Chapter 4. GDP by the income approach

4.0 Introduction

For 1995, the calculation of income-based GDP can be summarised as in the table below:

Table 100 GDP, income approach, 1995

	Value, DKK million	% of GDP
Compensation of employees	534 094	53
+ Gross operating surplus and mixed income	338 010	33
+ Taxes on production and imports	173 270	17
- Subsidies	35 618	4
GDP	1 009 756	100

With a highly developed and integrated system of national accounts such as the Danish, where the gross operating surplus component of income-based GDP comes from the same accounts at firm level as are used for the output-based estimate, the description of sources and methods will obviously overlap to a large extent. The description of the income-based estimate in Chapter 4 is much shorter than the output-based description, but this should not be taken as an indication of the relative importance of the two estimates in the Danish national accounts. Much of what has been said in Chapter 3 about business accounts is equally relevant to the income-based estimate and could in theory simply be copied into Chapter 4. Such repetition is not, however, desirable. Instead, it was decided that Chapter 4 should refer to the relevant sections of Chapter 3 where business accounts are concerned, and concentrate on the other parts of the estimate, i.e. first and foremost the compensation of employees.

4.1 Reference framework

4.1.1 Reference framework – the business register, the central population register and the salary information register (COR)

Since the compensation of employees as a component of income-based GDP equals total wages and salaries in domestic (resident) producer units, it is in principle the population of domestic producer units as registered in the business register which is the directly relevant reference framework for both the compensation of employees and gross operating surplus and mixed income when GDP is calculated using the income approach. Danmarks Statistik's business register was discussed in Sections 1.3.1. and 3.1.1.

The other two components, taxes and subsidies on production and imports, have an additional reference framework, namely general government and the EU institutions as recipients/payers. Central government authorities and the EU are, of course, covered by the statistics, and this section will therefore not dwell on the statistical units for the administration of taxes and subsidies on production and imports.

In theory, individuals need not come into the picture when GDP is calculated using the income approach, but in practice the main advantage of having an estimate based on the income which individuals have received alongside the output- and expenditure-based estimates lies in the very fact that it is based on individuals rather than enterprises. It is thus a robust and independent check on the estimates which use the other two approaches. The major difference between income- and output-based estimates is that for the latter the compensation of employees is worked out as expenditure for businesses and is based on business accounts whereas using the income approach it is worked out from the point of view of the income recipients, namely as the wages or salaries which employees have received according to tax statistics and other sources. Since personal information has been used as a starting point, it is appropriate to look at the reference framework constituted by those persons who, during a given period, have earned the income created during the same period in the resident producer units, i.e. on the economic territory.

Individuals may be either resident or non-resident. The problems arising with non-residents are dealt with in the discussion on the rest-of-the-world account. Here, we simply need to note that non-residents employed in Danish producer units and receiving wages or salaries from their employment are allocated for the purposes of the information forms an identification number which is the counterpart of the "person number" for Danish residents. For residents in Denmark, the reference framework is *Det Centrale Personregister* [the Central Population Register], compiled under legislation dating from 1967. All persons with a residence permit in Denmark are entered in this register and have a personal identification number ("person number") used by the public authorities for all administrative purposes. Asylum-seekers who have not received a residence permit are not allowed to be economically active and are not relevant in this connection. It is theoretically possible for persons living illegally in the country without a personal identification number to be earning substantial incomes from trade or industry, but in practice the widespread use of the identification number system means that we can rule out the possibility that anyone could earn income for any length of time from anything other than hidden economy or criminal activities without being registered in the CPR. If it does happen, the incomes thus acquired are considered to have been covered by the allowances for the hidden economy – cf. Section 1.7 – or to be criminal activity which for the time being is not supposed to be included in the estimates of GNI.

The wages and salaries etc. which employers in Denmark pay to their employees must, according to tax legislation, be reported annually on an “information form”. There are no exceptions to this rule, and there is no lower limit for the amounts involved. Income below the threshold for general income tax has to be reported on the information forms and a special, proportional “state participation tax” is collected on all earned income as from the first krone of earnings. In 1995, it was 6%. *Told&Skat* [the Central Customs and Tax Administration] records the information forms centrally in the salary information register (COR) on the basis of personal identification numbers. Along with the CPR, this register is the reference framework for the estimate of the compensation of employees in Denmark apart, of course, for the allowances for the hidden economy.

The subsections of Section 4.1 below outline the main features of the Danish system of statistics on persons which are relevant to the national accounts estimate of the GDP component "compensation of employees". These sections - with a few, mainly stylistic changes - have been taken from the book "Statistics on Persons in Denmark: A register-based statistical system", published by Danmarks Statistik in 1994, available in Danish and English. Below are extracts from those sections which are most relevant to national accounts.

Of particular importance for the estimate of this component divided by industry is the Workplace Module, which is one of the six classification modules in the system. The description is also relevant

to the estimate of the population of households for use in the Household Budget Survey (FU) which is an important source for the estimate of GDP from the expenditure side, as described in Chapter 5 below.

Sections 4.1.2 – 4.1.16 discuss the evolution of the combined statistical system, from the raw data via linkages to the end product in the form of a system of statistical registers. They discuss factors which have made it possible to use administrative registers for statistical purposes, the Act on Danmarks Statistik and the Central Population Register. The combined statistical system, consisting of registers of persons, property and businesses, is described, along with classification modules.

Below we describe the processes - linkages - by which administrative registers are converted into statistical registers. This section is largely a technical description which ties in with the concepts and methods of analysis which are established principles in the work of Danmarks Statistik.

4.1.2 Structure of statistics on persons

4.1.2.1 Basic information, statistical registers and classification modules

The basic information in the form of extracts from administrative source registers or other material collected is processed and organised into a number of statistical registers and classification modules.

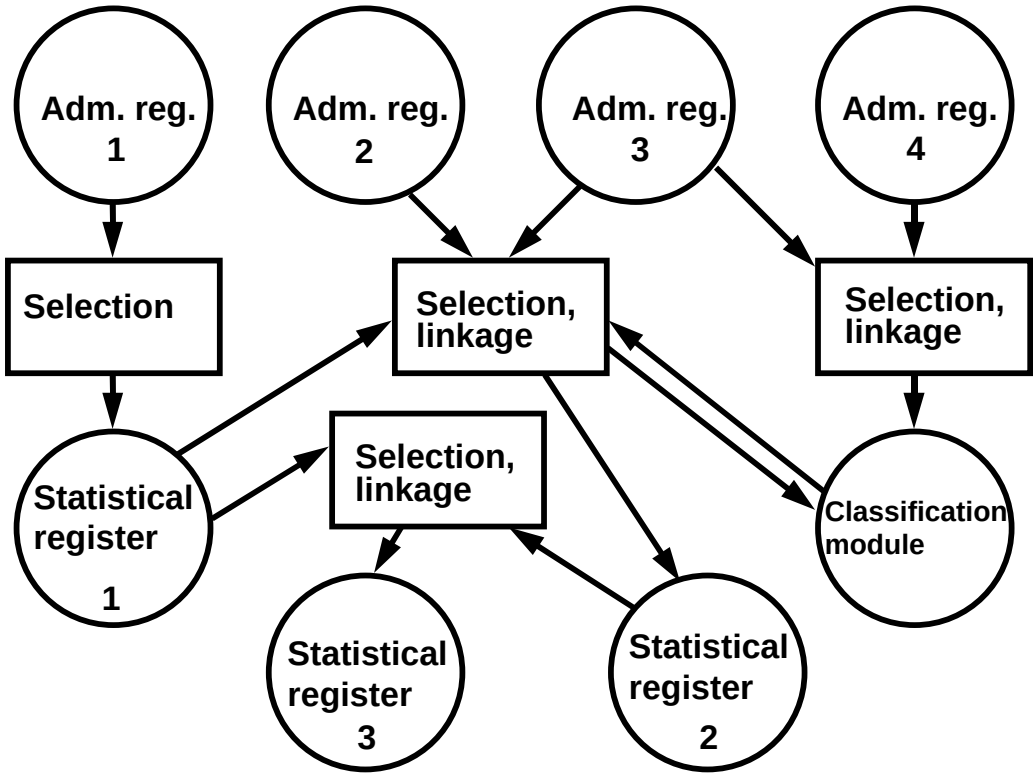
The individual statistical register is subject-related to the extent that it forms the basis for statistics within a certain field. This means that we work on the basis of registers which are clearly defined in terms of content and in which there is a clear connection between the source of the key information and the statistical register in question. The usual method is to collect subject-specific statistical data from a single source and to add to them a range of background information, often from other statistical registers. This combination of information then constitutes the register for the statistical area in question.

