estimates of, for example, number of employees by industry (as in Chart 1.2) can be based on a census, on a sample survey, or on a register survey. We can choose between these three different survey methodologies to estimate the same parameters. This is the reason why we have chosen to use the survey approach to administrative data – register surveys are only a new alternative to the two other well-established survey methods.

The fourth survey method above is the method that is used for the National Accounts. The National Accounts survey is based on a model-based compilation of macrodata (or estimates) from a system of economic surveys. Chart 1.4 compares the four kinds of surveys.

**Chart 1.4 The four different survey methodologies**



Sample surveys are based on a mathematical theory – probability and inference theory. Censuses and sample surveys are based on a non-mathematical survey methodology based on behavioural science – psychology and cognition are important aspects that are used to discuss errors that arise during the collection of statistical data through interviews and questionnaires.

Register surveys require a non-mathematical theory based on a systems approach. Macrodata surveys should also be based on a theory of systems of surveys. We discuss these issues later in this book when we introduce the concept of *survey system design*.

### 1.5.2 What is a register?

An administrative register is maintained to store records on all objects to be administered, and the administrative process requires that all objects can be identified. The following definition is valid for administrative and statistical registers:

A *register* aims to be a complete list of the objects in a specific group of objects or population. However, data on some objects can be missing due to quality deficiencies.

Data on an object's identity should be available so that the register can be updated and expanded with new variable values for each object.

– *Complete listing* and

– *known identities* are thus the characteristics of a register.

*Catalogue, directory, list, register, registry* are different terms for the same concept. We will only use the term *register*.

The following are examples of registers:

- Civic, civil or national registration of the population in a country results in registers of citizens, births and deaths.
- Income self-assessments from persons give registers of all taxpayers for a given year.
- In Sweden, enterprises with a turnover of SEK 40 million or more should report monthly. This gives monthly registers of all enterprises that have reported. For smaller enterprises, we obtain quarterly or yearly registers.
- All export and import transactions are registered by Customs. Monthly registers are created with all transactions for a specific month.
- A census file with data from a housing and population census is a register if there are identities of the persons in the file.

The identities used in register processing can either be identity numbers that are unique within a national administrative system or an identity number in a subsystem with keys to the identities in other systems. It is also possible to use identities defined by, for instance name, address, date of birth and place of birth.

These identities will be used in *deterministic matching* of the objects in different registers, where the aim is to find identical or related objects in two registers. In deterministic matching, two records are linked if the identifiers agree exactly. This is the most efficient method when the identifying variables are of good quality.

**Chart 1.5 Deterministic matching with Personal Identity Numbers, PIN**

| Population register |     |     | Administrative income register |        | Combined register after exact matching |     |     |        | Statistical income register after imputations |     |     |        |
|---------------------|-----|-----|--------------------------------|--------|----------------------------------------|-----|-----|--------|-----------------------------------------------|-----|-----|--------|
| Person              | Sex | Age | Person                         | Income | Person                                 | Sex | Age | Income | Person                                        | Sex | Age | Income |
| PIN1                | F   | 87  | PIN1                           | 167    | PIN1                                   | F   | 87  | 167    | PIN1                                          | F   | 87  | 167    |
| PIN2                | M   | 74  | PIN2                           | 215    | PIN2                                   | M   | 74  | 215    | PIN2                                          | M   | 74  | 215    |
| PIN4                | M   | 62  | PIN3                           | 94     | PIN3                                   | *   | *   | 94     | PIN3                                          | *   | *   | 94     |
| PIN5                | F   | 49  | PIN4                           | 341    | PIN4                                   | M   | 62  | 341    | PIN4                                          | M   | 62  | 341    |
| PIN6                | F   | 35  | PIN5                           | 298    | PIN5                                   | F   | 49  | 298    | PIN5                                          | F   | 49  | 298    |
| PIN8                | M   | 14  | PIN6                           | 277    | PIN6                                   | F   | 35  | 277    | PIN6                                          | F   | 35  | 277    |
|                     |     |     |                                |        | PIN8                                   | M   | 14  | *      | PIN8                                          | M   | 14  | 0      |

Because person PIN3 is not in the population register and person PIN8 is not in the administrative income register, the combined register after deterministic matching will have two records with missing values due to this non-match.

Many administrative registers consist only of persons or enterprises of a defined category. Only persons with income are in the administrative income register in the example in Chart 1.5. When such registers are combined with the population register, the non-match will generate missing values. Zero income must be imputed for persons not in the administrative income register, such as person PIN8. Person PIN3 is not in the population register and if that person is not found in any other register the non-match will result in missing values (\*) for sex and age.

### 1.5.3 What is a register survey?

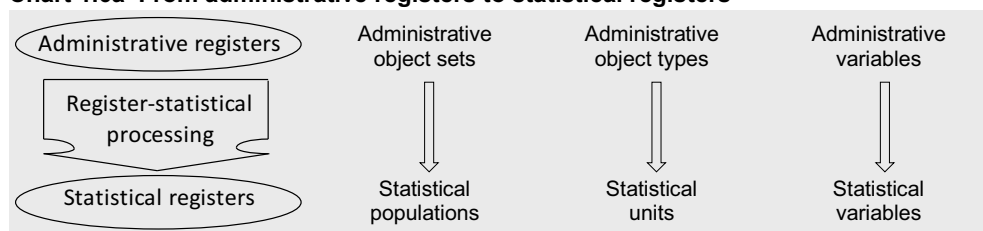
The original data are generated in public administrative systems. Definitions of object sets, objects and variables are adapted to administrative purposes. Every authority carries out controls, corrections and other processing suited to their administrative aims.

When an authority delivers data to a national statistical office, further selections and processing may be carried out to meet the needs of the statistical office. The authorities also have metadata as definitions, administrative rules and quality aspects, based on the administrative authority's experiences and investigations. This information is important for those receiving the data at the statistical office.

It is generally not a good idea to produce statistics directly from the received administrative registers because these are not adapted to statistical requirements. The object sets, object definitions and variables need to be edited, and as a rule it will be necessary to carry out some processing so that the register fulfils the statistical requirements for population, objects and variables. The register-statistical processing, which aims to transform one or several administrative registers into one statistical register, should be based on generally accepted statistical methods.

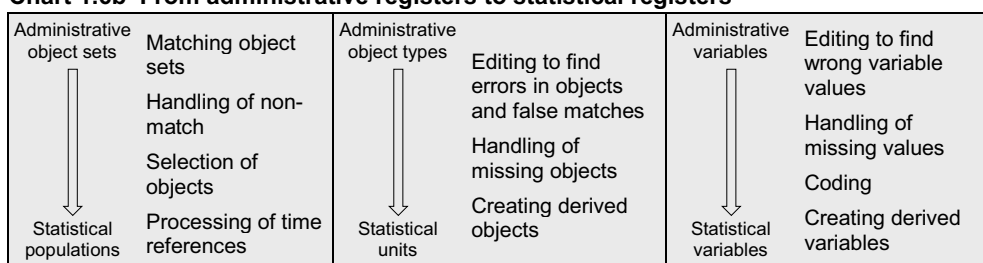
Chart 1.6a shows three important components of this work. We have found that people have a tendency to use administrative concepts as they are, and in some cases this can be acceptable – but in other cases it can be unacceptable. The three issues of how to define population, units and variables of a statistical register are important for the quality of the statistics to be produced with the newly created statistical register.

**Chart 1.6a From administrative registers to statistical registers**



A statistical population or administrative object set consists of  $N$  *objects* or *units* or *elements*. Of these three synonyms, we will as a rule use the term *object* for the units in an administrative object set and the term *statistical unit* for the units in a statistical population. The register-statistical processing is described in Chart 1.6b.

**Chart 1.6b From administrative registers to statistical registers**



### 1.5.4 The Income and Taxation Register

The Income and Taxation Register (I&T) is an important part of Statistics Sweden's register system. It is used to describe income distribution and for regional income statistics, and it is the basis for longitudinal income registers used by university researchers.

This register utilises many administrative sources, and many administrative variables are used to create important statistical variables. Besides these administrative sources, it is necessary to use the register system at Statistics Sweden: the Population Register is used to define the population of the Income and Taxation Register, and important classification variables are imported from other registers in the system to the Income and Taxation Register.

#### 1. *Data generation at the National Tax Agency*

The annual income self-assessment is based on tax returns from income earners and the taxation decisions of the local tax authority. Both the income earner and the tax authority use statements of earnings for salary, sickness benefits and interest payments that are the responsibility of employers, social insurance office and finance companies. The National Tax Agency ultimately compiles this information. Tax returns, statements of earnings and taxation decisions can be changed and supplemented. Data for one person can thus be very complex.