This clear connection between the basic register and the statistical register is still the dominant principle, but the need to view things in a broader context (horizontal integration) has meant that more recent statistical registers are increasingly derived from linked statistical registers. The registers may still be said to be subject-related, but the subject fields have taken on a somewhat different character. Similarly, for the purposes of longitudinal studies, versions of the same register for different years form the basis for the creation of a new register. This is called vertical integration.

In certain cases the basic information is processed centrally and stored in the *classification modules*. These modules serve as a resource for the creation of statistical registers and play a part in the linking of statistical registers for the purpose of generating anonymous data sets for individual special projects.

The statistical registers and classification modules together make up the complete personal statistics register system of Danmarks Statistik. The system is illustrated schematically in Figure 5, in which statistical register 1 is the unique and very direct form in which a statistical register is created, and statistical register 2 has a more complex source basis in a partially integrated process requiring the creation of a classification module, while statistical register 3 is an integrated register based on other statistical registers. The individual elements can be used in isolation or in conjunction with one another. The connection between them is as a general rule provided by the person number and in individual cases by the address code.

Figure 5 System of personal statistics



The Register of real estate statistics is one element in the system of personal statistics. It contains a key linking the owner's person number to the property identifier and this can be used to extract information on immovable property for use in personal statistics. In addition the *Register of statistics on workplaces* provides a link with business statistics through the *workplace identifier* (cf. Section 4.1.2.4). Using this identifier, information contained in the Business Register can be transferred to personal statistics to assist in the classification of persons according to their employment circumstances.

4.1.2.2 Classification modules

A small number of data sets are included in many different statistics and thus have the character of basic data. Examples would be the basic information on family circumstances, education and labour market experience of individuals. These data are processed centrally and are fed into the statistical system in the form of classification modules. A new version of the classification modules is normally produced each year. They are in the nature of statistical registers whose purpose is to supplement other statistical registers and other statistical functions, providing certain relevant background information for general use by the direct transfer of module information to a statistical register or by the processing of data already contained in statistical registers according to particular rules, typically by the use of conversion keys. If the information merely has to be incorporated into an anonymous statistical product, the module in question will be fed into the linkage process by which the basis for the product is created.

The following six classification modules are used in the statistical system:

- The *Population/Date Module* defines the population on a given date and forms part of the Population Statistics Register.

- The *District Coding Module* defines geographical areas on the basis of address codes.

- The *Household and Family Module* groups persons into families and households in accordance with the definitions applicable.

- The *Educational Classification Module (UKM)* is used to classify persons in terms of the highest level of education attained or, where appropriate, in terms of vocational training completed or education in progress.

- The *Employment Classification Module (AKM)* classifies individuals in terms of their most relevant employment (employment status, occupation and branch of their main place of employment) and the extent of their employment or unemployment in the course of a calendar year.

- The *Workplace Module* identifies the individual employment relationships of employees with a person number and workplace identifier, thus making it possible to classify persons in terms of individual enterprises and to classify enterprises in terms of workforce composition. In addition, the address code of the workplace can be compared with the home address of the employee, thus creating a basis for commuter statistics.

The function of the classification module as a general basis for information for frequent use in different fields of statistics presents a number of advantages. The main ones are: better data quality, greater effectiveness and consistency of different statistical presentations.

The statistical concepts must as far as possible accord with the world described by the statistics. The quality of the statistical information is evaluated with reference to this criterion. Information drawn from administrative systems may, taken in isolation, be inadequate as statistical data to meet specific needs. From the available administrative information, the most relevant data that can be transferred directly to the classification module are selected. Some data, however, are combined with others to form new variables which are considered better able to describe the characteristics of interest. Often the data to be combined in a new variable come from several different source registers. In order to form combined variables, data must be processed in accordance with a fixed set of rules which are formulated in terms of the definition of the concepts. This process makes for a high level of quality in the use of the information. The module must be managed so that necessary changes in the administrative basic data are taken into account, and the necessary investigations carried out, applying the rules for the processing of the data in such a way that concept definitions as far as possible remain consistent with those for previous years, provided, of course, that they are still relevant.

In the establishment and management of the module, those procedures are applied which are necessary to create the desired background variables once and for all, which makes for greater efficiency in the production of statistics. In this way the often very complicated and demanding data processing associated with these procedures only has to be carried out once instead of every time data from the classification modules have to be used for a given set of statistics.

Use of the modules' standard classifications ensures that the various statistical reports of which classifications of the type in question form part remain consistent. Definitions, times of updating after

changes etc. will be the same. The consistency achieved is of very great importance for the usability of the statistics. It also has the advantage of facilitating documentation, since reference can be made to the module. In addition, users of the statistics are better able to familiarise themselves with the concepts involved.

Classification modules must be so designed as to provide the broadest possible coverage with regard to the use of the classifications. This means that there must be flexibility in the classifications of the modules, so that adjustments specific to the task in hand can be made. It also becomes possible to ensure, to a certain extent, that application of the modules does not lead to a rigid system forcing users to adopt classifications which perhaps do not entirely suit their purpose. The demand for flexibility may seem incompatible with the concern for consistency from one statistical report to another, but there is a wide range of fields in which a selected standard classification both can and should be used. Where that is not relevant, it should be possible to use elements of the module, provided such use is compatible with the purpose of the report.

An example here is the use of the Employment Classification Module (AKM) in the labour force statistics. In the standard classification of the AKM, employee groups are defined on a whole-year basis, taking into account what their main activity was in the course of a year, whereas in the labour force statistics (statistical register) employees are defined in accordance with rules relating to a given date. This difference in the definition of employees does not prevent the AKM job classifications from being transferred to and used in the labour force statistics. Job classifications are created for all persons for whom the basic material allows this to be done, both employed and non-employed persons. This information is therefore also available for employees defined differently.

The balance between consistency and flexibility must be such that the same standard classification is always used in different statistics, unless there are material grounds for using a different classification for particular statistics. In such cases it must be possible to account for the differences.

4.1.2.3 Statistical registers

The sources for the data content of the statistical registers may be administrative registers, classification modules, other statistical registers or data collected directly through interviews or postal questionnaires.

As already stated, the individual statistical register is defined in relation to a particular subject area, so that the definitions of the content and population of the register take account of what is relevant within this particular area. Up to a point, the subject definition of the administrative registers coincides with the classification of the combined statistical basis in the statistical registers, which means that in the case of many statistical registers there will be one main source and one or more other sources providing supplementary information. The registers defined in terms of subject areas or case types in the local authorities' joint computer systems and their use, especially within the various fields of social statistics, are examples of this.

The apparently clear connection between the main source and the statistical register does not mean, however, that the administrative information can be used straightaway. Problems may arise, for example, where time-related information is concerned. In administrative practice, the time reference may be determined by administrative action which takes place after the time of the event or observation to which the statistics in question relate. Furthermore, concept definitions in the administrative system may differ from the statistically relevant concepts in a way that directly reduces the quality of the statistics, in certain cases to an unacceptable degree if the administrative information is used in its raw form. Finally, there are instances of inadequate updating, with

consequent errors and defects in the source information. This is observed in administrative systems with mixed automatic and manual reporting, or where errors in the information supplied do not have administrative consequences.

An example of the latter is the income tax system, in which it has hitherto been of no administrative importance how a person's basic income ("A-income") is specified. The tax assessment will give the same result.

It is thus often necessary, when setting up and updating a statistical register, to draw on several sources, in order to ensure the correct timing of events, to achieve the most relevant concept definitions and, as far as possible, to compensate for errors and deficiencies.

Statistical information is collected directly in connection with surveys of representative sections of the population, either by postal surveys in which the respondents themselves fill in a questionnaire received by post and return it to Danmarks Statistik, or by computer-assisted telephone interviews. Examples of interview surveys are the Labour Force Survey, the Household Budget Survey (FU) and the Omnibus Surveys. Postal surveys are not used today within the field of personal statistics but are the main method of collecting business statistics. In the production of statistics of this type, use is made, where possible, of information which already exists in various statistical registers. Of course, this involves an assessment of the quality of the register information, since the purpose of direct contact may be to compensate for inadequate register data. Statistical registers have a special role to play in connection with estimation procedures, the purpose of which is to correct skewness caused by non-response in interview surveys.

The need to investigate causal connections between different phenomena has led to the establishment of registers with horizontally related content. The subject-based structure of the register system means that new linkages must be made if a specific need arises to see subjects in relation to one another, and hence to obtain data compilations beyond what is offered by the individual register. Tasks of this nature have been and are still being performed on the basis of anonymous data sets derived from linkages, but the possibilities of analysing processes which the system now offers have aroused considerable interest, especially in research circles, and this has led to the creation of "*integrated registers"*. Registers of this type are, on the one hand, part of the general statistical apparatus and, on the other, a consequence of special data needs arising from particular service assignments.