#### 2. *Microdata deliveries to the Income and Taxation Register*

The Swedish National Tax Agency annually creates databases that contain information on Sweden's population. The data files for one year – containing around 9 million records, each with around 300 variables – are delivered to the Income and Taxation Register at Statistics Sweden.

#### 3. *Metadata to the Income and Taxation Register*

Record descriptions with names and definitions of variables accompany the deliveries from the National Tax Agency. Tax return forms, statement of earnings forms, taxation decisions and tax return instructions are also necessary for the correct interpretation of the data.

#### 4. *Editing of data*

The I&T Register receives data from many different suppliers outside and inside Statistics Sweden. External data are edited. Data from other Statistics Sweden registers have already been edited. Contacts with suppliers are important to obtain knowledge of changes in the administrative system, which in turn is important to ensure the quality of the register statistics – administrative changes should not be interpreted as actual income changes.

#### 5. *Matching and selections*

There is a large number of registers that should be processed to create the different sub-registers that are included in the Income and Taxation Register. Records from different sources are matched using Personal Identification Numbers (PIN), and aggregation is carried out at the same time, i.e. all the statements of earnings data for a specific person are aggregated so that the person's income from work can be put together. One type of processing is to select persons aged 16 and older who were also parts of the population on 31 December.

## 6. Derived objects are created

More information on certain relations helps to form household units. Between adults, the relations *married* or *cohabiting adults with children in common* result in their placement in the same family unit. These relations are derived from the family members' personal identification numbers; these reference variables are found in the taxation data and in Statistics Sweden's Population Register.

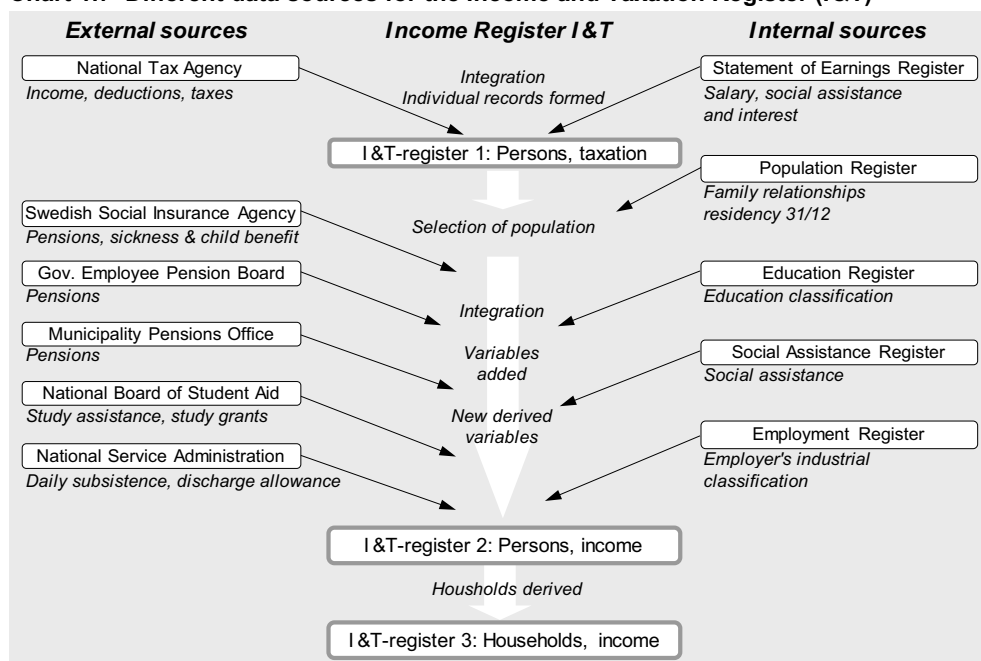
## 7. Derived variables are created

A large number of derived income variables are formed. For instance, the wage or salary amounts are aggregated from the different earnings data to become an individual's *income from work*. Every person's total income from work and capital plus transfer payments minus tax becomes the person's *disposable income*. For households, variables such as *household type*, *number of consumption units* and *disposable income* are formed.