A special form of basic data in the statistical register system consists of *supplied data sets.* For research projects conducted outside Danmarks Statistik, data may have been collected which it is desirable to link with information in the registers of Danmarks Statistik to make for a better understanding of a particular subject. If person numbers are included in the primary material, such linkage is possible and the research material can be supplemented by other information or classifications. The product of the linkage does not, however, form part of the register system, since it will be rendered anonymous. Hence it can only be used in isolation for analyses connected with the specific research project in Danmarks Statistik or for special data processing operations.

Having a number of statistical registers divided according to subject or designated for particular purposes is open to question from a structural point of view. The immediate consequence of such a situation is a considerable degree of redundancy in the system. The same information on a given unit will appear in more than one register and, in the case of certain information, in a good many registers. The expediency of this must be viewed from a number of angles. From a technical/economic point of view, it is obvious that a given item of information should only be recorded in one place. This will save space and documentation and management will be simplified. Availability for use is somewhat more difficult to assess, since it will depend on the technical solution chosen and the conditions of

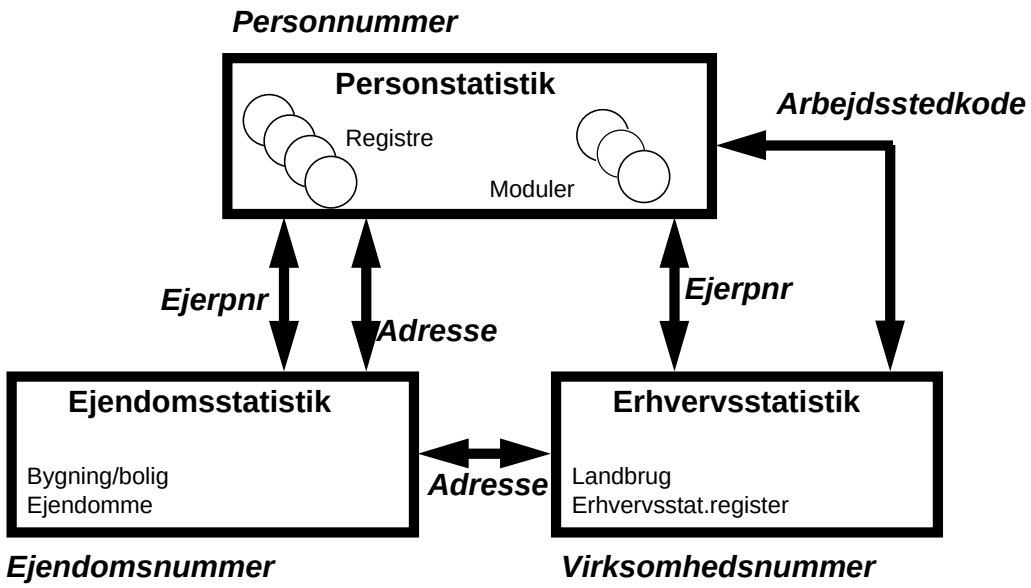
access. The latter are a particularly important element in the argument favouring the present system. One consequence of dividing the system up into a number of registers is that it has to be specified what each of them may contain, what it may be used for and who may use it. Restricting access to a register, in particular, ensures that the circle of users is small and that the same person does not have access to many registers. Different statistical registers can only be linked after a formal decision has been taken, and particular rules must be followed. This makes for transparency in the use of different information in a statistical context and minimises the possibility of breaches of security. At the same time availability is good, since the individual register is immediately usable within the area it covers. All things considered, Danmarks Statistik takes the view that the organisation of registers by subject is most appropriate and should be maintained. The concern for data protection is, of course, also crucial to this assessment.

4.1.2.4 Structure of the combined statistical system

The basic element in the statistical system is the person number, which is the common identifier for statistical units in the registers of personal statistics. Using this number it is possible in principle to combine any information. In practice this does not happen, since the linkage of information covers only those data which are relevant to specific statistical tasks. There are limits to what can be combined, since combinations of information must be objectively justified and it must not be possible to acquire knowledge of individual persons through the statistics.

As already stated, the system also includes keys to property/housing units and workplaces via the address code and the workplace code. It is thus possible to transfer statistical information on housing and business circumstances to personal statistics, and personal information can be transferred to housing and business statistics. Figure 6 shows the combined statistical system in diagram form, with an indication of the identifiers which link elements of the system together.

Figure 6 The combined statistical system



Key to Figure 6

Personnummer = Personal identification number (PNo)

Personstatistik	=	Statistics on persons
Registre	=	Registers
Moduler	=	Modules
Arbejdsstedkode	=	Workplace code (identification number)
Ejerpnr	=	Owner's PNo
Ejendomsstatistik	=	Statistics on real estate
Erhvervsstatistik	=	Statistics on businesses
Bygning/bolig	=	Building/dwelling
Ejendomme	=	Real estate
Landbrug	=	Agriculture
Erhvervsstat. register	=	Register of business statistics
Ejendomsnummer	=	Identification number of real estate
Virksomhedsnummer	=	Business identification number

Address codes link personal statistics to real estate statistics and real estate statistics to business statistics. There is a person number link between real estate statistics and personal statistics in the case of owners of immovable property. Similarly, there is a link between the business statistics and personal statistics for business owners. Finally, there is a connection between employees and business enterprises through the workplace code, so that information on persons employed can be linked to information on places of employment.

4.1.3 Linkage of registers

The linkage of registers is the process or method by which information from different statistical registers on a particular statistical *unit* is linked together or by which information on different *units* which have particular relationships is linked together. Linkage is also called integration.

The purpose of this interlinking of information is to acquire new knowledge on the units we are interested in. By means of the linkage, existing information is processed and this enhances the value of the information.

The linkage of statistical registers is of inestimable value as an alternative to traditional data-gathering. For linkage to function satisfactorily, however, the identification of persons and other units in the data extracted from the source registers in question must be clear and well-defined. If there are deficiencies in this respect, the cost of correction can be considerable.

Linkage will also reveal errors and inconsistencies in other information, and a decision must be made as to what corrections are needed.

In decisions regarding the correction of errors, it is important to be clear about the objective of the linkage, i.e. what the linkage must help to elucidate. Error correction will aim in the first instance to locate combination errors which arise in the process of linkage itself. It is not necessarily the case that this will provide edited data which will automatically be suitable for any other linkage that *might* be contemplated.

Linkage or integration may arise in many different physical/logical manifestations; it may not necessarily result in any permanent physical linkage. Four different types can be distinguished:

– *Logical linkage.* This refers to the situation in which it is decided, on the basis of a logical analysis, that it is possible to integrate two registers. The result of logical linkage is a diagram showing the relations between the logical units.

– *Ad-hoc linkage.* In order to compile a particular set of tables containing information from more than one register, the registers in question are linked in a single operation for the purpose in hand. A variation of this arises where a permanent facility exists for the compilation of tables of a certain type from a given set of registers. In both cases, physical integration exists only while the data processing is in progress. Two statistical service systems, the Local Authority Annual Service (KÅS) and Market Statistics, are examples of this.

– *Permanent linkage.* All relevant information on the units in question is linked once and for all in a permanent register. The Register of Statistics on Hospitalisations is an example of this.

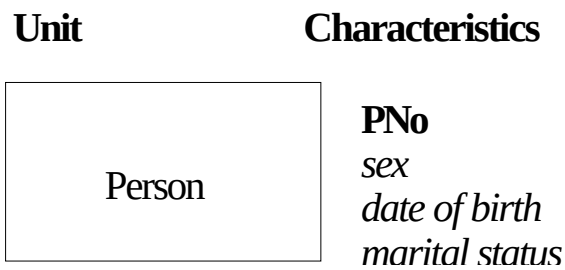
– *Creation of a database.* If the registers which are to be linked are set up along the lines of a relational database, further physical linkage may become superfluous, since only logical linkage is necessary (see above). The required relationships between units are determined across the registers by logical linkage with the inclusion of relevant keys which can then be used for searches of the database. Of course, this means that the registers must be set up in formats which are accessible for the relevant query language. One variation of this form of integration would be those cases in which registers are described in a common meta-information system - where they are defined in terms of logical linkage and information is provided on their physical storage.

What type of linkage is to be preferred depends on the objective of integration in specific instances, the quantity of data involved, facilities required, resources, machines, software, etc. Ad-hoc and permanent linkage are the most widely used forms today, but databases are likely to be more extensively used as costs associated with physical storage media come down and database tools become more effective.

4.1.4 Content analysis

Some important concepts relating to content analysis are discussed below.

Figure 7 Illustration of a statistical unit



Units are the persons or objects about whom or which we wish to present statistics, i.e. enumeration units. A unit is represented graphically by a rectangle - cf. Figure 7.

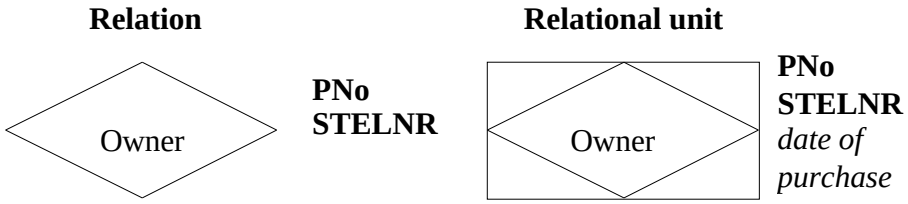
Characteristics or *variables* are used to describe the *unit*, e.g. age or magnitude. A distinction is made firstly between characteristics that identify the individual unit, the *key* or *identifier* (e.g. person

number), and secondly characteristics that describe the unit in other respects, *descriptive characteristics.* Characteristics can be *qualitative,* e.g. sex (male, female) or *quantitative,* e.g. income (DKK 200 234) or age (29 years).

The *key* is shown in the diagram in bold (e.g. **PNo**) and descriptive characteristics in italics. If we wish to compare information on a unit from a number of different sources, it is the *key* that we use to establish the link. A collection of units with particular characteristics is referred to as a *unit group* or *population.* If we wish to investigate links between different units, i.e. relations, we use the keys from the units involved.

A relation characterises a link between two or more *units,* for example "a person resides in a dwelling" or "a person owns a bicycle". A relation is identified by the combination of the keys of the units in question, e.g. "person: PNo owns bicycle defined by STELNR' (STELNR stands for "frame number". In Denmark, every bicycle sold is registered with the owner's name and the bicycle frame number). A relation is represented graphically by a rhombus as shown on the left in Figure 8. The identification of the relation consists of the identifiers of the units involved in it, in this case PNo and STELNR.

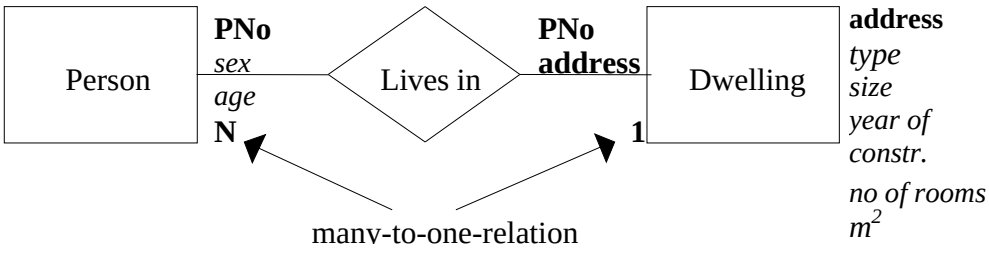
Figure 8 Illustration of relations and relational units



In addition to the keys, a relation may contain specific information on the nature of the relation. In the example above, the relation contains information - shown to the right - on the date of purchase of the bicycle. If a relation contains information additional to the keys involved, it is referred to as a *relational unit* and is represented graphically as a rectangle with a rhombus inside it.

The logical connection between *units, characteristics* and *relations* is illustrated in Figure 9.

Figure 9 Illustration of a unit diagram



Relations between units are often divided into the types *"one to one"*, *"one to many"* and *"many to many"*. This indicates how the units relate to one another numerically. One person may have at most one spouse, whereas one person can own several bicycles. Many secondary schools offer many secondary education courses. "One" and "many" relations are denoted by 1 and N.

The figure, which shows the link between persons and dwellings, may be read thus: persons live in dwellings; many persons may live in *one* dwelling, and one person can only have *one* domicile; the specific persons who live in specific dwellings are determined by the PNo and address code.

4.1.5 Creation of statistical registers

Almost all person-related *statistical registers* at Danmarks Statistik are, as already stated, produced on the basis of *administrative registers* maintained by central or local government departments, possibly supplemented by information collected manually from other sources (e.g. statistics on live births: CPR and registrations of births).

The administrative source registers may be functional registers used on a day-to-day basis, which are dynamic registers in the sense that they continuously receive *transactions* (updates) of information on *events* affecting the register *units*, so that the register at any given time reflects administrative reality. The source register may, however, also be a centrally integrated register that periodically receives information from the functional registers. For example, the National Patient Register is a register of the Department of Health which regularly receives extracts from the patient registers of individual hospitals.

From the source register, Danmarks Statistik receives extracts for statistical processing, which in many cases (for practical reasons, among others) only reflect the state of the source register on a particular date or for a particular period. Danmarks Statistik typically receives an extract one or more times a year from the administrative source register which contains the relevant information. The extract thus represents the content of the source register only on the date in question or for the period in question. Three main types of extract may arise:

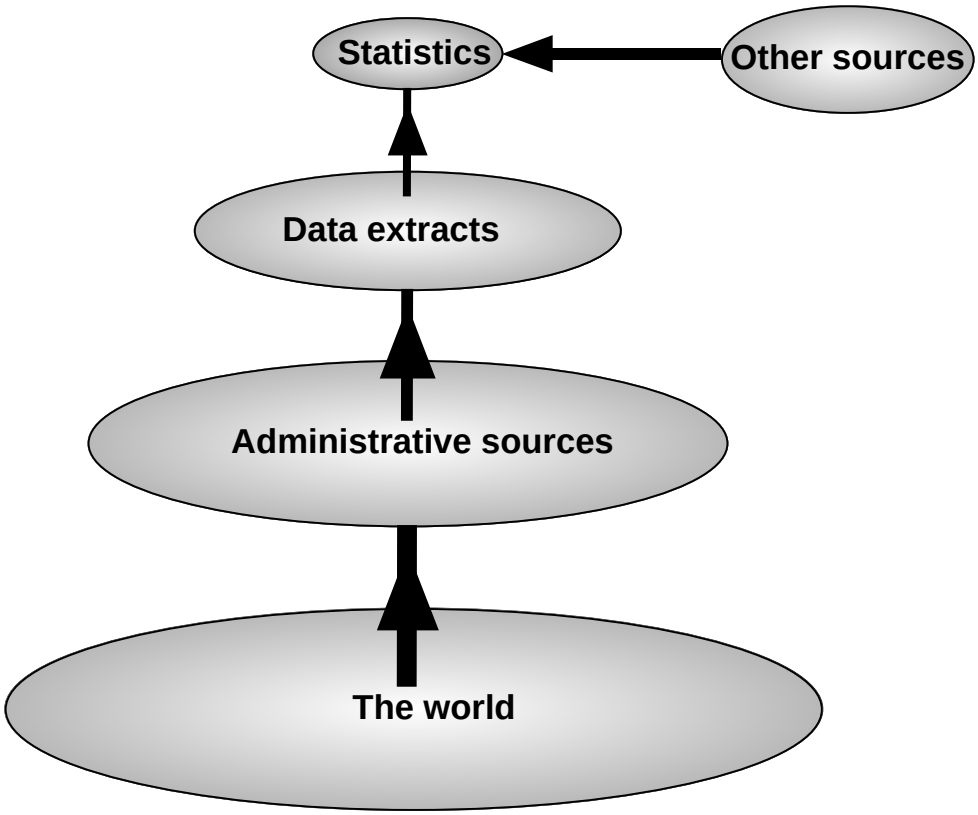
– extracts for the *status of the population on a given date* (e.g. persons drawing a social pension and the amount of the pension in January);

– extracts for the status of the population on a given date, supplemented by information on *the trends in key characteristics over a period* (e.g. Social Assistance Act statistics in the year 1992);

– extracts for *transactions covering all relevant units over a period* (e.g. use of hospital services).

The extract is edited, and often compared with other information and registers, finally ending up as the completed *statistical register*. This is an elaborate and complex stage in the process.

Figure 10 From the phenomena of reality to statistics



The statistical register must be able to describe part of "the world". As we can see from Figure 10, it is a long way from "the world" to the statistical register and there may be many slips along the way.

One frequent problem is that we have information only for given periods or dates. The task may be

- to describe a continuous process on the basis of two status extracts by way of additions and withdrawals, or
- to compare status data from a single date with continuous data over a period.