Chart 1.7 shows how the Income and Taxation Register receives administrative data from a variety of different external sources and some Statistics Sweden registers. The middle column shows the different phases when the sources are used during the process to create the new statistical register.

This example shows the importance of the principles in Chart 1.1. Statistics Sweden has access to many administrative registers with variables describing different kinds of income. The object set and the administrative variables have been processed to meet statistical needs. Many sources have been used to produce a statistical income register with rich content. The population in the income register is consistent with other statistical registers within the register system.

**Chart 1.7 Different data sources for the Income and Taxation Register (I&T)**



### 1.5.5 The Quarterly and Annual Pay Registers

Aggregate wages and salaries by economic activity and institutional sector are important inputs for yearly and quarterly National Accounts. One quarterly and one annual register survey produce these estimates and we will use these surveys to illustrate the different ways a source can be used as noted in Section 1.3.

#### *A simple and straightforward approach*

An administrative register of good quality can be used almost as if it were for statistical purposes. If content and coverage are sufficient, it will be relatively easy to use the administrative register and produce statistics of good quality. The different steps of the yearly survey are illustrated in Charts 1.8a–1.8c.

The administrative Annual Pay Register is first *edited* to find incorrect or unreasonable values of aggregated wages and salaries (WagesYear). The identity number (BIN) is first checked to ensure that all values have the right format and acceptable values. Preliminary tax (Prel-tax) should be between 30% and 35% of WagesYear and this relation is used to edit WagesYear. For the enterprise with identity number BIN05, we find a 1000-factor error; 2 is therefore replaced with 2 000 and the variable W-imp is created to show imputed values of WagesYear (Chart 1.8b).

The registers in Chart 1.8a are then *matched* with the identity numbers BIN. Records with the same value of BIN are combined in a new register that is shown in Chart 1.8b.

**Chart 1.8a The sources of the Annual Pay Register**

| 1. Business Register |         |      | 2. Administrative Annual Pay Register |           |        |          | Sector:                     |
|----------------------|---------|------|---------------------------------------|-----------|--------|----------|-----------------------------|
| BIN                  | Sector  | ISIC | BIN                                   | WagesYear | Source | Prel-tax |                             |
| BIN02                | 6       | 52   | BIN01                                 | 25        | I      | 8        | 1 Non-financial enterprises |
| BIN03                | 1       | 51   | BIN03                                 | 1 667     | I      | 544      | 2 Financial enterprises     |
| BIN04                | 7       | 91   | BIN04                                 | 796       | I      | 252      | 3 Government                |
| BIN05                | 1       | 70   | BIN05                                 | 2         | P      | 689      | 4 Municipalities            |
| BIN06                | 1       | 45   | BIN06                                 | 92        | I      | 29       | 6 Self-employed             |
| BIN07                | 1       | 51   | BIN07                                 | 4 758     | I      | 1 565    | 7 Non-profit organisations  |
| BIN08                | 1       | 60   | BIN08                                 | 39        | P      | 12       |                             |
| BIN09                | 1       | 28   | BIN09                                 | 452       | I      | 142      |                             |
| BIN10                | 1       | 74   | BIN11                                 | 289       | P      | 95       |                             |
| BIN11                | 1       | 27   | ...                                   |           |        |          |                             |
| ...                  |         |      | Count                                 | 305 411   |        |          |                             |
| Count                | 331 518 |      |                                       |           |        |          |                             |

When the 331 518 records with active employers in the Business Register are matched with the 305 411 records in the administrative Annual Pay Register we find that there is a lot of *non-match*. There are 33 543 records in the Administrative Register missing in the Business Register; this is an indication of *undercoverage*. And 59 650 records are missing in the administrative register; this is an indicator of *overcoverage* in the Business Register. The non-match is shown to the right of Chart 1.8b below.

This non-match above gives rise to *missing values* (\*) in the variables Sector and Economic Activity, ISIC. After imputations the final statistical Annual Pay Register in Chart 1.8c can be created.

**Chart 1.8b After matching****3. After matching of register 1 and register 2**

| C1    | C2     | C3   | C4    | C5        | C6    |
|-------|--------|------|-------|-----------|-------|
| BIN   | Sector | ISIC | BIN   | WagesYear | W-imp |
| *     | *      | *    | BIN01 | 25        | 0     |
| BIN02 | 6      | 52   | *     | *         | *     |
| BIN03 | 1      | 51   | BIN03 | 1 667     | 0     |
| BIN04 | 7      | 91   | BIN04 | 796       | 0     |
| BIN05 | 1      | 28   | BIN05 | 2 000     | 1     |
| BIN06 | 1      | 45   | BIN06 | 92        | 0     |
| BIN07 | 1      | 51   | BIN07 | 4 758     | 0     |
| BIN08 | 1      | 60   | BIN08 | 39        | 0     |
| BIN09 | 1      | 28   | BIN09 | 452       | 0     |
| BIN10 | 1      | 74   | *     | *         | *     |
| BIN11 | 1      | 27   | BIN11 | 289       | 0     |
| ...   |        |      |       |           |       |

|    | Name      | Count                 | Missing |
|----|-----------|-----------------------|---------|
| C1 | BIN       | 365 061               | 33 543  |
| C2 | Sector    | 365 061               | 33 543  |
| C3 | ISIC      | 365 061               | 33 543  |
| C4 | BIN       | 365 061               | 59 650  |
| C5 | AggrWages | 365 061               | 59 650  |
| C6 | W-imp     | Imputed values, wages |         |

The non-match in Chart 1.8b arises because the Business Register and the administrative Annual Pay Register give conflicting pictures of the population of active employers. Such conflicts always arise when different sources are compared – and a decision must be made regarding which source should be trusted.

If the Business Register in C1–C3 is trusted, then enterprise BIN01 should be excluded and missing values should be imputed for enterprises BIN02 and BIN10.

If instead the Pay Register in C4–C6 is trusted, then enterprises BIN02 and BIN10 should be excluded and missing values for enterprise BIN01 should be imputed as in Chart 1.8c.

**Chart 1.8c The statistical Annual Pay Register****4. Statistical Annual Pay Register – Overcoverage has been taken away**

| BIN   | Sector | S-imp | ISIC | I-imp | WagesYear | W-imp |
|-------|--------|-------|------|-------|-----------|-------|
| BIN01 | 6      | 1     | 01   | 1     | 25        | 0     |
| BIN03 | 1      | 0     | 51   | 0     | 1 667     | 0     |
| BIN04 | 7      | 0     | 91   | 0     | 796       | 0     |
| BIN05 | 1      | 0     | 28   | 0     | 2 000     | 1     |
| BIN06 | 1      | 0     | 45   | 0     | 92        | 0     |
| BIN07 | 1      | 0     | 51   | 0     | 4 758     | 0     |
| BIN08 | 1      | 0     | 60   | 0     | 39        | 0     |
| BIN09 | 1      | 0     | 28   | 0     | 452       | 0     |
| BIN11 | 1      | 0     | 27   | 0     | 289       | 0     |
| ...   |        |       |      |       |           |       |

Count 305 411

The choice between these alternatives should be based on a clear understanding of *the administrative system*. In this case, it does not cost anything to remain as ‘active employer’ in the system and report and pay 0 every month. Enterprises that have been active employers in the past may thus choose to remain in the Business Register as ‘active employers’ even if they have no employees this year. For this reason we have chosen to trust the Annual Pay Register.



When we compare C4 and C10 in Chart 1.8d, we find that aggregate wages are almost the same with the exception of enterprise BIN07. We have added three records that show large differences. For enterprise BIN20, it was found that insurance companies (ISIC=65) report insurance benefits in the same way as wages in the monthly tax reports, as tax is paid in the same way as for wages.

In addition, the differences are large for enterprises BIN30 and BIN40 – this is an example of a takeover at the middle of the year – after the takeover enterprise BIN30 sends in tax reports for all employees at both units. As a result, the Annual and Quarterly surveys will give inconsistent estimates of aggregate wages by industry.

Hence, the answer to the first question is: yes, the Annual survey can be improved if we also use the Quarterly survey. The imputed value for enterprise BIN05 can be replaced with the value from the Quarterly survey; the values from the Quarterly survey give better estimates by industry for the last two enterprises in Chart 1.8e; and coverage has been improved with 2 812 more enterprises.