4.1.6 Administrative data as a source of statistics

The rest of Section 4.1 discusses the requirements which have to be met to enable official statistics providing information on society to be based successfully on administrative registers, what types of data must be considered key data and how different types of data can be included in the basic statistics. We also discuss a few practical problems arising with corrections to register-based statistics and how they may be solved. Alternative solutions are also discussed, for when the information required is not directly available in the administrative registers.

The requirements which must be made of the basic data, i.e. the administrative registers, have in part to do with the coverage of the data in terms of content and in part with qualitative aspects. It is impossible to lay down very precise requirements which should be met in all circumstances if worthwhile statistics are to be produced on the basis of administrative registers. It is also clear that even isolated individual administrative registers can be used, and are being used all over the world, as

a basis for statistics, and some of the oldest statistics in Denmark were based on records such as church registers.

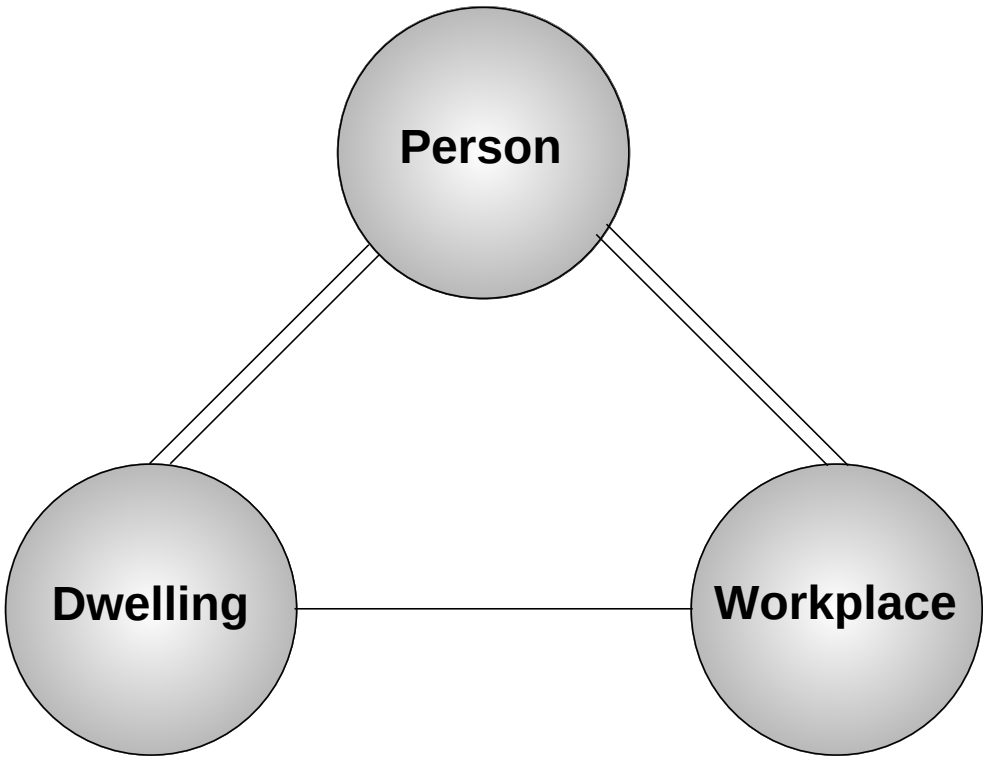
In the following sections some requirements are set out which must be met if there is to be any possibility of creating a *coherent, comprehensive system of personal statistics* based mainly on registers, i.e. a system of the kind developed in Denmark.

4.1.7 Content

First and foremost, the registers must contain data which cover the most important subjects in a statistical system, to elucidate patterns and trends in society.

In addition, there are many different units which it may desirable to study through statistics, but three central units are essential to the structuring of the statistics: persons, enterprises and dwellings. In personal statistics, persons are, of course, at the centre, but the other two units are also relevant to this field of statistics. This is shown, for example, by the fact that a traditional population and housing census also records data on the homes in which the people live and the enterprises (workplaces) in which they work. A large proportion of users of statistics, in both the public and the private sectors, need information on precisely these relationships.

Figure 11 Identification relationships



There are, of course, many other units of which it is important to have records. Some of these units are *events* which relate to persons, for example traditional demographic events (births, deaths, etc.) or developments of various kinds in social circumstances. Among the other units of significance, motor vehicles feature prominently.

For statistical use, it is important for the units to be well defined. This is not a problem for "natural" units such as persons or motor vehicles which, compared with other units, more or less define themselves. In addition, the concept of a home or dwelling is fairly easy to deal with, although there are some borderline cases in which the existence or specification of a dwelling can only be determined on the basis of detailed rules. Some units manifest themselves as combinations of natural units, e.g. families and households.

It quickly becomes more problematic when other types of unit are considered, and here business units are a good example. An enterprise or workplace is an abstract concept. Its existence can only be established and its nature defined in accordance with very precise rules. Lack of clarity and disagreement over such definitions have for decades rendered difficult the use of registers in business statistics, and these problems were not solved until the start of the 1990s, with the compilation of the present business register.

Registers should, ideally, contain information on all units in the country of the types specified, i.e. they must be exhaustive. Incomplete coverage will often be systematic. For example, if certain municipalities are not covered, the statistics produced on such a basis will immediately be affected by skewness, for which it may be difficult to compensate.

One of the great strengths of register-based statistics, as compared with traditional questionnaire-based statistics, is that it is often possible to achieve total coverage, or something very close to it. In questionnaire surveys, a representative sample of the population usually has to suffice, and a degree of non-response is inevitable. More often than not, this skews the results in such a way that those who are covered by the survey differ systematically from those who are not. The rate of non-response is often significant, typically from 15 to 50%, when participation is voluntary and the survey is conducted among private individuals. In some surveys in other countries, rates of non-response ranging up to 80% have been observed, which in most contexts make the results entirely worthless.

An important precondition if statistics based on registers are to give comprehensive coverage is that the data contained in administrative registers should be extensive and should cover many variables relating to the relevant units. Gaps either necessitate the additional collection of information using traditional methods, or limit the content of the statistics.

4.1.8 Identifiers

Identifiers play a very considerable role in both the maintenance of administrative registers and their statistical use. For statistics, it is particularly important that it should be possible to employ the common identifiers used in more than one register, in order to link information on a given unit which comes from various sources. Since the administrative registers are not set up for statistical purposes and therefore do not usually contain the desired data combinations in directly accessible form, an important part of the production of almost any register-based statistics is to link data from different registers by means of common keys.

The most important key in this connection is, of course, the *person number*. The *street address* and *workplace number* are also essential in determining relationships between persons, dwellings and workplaces.

4.1.9 Time references

The time dimension plays a very special role in statistics revealing patterns and trends in society, and in all areas of coverage it has to be possible to make comparisons over time. In statistical publications or databases, the time dimension is, broadly speaking, ever-present. It is therefore vital for statistical usability that reliable information on various dates be included in the registers.

This applies first and foremost to the dates of changes or events. Among the main events of interest here are the "births" and "deaths" of units, but it is also important to date other changes affecting units. What we are concerned with here is the actual date on which an event leading to a change in an item of data took place, e.g. the date of a change of address or of the branch of an enterprise.

It may be extremely difficult in practice to pinpoint the real date of an event. The branch within which an enterprise operates is a variable which may change gradually, as the activities of the enterprise shift their emphasis, and it is perhaps not possible to determine exactly when a new activity becomes dominant. But in the case of many other variables, too, which in principle can be observed at any given time and in which a change can therefore be dated with reasonable accuracy, problems arise in practice in the recording process. This may be because the time when a particular value became effective may not necessarily concern the administration using the data: all it needs to know is the most up-to-date value at any given time. In other cases problems may arise because it is not possible or practicable to record events in real time: it may only be possible to ascertain data changes from responses to enquiries, for example once a year in connection with a tax return.

If it is not possible to record accurate dates, approximations remain an option, of course, and are much better than nothing at all.

In addition to dates of events or changes there is a need for *registration dates*, i.e. an indication of when the data value in question was entered in the register. This is of crucial importance when the register has to function as a *statistical model* of the reality the statistics are to describe - see Section 4.1.10.

The ideal situation, therefore, is that any item of information in the administrative register should be accompanied by two dates: the *date of an event or the date on which a change took effect* and a *date* of *registration.* Registers in reality often deviate significantly from this ideal.

Delays in the updating of administrative registers mean in many cases that several editions of the figures have to be published: *provisional figures,* which have to be issued because of the need for current figures but which lack precision due to inadequate updating, and *final figures*, which are more extensive but do not appear until some time later. Sometimes there may be a number of interim versions. The existence of several versions of the figures may, however, cause great confusion and inconvenience to users and should for that reason be avoided if at all possible.