**Chart 1.8e Final version of the statistical Annual Pay Register**

Aggregate wages are based on the best of Annual and Quarterly sources

| BIN   | Sector | S-imp | ISIC | I-imp | AggrWages | W-imp | W-source |
|-------|--------|-------|------|-------|-----------|-------|----------|
| BIN01 | 6      | 1     | 01   | 1     | 25        | 0     | Y        |
| BIN03 | 1      | 0     | 51   | 0     | 1 667     | 0     | Y        |
| BIN04 | 7      | 0     | 91   | 0     | 796       | 0     | Y        |
| BIN05 | 1      | 0     | 28   | 0     | 2 074     | 0     | Q        |
| BIN06 | 1      | 0     | 45   | 0     | 92        | 0     | Y        |
| BIN07 | 1      | 0     | 51   | 0     | 4 758     | 0     | Y        |
| BIN08 | 1      | 0     | 60   | 0     | 39        | 0     | Y        |
| BIN09 | 1      | 0     | 28   | 0     | 452       | 0     | Y        |
| BIN11 | 1      | 0     | 27   | 0     | 289       | 0     | Y        |
| ...   |        |       |      |       |           |       |          |
| BIN20 | 2      | 0     | 65   | 0     | 627       | 0     | Y        |
| ...   |        |       |      |       |           |       |          |
| BIN30 | 1      | 0     | 74   | 0     | 2 067     | 0     | Q        |
| BIN40 | 1      | 0     | 73   | 0     | 605       | 0     | Q        |
| ...   |        |       |      |       |           |       |          |

|          |                      |         |                              |
|----------|----------------------|---------|------------------------------|
| S-imp    | Sector imputed       | 6 795   | only in Annual               |
| I-imp    | ISIC imputed         | 299 199 | both in Quarterly and Annual |
| W-imp    | AggrWages imputed    | 2 812   | only in Quarterly            |
| W-source | Source for AggrWages | 308 806 | total number of records      |

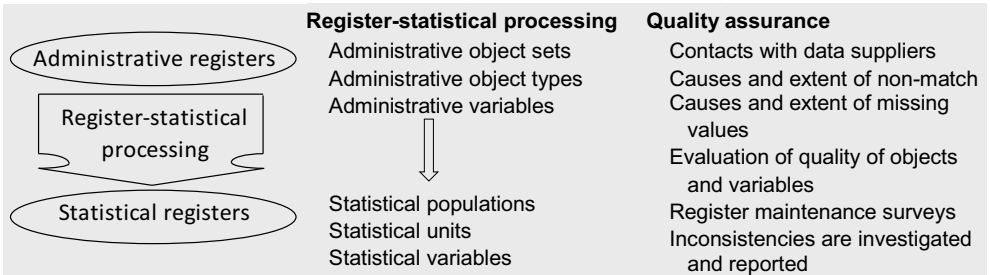
Let us turn to the second question: can the Annual survey be used to improve some parts of the production system at our national statistical office?

During work with the Annual Register, we found coverage errors in the Business Register. As the Business Register is used as a basis for all business statistics, the accuracy of many economic surveys will be improved if data from the Annual and Quarterly surveys can reduce these coverage errors. When data from the Annual

and Quarterly registers were compared, we found a serious misunderstanding regarding insurance companies. Both wages and insurance benefits were reported as wages in the monthly tax reports on aggregate wages. With more variables for the same enterprise, we have greater possibilities of finding and correcting errors in variables, but we can also obtain information on units in the population that have changed in some important way, e.g. that one enterprise has taken over another.

During the work with register-statistical processing, we have had the opportunity to observe errors and anomalies in the data we use. These observations should be developed into systematic work with quality assurance, as indicated in Chart 1.6c.

**From Chart 1.6c From administrative registers to statistical registers**



**1.6 Comparing sample surveys and register surveys**

Planning a register survey is completely different from planning a sample survey. The first step with a *sample survey* is to decide on the population and which parameters should be estimated for which domains of interest. This determines the character of the survey with regard to sample design and estimation. The population and parameters are defined first, followed by the collection of data. In general, *one* survey is considered at a time with a limited number of parameters and domains of interest. When working with sample surveys, the concept pairs *population – sample* and *parameter – estimator* are often used.

With a *register survey*, the starting point is very different, as data have already been collected and are available in different registers where they are not adapted for any specific survey. Using the available registers, the objects and variables that are relevant for the survey are selected. Sometimes it is necessary to form new variables from the existing registers – and sometimes even new objects (statistical units). But, in the beginning, the data already exist and then the population is defined. The plan of tables is then decided without thinking in terms such as parameter estimation and domains of interest. Sampling error does not limit the possibilities for choosing domains for the analysis and reporting of results. When working with register-based statistics, the distinction between *population* and *sample* has no relevance and it is not usual to use the terms *parameter* and *estimator*.

When designing a statistical register and a register system, it is desirable to make them flexible so that they are as widely applicable as possible. Therefore, an important part of register-statistical methodology work is to structure and improve the

whole, i.e. find the best possible design for the entire register system. This includes long-term work to monitor and influence the access to administrative data for statistical purposes.

Although much statistical methodology can be considered the same for sample surveys and register surveys, such as problems with errors (excluding sampling error) and work with analysis and presentation, the way of thinking is different, as sampling error and design problems are so crucial for sample surveys. *System-oriented thinking* is fundamental for register-based statistics. To improve quality, it is insufficient to look at one register at a time; instead the system should be seen in its entirety. Special attention should be paid to the quality of the base registers and the identifying variables that act as links between the different registers.

There are different preconditions for *the editing process* for sample surveys and for register surveys. If unreasonable values are detected when checking a sample survey, it is possible to re-contact the relevant respondent. Alternatively, it is possible to repeat a question during an interview if a response seems anomalous. These solutions are not possible with register surveys. However, questions can be put to the register-providing authority on what is considered a reasonable value and how errors can occur. It is important that staff at the statistical office know which variables have been checked and corrected by the authority.

Register surveys have editing problems in other situations as well. When receiving large registers, it may be necessary to check and correct technical errors. For example, data that have been scanned can still have values in the wrong positions. In certain cases, data are not collected centrally but by different regional authorities. Then, the statistical office can investigate whether there are structural differences between the different authorities regarding time delays, response patterns, etc.

The problem of *nonresponse* also has a different character for sample surveys and register surveys. Missing values also exist in register surveys, but persons who could not be contacted or refused to respond do not cause the same problem as in sample surveys or censuses, and it is not necessary to send out reminders or decide when to bring an end to the data collection. In sample surveys and censuses, estimates are revised to counteract the effects of nonresponse. This does not always occur with register-based statistics.

*Measurement errors* exist both in sample surveys and register surveys. For sample surveys, we focus on minimising measurement error by testing and improving the questionnaires. We cannot, however, work in the same way with register-based statistics, because the administrative authorities have designed the questionnaires. However, a national statistical office should be able to influence different authorities so that the questionnaires and instructions can be improved.

The *presentation* of survey results has somewhat different prerequisites. We cannot present results from sample surveys for groups that are too small, as the limits are set by the sampling error. Therefore, the tables presented cannot be broken down by many variables at the same time nor have too many cells. However, with register-based statistics, it is possible to produce large detailed tables that are broken down in many ways. Such tables are often difficult to interpret, and place

great demands on the method of presentation. There are also limits for how detailed the tables can be due to *risks of disclosure* for the separate individuals and enterprises.

These examples illustrate how sample surveys and register surveys often have different types of methodology problems. Methodology development in the areas of editing, nonresponse and measurement error has, until today, principally concerned sample surveys.

To avoid mixing together the different concepts, it is important to clearly distinguish between the three types of surveys that exist at a statistical office. Here we compare the similarities and differences in the conditions for sample surveys, censuses and register surveys. The latter, which will be discussed in the following chapters, has been shaded in Chart 1.9.

Data collected for censuses and registers based on administrative data can be included in the register system, if the data contain identifying variables so that the data can be linked with other registers. However, sample surveys are not part of the register system. Sample surveys can use registers in the system, but as sample survey data cannot be used by registers in the system, they are not included in the register system.

**Chart 1.9 Similarities and differences between the different types of surveys**

| Chart 10: Similarities and differences between the different types of surveys |                                                                                     |                                                         |
|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|---------------------------------------------------------|
| Sample survey                                                                 | Census                                                                              | Register survey                                         |
| Not included in register system                                               | Included in register system – can be used for other register surveys                |                                                         |
| Uses the register system to define populations and as a source of variables   |                                                                                     |                                                         |
| Sample design, estimation, measures of uncertainty                            | System-oriented thinking and coordination with other register surveys are important |                                                         |
| Own data collection – produce own questionnaires                              | Uses other administrative registers                                                 |                                                         |
| Editing – can contact respondents                                             | Editing – can contact register-providing authority                                  |                                                         |
| Nonresponse – reminders, when to stop data collection?                        | Non-match related to missing values or under-coverage                               |                                                         |
| Quality flaws – sampling errors, measurement errors                           | Quality flaws – measurement errors                                                  | Quality flaws – relevance errors, lack of comparability |
| Small tables – cannot give estimates for small groups                         | Presentation – large tables with many cells, disclosure risks                       |                                                         |

## 1.7 Conclusions

A theory can be defined as a system of principles, and we discuss the principles that should be used for register-based statistics.