4.1.10 The register as a statistical model

When a register system is to be used as a basis for statistics, it must be possible to view the register as a model of the group of units the statistics are to describe,

In some cases, the data changes recorded in the register may also need to be viewed as models of events described in the statistics. This applies, for example, to population statistics, where it should be

possible for reports of events in the population (when people move house, marry etc.) to be derived from data changes in the Central Population Register (CPR). Registers must therefore be set up in such a way that reports on the population covered by the register and on movements in it can be pinpointed with respect to the dates on which the events actually took place.

For statistics on movements, it must be possible to make a distinction in the register between real events and corrections and data changes of a more technical nature. The content of this latter requirement can best be illustrated by certain problems which have actually occurred in the use of the CPR for population statistics, even though the CPR must be said to meet the needs of statistics production to a very high degree.

Actual corrections of wrongly reported data pose a problem for statistics, however, even though it is possible to distinguish them from real events. Corrections have to be regarded in many cases as modifications of events reported earlier: an address correction may be a correction to a report on a removal. A correction may also cancel an event reported earlier, which occurs in the CPR system, in particular, when people who have notified the population register that they are to emigrate change their minds and remain in the country.

Corrections to events are difficult to manage because, in reality, they can be reported with random time-lags. In the population statistics, it has been decided to exclude them entirely from reports of movements, but this procedure is only feasible because the number of corrections is small compared with the number of events.

4.1.11 Stability

An important property of any statistics is that they can describe a process over time, i.e. can show how a particular magnitude develops from month to month and from year to year. It is also of great importance, therefore, that concepts in the administrative registers should remain constant over the longer term. Otherwise, major problems can arise in securing comparable figures from one period to the next. In some cases, but far from all, it is possible to adjust for changes with greater or lesser precision. There are also examples of situations in which it was found necessary to abandon the publication of certain statistics in a particular period because it was impossible to adjust for the consequences of changes in registers.

4.1.12 Quality

The quality requirements imposed by statistical use coincide to some extent with the requirements which must be met if the primary purpose of the registers is to be served.

The first quality requirement to be met by basic data is that they must be *relevant* to the statistics, i.e. the data must relate to the concepts on which the statistics are to provide information. It is much easier to check for relevance in traditional statistics which are based on data specially collected for the purpose. When the statistics are to be based on data collected for a different purpose, whatever is necessary or desirable for that purpose usually has to suffice. In favourable cases, however, the statistician may be fortunate enough to exert an influence on the content of registers.

It is also possible to let different data sources supplement one another through linkage, so that the desired information is provided by a comparison of a number of data from different registers which document the subject more or less well. The use of several data sources is covered in more detail in Section 4.1.13.

On the subject of relevance, it must also be borne in mind that a statistical system which is exclusively or to a very large extent based on the registers of public authorities may end up giving a picture of the world as seen through the eyes of those authorities. Concepts or groups of persons which do not exist in the frame of reference of the tax authority or social affairs administration may also not show up in the statistics. One example might be homeless persons.

To exclude this risk, it is important that there should be certain basic registers, such as the CPR, whose task is to record all units without reference to any specific administrative purpose. We also know for certain that many records from different authorities can be linked so as to provide statistics which are not tied to any single administrative viewpoint. At the same time, it is important to bear in mind that the registers cannot or must not give us a complete picture of society, and that it is therefore necessary to collect data specifically in order to reveal aspects which are not covered by the registers. These additional data can be collected through interview or questionnaire surveys, and they should ideally include identifiers which make it possible to compare them with the register data.

The next requirement to be met by basic data for statistics concerns reliability. This applies whether the data come from questionnaires or registers. One must be as certain as possible that the recorded data faithfully reflect the circumstances investigated. If a person is stated to be a painter and is in fact a government statistician, there is a risk that incorrect conclusions will be drawn from the final statistics.

It can, of course, be argued that, as long as the errors are not systematic, they will not result in skewed boundary distributions, since errors in opposite directions have a tendency to cancel one another out. But when, as is often necessary, we look at *statistical links* between several variables, such errors become a problem, since they can distort the assessment of those links.

If the errors tend systematically in a particular direction, the problems are immediately magnified, but their systematic nature may make it easier to correct them.

Finally, it is important that data be recorded with a degree of precision suited to the needs of the statistics, i.e. the scale used should be sufficiently detailed. For example, the degree of specification in the occupational information must be fairly detailed if the information is to be used for epidemiological studies, where the risk of disease is to be viewed in terms of job-related effects.

The need for precision remains even when the information is not to be used in statistical reports at the recorded level. In many cases, groups are formed on the basis of variables derived from discrete items of data. The definition and application of such variables in statistical contexts depend crucially on the precision of the basic material.

4.1.13 Use of multiple sources

When data are collected for statistical purposes using traditional collection methods, the questions and guidelines for answering them are formulated in such a way that information is obtained which accords with the definitions used in the statistical field concerned. The information-gathering can be driven, so to speak, in such a way that we obtain precisely the information we need.

The use of administrative registers as a basis for statistical data does not provide the same scope for control. It is possible to search the information available and select those items which best suit the intended purpose. It will often be the case, however, that even the most suitable information does not cover the immediate need to a sufficient degree, since there may be divergent definitions, limited

coverage, errors or inadequate updating. In such a situation, one possibility is to draw on information from several administrative sources.

Where the need for a specific item of information for one or more statistical functions cannot be covered immediately by direct reference to the information available, the problem may in certain cases be solved by the formation of "*derived variables"*. Such a variable is obtained by chaining together different items of information which, together, provide the information required, if not with absolute precision at least with good approximation.

When a derived variable is to be set up, the starting point is to establish a definition of the concept one wishes to use in the statistics. Administrative information is then sought - possibly information which is already used for statistics - which can help to meet the definitional requirements. A distinction can be made between three types of variable formation: *creation of a new variable, prioritisation of information for a given variable* and *adjustment or correction of specific information.*

The creation of a new variable may be the solution if enough of the information required is available but is not contained in a single item of administrative data. Several items of information, collected and processed in accordance with administrative rules, can each contribute to providing an overall description of individuals in terms of the desired statistical concept. The process can be compared with a jigsaw puzzle in which each piece forms part of the total picture but in isolation does not give an impression of the complete picture.

One example of a new variable of this kind is the employment status code, which forms part of Danmarks Statistik's Employment Classification Module. This classifies persons according to their main form of employment in the course of a calendar year and comprises the following groups: sole proprietors, assisting spouses, wage- and salary-earners, pensioners and other persons not economically active. Information such as this is clearly of a statistical nature and will not be found in the same form in administrative registers. It includes, for example, income received in a breakdown by type and, where appropriate, the obligation to submit accounts, derived from the registers of the income tax system, and possible details of business ownership derived from the Central Business Register. Such information is suitable for use in forming an overall assessment of the main form of employment for a given individual during the course of a particular year. It provides a number of indicators which are weighted by means of a set of rules, resulting in the classification of persons in terms of employment status.

In most cases this classification is quite uncomplicated. All indicators will point to the same group. It may be, for example, that the only classifying income consists of wages received and that the person concerned is not the owner of a business. This person will be classified as a wage-earning or salaried employee, provided the wages or salary received are of a certain amount. Otherwise, the person is classified as being not economically active.

It becomes more complicated when more than one classification is possible. It may be that a business owner has received salary and pension payments in the year in question. In this case, the information will be weighed up carefully. Ideally, the duration of the individual situation should decide the weighting, but this is not possible. The best approximation to the ideal is to weight according to the volume of the relevant types of income: business profits, wages/salaries and pension payments. That is in fact the procedure followed up to a point, but it can give rise to serious misinterpretations when the amount of work performed by a self-employed businessman is assessed on the basis of the business's profit or loss in a single year, since a loss can arise despite considerable labour input. Substantial weight is therefore assigned to information on business ownership from the Central

Business Register when self-employed persons are defined. This is done by classifying the business owner as self-employed unless any salary drawn is greater than the profit of the business and the ATP contribution indicates full employment over the entire year as a wage- or salary-earner or if the profit of the business is zero or less and the person in question received a pension. These two exceptions give rise to classification as an employee or pensioner, respectively.

When a new variable is formed, it is essential for the necessary components to be present for all persons, and that, for the intended purpose, quality is good. In the examples quoted, the information on business ownership can be found only for persons who, according to the rules, must be included in the Central Business Register. These are employers and persons subject to VAT. A small group of businesspeople are not included in this population - for example, persons operating within the field of passenger transport and providers of certain health services. In this case, information on the obligation to submit accounts is used, and greater significance is attributed to profits or losses recorded by the business. The separation of wage and salary income from other types of income taxed in the same way is influenced to some extent by the fact that the distinction is of no great relevance to the assessment work of the tax authorities, so that a certain degree of uncertainty arises here. It may be concluded, however, that, despite the deficiencies indicated, the result is in good agreement with the definition laid down.

The use of information from multiple sources may be the solution when the desired information is not immediately available in a form suitable for the statistical purpose. It must be realised, however, that, even if we take as our starting point a definition fixed on the basis of the statistical requirements, the data available and the processing which can be given to them may very easily lead to departures from this definition. When selecting information and determining the rules for its processing, it is important to take great care to ensure that the deviations are not too wide. The result of this could be that the statistics convey something other than what they are believed to convey.

The actual definitions associated with the derived concepts may often be somewhat ingenious, since they reflect the set of rules applied. This may create unnecessary confusion in the minds of the users, which is why they appear in a more simplified form in their presentation, as is also the case with traditional data collection. When information from multiple sources is used, a more differentiated picture of the individual's circumstances can be obtained. For example, the COR may provide information on wages earned by a particular individual and the Business Register may show that the same person has been active as a business owner. The set of rules must determine how the information should be weighted when such a person is classified in terms of employment status as an employee or as a self-employed businessman.

These differences in information also arise in the interview situation, but they are known only by the interviewee, so it is left to him or her to choose the classification and hence the weighting of the information. This choice can be highly subjective, since it is impossible to indicate the precise delimitation of individual groups in a classification. A person is classified as an employee if he or she has indicated as much in answer to a question.

For presentation purposes, register-based employment status is defined as a person's main activity during a given calendar year. Further details can, of course, be made available if special investigations call for an analysis of the demarcation lines between different groups.

4.1.14 Integrated data collection

When an administrative authority obtains information on behalf of another authority at the same time as it collects information for its own use, this is called *integrated data collection.*

This means that an administrative authority collects information from firms, institutions etc. which it will itself use to perform its own tasks, while at the same time collecting information which it does not itself need but is to be used by another authority, such as Danmarks Statistik. The two data sets are collected in one and the same integrated operation.

The purpose of integrated data collection is to limit the costs of data-gathering. The total cost to society is reduced because it is far cheaper to gather information in a combined process than in several processes. This applies, in particular, where the source information does not exist in computerised registers and where resource-intensive data collection methods using questionnaires or interviews therefore have to be used.

An added advantage is that integrated data collection helps to ensure that the information gathered is consistent and covers the same units, which would not necessarily be the case if collection had been undertaken by two institutions in separate processes.

4.1.15 Legal basis

The *Act on Danmarks Statistik* contains provisions on what data may be collected by the institution but gives no guidance as to collection methods. Thus the Act specifies what information or types of information can be collected from business enterprises, institutions etc., but it does not state whether the information is to be collected by means of questionnaires, whether data which institutions or businesses have in their possession are to be supplied directly to Danmarks Statistik by electronic data transfer, whether information is to be collected by another authority using integrated data collection or, indeed, whether some entirely different method is to be used.

The *Public Authorities' Registers Act* contains provisions on the processing of personal information on electronic media. The Act aims to ensure that the public authorities set up and use computerised registers containing personal information in such a way that the legal rights of individual citizens with regard to the protection and integrity of their data are not violated, but without any loss of efficiency for the administration. Under the Act, registers to be kept by a government authority may only be set up after approval by a minister. Before a register is brought into use, rules must be drawn up for its structure and operation. For registers to be kept by a local authority, decisions on its establishment and the formulation of the rules governing it are normally taken by the local authority concerned. Rules governing registers must be approved by *Registertilsynet* [the Data Surveillance Authority].

Part 3 of the Public Authorities' Registers Act contains provisions on the recording and storage of information. Subsection 1 of Section 9 states: "Only such data shall be registered as are clearly relevant to the purposes of the authority concerned. Other data relevant to the purposes of another authority may be registered provided the register is blocked in such a manner that these other data can be used only by that other authority." This will normally be the case where an authority collects a range of information from citizens (or business enterprises, on employees for example) for future use in its own administration, whereas the other authority has to use a particular item of information in much the same area.

4.1.16 Examples of integrated data collection - the Workplace Project, the ERE statistics

In a number of cases Danmarks Statistik needs both information collected for its sole use, which must not therefore be used by the collecting authority, and information which the administrative authority collects for its own use. In these cases, the combined set of information will make up the data supplied to Danmarks Statistik and included in its statistical registers. This applies in the following example, which is particularly relevant from a national accounts point of view, since it refers to the production of the basic figures for one of the most important sources for the income-based estimate of GDP, the ERE statistics*.

The aim of the Workplace Project - which, somewhat misleadingly, is called a "project" for historical reasons, even though for many years now it has not been a development project, but an ongoing operation - is to produce a list of all workplaces and to assign all employment relationships to the correct workplace. Information for the Workplace Project is obtained by integrated data collection using the information sheets which employers have to submit each year to *Told&Skat* in respect of each employee. The information sheets contain details of wages, pensions, fees, allocations from funds etc. for the individual employee or recipient of public benefits. The latter is identified by his or her person number while the employer is identified by the SE number, which is supplemented in the Workplace Project by a code for the place of employment if the employer in question has more than one workplace. The information sheet also provides information on the period of employment, ATP contributions and income tax deducted at source ["A-tax"], along with certain other forms of income from which tax is not deducted at source.

The Workplace Project is carried out once a year by *Told&Skat* and Danmarks Statistik working in cooperation. *Told&Skat* sends out the blank information sheets in November/December each year, together with guidelines for enterprises on data to be reported for the year. These forms are accompanied by workplace lists for enterprises that have more than one workplace. The workplace lists are produced by Danmarks Statistik and contain a list of the workplaces which Danmarks Statistik has recorded for the enterprise in question as they appear in Danmarks Statistik's business register. The lists contain the name, address and branch (branch name and code) of each workplace. Finally, they contain the code for the workplace, which must be indicated on the information sheet for each individual employee. Employers must update the list with appropriate information on the individual workplace, should there be errors or inconsistencies in the list. They must also add new workplaces and indicate others which have closed so that, after any corrections needed, the list contains complete information on all workplaces which have been active during the course of the year.

---

* *Erhvervsbeskæftigelsesstatistik*, EBS in Danish, translated in Danmarks Statistik's "Declarations of content" as "establishment-related employment (ERE) statistics".

Information on wages and salaries is collected by *Told&Skat* for the purpose of payment demands and verification and for use in assessment work, which is carried out in coordination with the assessment authorities at local government level.

The workplace codes are collected for use by Danmarks Statistik. Information is also collected on the employee's period of employment during the year in question for the exclusive use of Danmarks Statistik.

Danmarks Statistik receives a copy of all data on the information sheet, both the information collected by *Told&Skat* for use by Danmarks Statistik alone and the information collected for the administrative use of *Told&Skat* itself.

For Danmarks Statistik, the information is complete in itself. If it were to collect the information itself, the administrative information would be just as meaningful as the information on workplaces and periods of employment. The interpretation of the information collected for Danmarks Statistik alone depends on the value of the other information.

As far as the public sector is concerned, information on workplaces etc. is to a large extent already recorded in the computerised wage and salary payment systems of central and local government. The Workplace Project therefore uses information drawn directly from these systems which, however, must first be supplemented and systematised. This is a laborious process.

Information from the Workplace Project is used by Danmarks Statistik for statistical reports. In the first instance, it provides the basis for individual data in the personal statistics system covering labour market information for a whole year and for a single point in time (end November). This information is thus included in the classification basis of the personal statistics system.

However, the information also contains the crucial link between personal statistics and business statistics, since the SE number, with its associated workplace code, provides information on links between persons and specific enterprises and workplaces. This provides a basis for updating units in the business register, and the aggregate personal data update the employment information in the business register at workplace level.

Information at the personal level thus provides scope for business classifications at the level of individuals and, by way of the associated person numbers, the business information makes it possible to link personal variables from the personal statistics register system to business units.

The Register-based Labour Force Statistics (RAS) include workplace information. The RAS describe the population's links with the labour market at a given point during the year, i.e. end November, corresponding to the information on period of employment from the information sheet. The basic data are provided by the basic information from the Workplace Project, together with extracts from a number of administrative registers and statistical registers, information from which is processed to indicate labour market involvement in the last week of November of the year in question for all persons resident in Denmark on 1 January the following year.