- The register-based statistics in the Nordic countries were developed by subject-matter specialists from the 1960s and onwards. A number of very creative entrepreneurs have contributed to this development.
- Many statistical registers were developed, ad hoc methods were used – in most cases, good ad hoc methods were developed.
- Our work started with the book by Statistics Denmark (1995). We have been searching for the principles underlying the Nordic register-based statistics – both the principles actually used, but also good statistical principles that should generally be used. We wanted to transform good ad hoc methods to general statistical principles and find alternatives to the bad ad hoc methods we found.
- Our main conclusion is that the success of register-based statistics in the Nordic countries is explained by the efficient register systems that were developed. A system of statistical registers is a new concept within statistical survey theory. The register system ensures that microdata can be integrated and used effectively and opens new possibilities for quality assurance.

### *New competence and a new paradigm*

Countries where the national statistical office wants to develop register-based statistics should have a long-term policy and a short-term policy. The long-term policy includes strategic planning and building close contacts with politicians and administrative authorities to promote good preconditions in the future.

The short-term policy should be to gain access to administrative registers as soon as possible and start the work of building the new competence that will be needed. The new competence should be based on a new paradigm – instead of the traditional *one survey at a time* paradigm, a new *system-oriented* paradigm is necessary. When administrative registers are used, it will be possible to combine sources and then consistency and coherence will be central issues.

### *The importance of the register system*

After more than 20 years at Statistics Sweden, we have learnt two facts:

- The register system is an extremely valuable asset that today is widely used – and can be used even more – to produce statistics and improve quality.
- It is costly and difficult to change the production system when causes of errors are detected. When we find inconsistencies between two surveys, there are two managers, both with their own budget, who should act together. And there are often expensive IT systems that must be changed.

The conclusion is that it is important to do the right thing from the beginning. New register countries that are unaware that a system of registers has started to develop and grow at their national statistical office can save large amounts of money and time if they have a plan for their future system right from the beginning of their work with administrative sources.

### ***Register surveys – basic principles***

In Chart 1.3 we mention two basic preconditions, the identity number principle and the legal principle. Administrative data with good identifiers and an appropriate legal framework are necessary to develop an efficient national statistical system.

In Chart 1.1 we note four basic methodological principles: that administrative data should be transformed into statistical registers, the importance of a coordinated register system, and the consistency and quality principles.

These six principles are used in the rest of the book and we follow the structure illustrated in Chart 1.6, where the work with creating statistical registers is divided into three parts: creating population, units and variables. Books on sampling theory start after the creation of population, units and variables, but in register surveys the main part of the work consists of creating populations and variables, which is described in Chapters 7 and 8. When we discuss the quality of register-based statistics, we also use this structure and analyse the quality of register populations and the administrative units and variables.

It is important to remember that an administrative register or source can be used in four ways:

1. *Completely alone, if there are no identifiers.*
2. *Alone, but in combination with a base register.*
3. *In combination with a base register and other administrative sources.*
4. *To improve other surveys, i.e. to improve the production system.*

The examples in Section 1.5.5 regarding the Quarterly and Annual Pay Registers illustrate that many ways of using an administrative source are possible. We should not only consider one source and one register survey at a time, but also think of other sources that could be combined together and whether a source can be used for other purposes.

Administrative source 2 in Chart 1.10 is the main source for the Quarterly Pay Register, but it can also be used for the Annual Pay Register. When we use the Quarterly register for quality assessment, we find coverage problems with the Business Register. When we use the Annual Pay register for quality assessment, we find errors regarding data from insurance companies in source 2. If sources 1, 2 and 3 are combined and used in new ways, errors can be reduced. The example illustrates that a system-oriented approach is essential when we design register surveys.

**Chart 1.10 Consider many sources and registers simultaneously**