The material from the Workplace Project and the information sheets is also used to produce the ERE, the purpose of which is to show employment from the point of view of business units. Workplaces are classified, *inter alia*, on the basis of their employment levels at the end of November. The ERE statistics were first produced for 1990 using the method described here. The definitions used are the same as for the RAS, and the two sets of statistics are coordinated**.** A basic register is compiled from which extracts are taken for both register-based labour force and ERE statistics. The latter give details

of both primary and secondary employment, i.e. a person can be recorded with two employment relationships.

One way of looking at it is that the RAS detail labour market status and employment seen from the point of view of the individual, whereas the ERE statistics show employment from the point of view of enterprises. This latter view is the one relevant to the estimate of the compensation of employees as a component of GDP in the generation of income account, which in the national accounts is a "domestic" concept, i.e. total wages and salaries in all resident producer units, regardless of whether this income goes to residents in Denmark or non-residents, whereas the compensation of employees in the allocation of primary income account is a "national" concept, i.e. total wages and salaries which Danish residents have earned in a given period, regardless of whether they come from Danish or foreign producer units.

4.2 Valuation

Both the compensation of employees and gross operating surplus and mixed income are by definition estimated at factor cost. Gross operating surplus and mixed income are calculated from an estimate of value added at basic prices as calculated in the target total module (MTM). Consequently, there are no valuation problems specific to the income-based estimate in the Danish national accounts, one reason being the integrated estimate of gross operating surplus based on business accounts. It would have been a different matter if this estimate had instead been based on corporation tax and personal income tax returns. The tax concepts of income used for these returns are often very different from national accounts concepts.

4.3 Transition from accounting and administrative concepts to ESA 95 national accounts concepts

The main statistical source for the estimate of the compensation of employees is the ERE, the establishment-related employment statistics. The "wages and salaries" concept in these statistics is not the same as the "compensation of employees" concept in the national accounts. Section 4.7 discusses the transition to national accounts concepts as part of the discussion of the processing of primary statistics. Table 101 shows the result of the transition from primary statistics to the national accounts calculation.

Table 101 Connection between the ERE and the national accounts

	Wage/salary component	DKK million
	Compensation of employees in the national accounts	534 094
	Total wages and salaries in the ERE	490 966
	of which employer-administered capital pensions	*7 209*
-	Tax value of fringe benefits	2 862
+	Long-service gratuities and severance pay	1 905
+	Gifts in kind	12
+	Wages from the black economy (1 660 + 1 235)	2 895
+	National accounts calculation of fringe benefits	6 788
+	Employer contributions to the ATP	4 906
+	Employer contributions to pension schemes etc.	26 284
+	Industrial accident insurance	1 631
=	Corrected total wages and salaries for the ERE	532 525
	Percentage discrepancy	0.29%

For gross operating surplus and mixed income, reference should be made to Section 3.3, which discusses in detail the change from business accounting concepts to national accounts concepts for the compilation of the intermediate system and the target total module.

4.4 Direct versus indirect methods of estimation

All income components other than that part of gross operating surplus for which figures are imputed (surplus on the imputed rental value of owner-occupied housing, consumption of fixed capital relating to non-market output, etc.) are in principle estimated directly as income created by the production process.

4.5 Direct estimates of levels as opposed to projections

With the exception of allowances for the hidden economy, income-based GDP is in no case estimated using projections, but is a direct estimate of levels based on total coverage of wages and salaries in the primary statistics.

4.6 Most important allowances for exhaustiveness

As Table 19 in Section 1.7 shows, there are explicit allowances for fringe benefits and the black economy.

4.7 Compensation of employees

4.7.1 Statistical sources

By far the most important statistical source for the estimate of the compensation of employees is the ERE statistics, employment and payroll statistics based on various administrative and statistical registers.

The ERE statistics give detailed information on the number of workplaces, the number of jobs and total annual wages and salaries divided by geographical area and industry. The statistical unit is the *workplace*, i.e. the local kind-of-activity unit. The most important basis for the statistics is the information form which all employers have to submit every year to *Told&Skat* for each wage- or salary-earner employed. The information form for 1995 is reproduced in Annex 7. These forms cover all employees without any lower limit whatsoever for amounts or length of employment, and both Danish and foreign residents. This makes them a virtually ideal source for the calculation of income in the form of wages or salaries other than that which comes from the hidden economy, which is, of course, withheld from the tax authorities.

A brief description of this key source in the form of a "declaration of contents" can be found in Chapter 11. Below is a detailed description of the source, taken from the publication of statistics in the series *Statistiske Efterretninger* (with a few stylistic changes in the Danish).

The ERE statistics are annual and give information on:

- the number of workplaces and employment relationships as at the end of November;
- full-time employment and total wages and salaries during the year.

The ERE statistics are an extension of the workplace statistics produced for the period 1980-1989. The extension consists in the fact that:

- workplaces at the end of November are those of all enterprises where people are employed – including those where only the owner is employed;
- employment relationships at the end of November include the self-employed and assisting spouses as well as employees;
- employment relationships at the end of November can be described in terms of the following variables: employment status, sex, age and education.

The 1995, ERE statistics were compiled from the same basic material as was used for the register-based labour force statistics (RAS) on 1 January 1996, and the two sets of figures are coordinated.

Those on leave of absence as at the end of November 1995 are considered by the RAS to be in employment, but are not so considered by the ERE statistics. Of the self-employed, the ERE statistics count only those who are either liable for VAT in connection with their primary link with the labour market or are employers.

The basic data are derived from various administrative and statistical registers - primarily information given by employers to *Told&Skat* on wages and salaries paid out and ATP-contributions paid in (Salary Information Register) - from the business register and from the employment classification module. For further processing and definitions, information from various registers is added.

Since the ERE statistics provide detailed information on the number of workplaces and the number of employment relationships divided by geographical area and branch group, updated information is required on all employment units right down to workplace level. The statistics are updated annually in connection with the Workplace Project.

For employment relationships with *private employers*, the Workplace Project is based on the information forms which all employers have to send in each year to *Told&Skat* for each wage- or salary-earner employed.

The forms identify both the employee (person number) and the employer (employer number). For employers with more than one workplace, a workplace code is used to mark which of the workplaces the information form refers to.

Every year, Danmarks Statistik sends out with the information forms to each employer who has more than one workplace an "*Arbejdsstedsfortegnelse*" [workplace list] showing which workplace codes the employer has to use for the individual workplaces. The workplace list is based on information from *Told&Skat* on active enterprise code numbers, combined with information on workplaces from the business register.

Employers are asked to update the workplace list to add workplaces which have been set up or delete those which no longer exist, so that the enterprise's corrected list includes those units which, during the information period in question (the year) have had employees linked to them (employment relationships). For each employment relationship, the employer then inserts the relevant workplace code.

In the case of *government employers*, the Workplace Project is not based on the information form. The individual employment relationships and their links with the employer are reported to Danmarks Statistik from the government wage and salary systems, together with information for wage and employment statistics.

The workplace information is collected directly from public employers. A special "workplace list" is sent out, which includes the administrative codes under which the workplaces are registered in the wage and salary systems. The workplace list is used to update the information on the public workplaces in Danmarks Statistik's business register.

The necessary, thorough error-check on the information from the various sources begins as soon as the basic data are available, around four months after the end of the statistical year. Errors are corrected.

This process includes:

- checking and correcting for changes in the data structure in the donor registers;
- a computerised preliminary check on the validity of the information reported;
- a probability check, *inter alia* by comparing figures with previous years' figures;
- contact with some employers for further details and correction of the information reported;
- checks on the information from the statistical registers used, both for validity and to ensure that all the information refers to the same date.

In some cases where firms have more than one workplace, there is no information as to which workplace the individual employment relationship should be linked to. If it is not possible to collect information on this point via the error correction process described above, the employment relationship is automatically allocated to the most probable workplace, on the basis of information such as the size of the workplaces in question and their distance from the municipality of residence.

A workplace is defined as part of a firm which is strictly delimited in terms of organisation, located at a given address and produces one - or one main - type of good or service. It corresponds to the producer units (local kind-of-activity units) in the national accounts.

The number of workplaces at the end of November is taken to be the number in which there is employment at that time. A workplace may have one or more paid employees or a self-employed person.

In certain cases, the employees do not work at any of the employers' registered workplaces but work from home or near their home or at more than one workplace. This happens with cleaning, for example, or with insurance or social or healthcare work (childminders, health visitors). In these cases, the employment relationships are assigned to "notional workplaces", which are not, however, included in the estimates of the number of workplaces.

The starting point is a common definition of the number of persons in employment as at the end of November, for both the ERE and the RAS statistics.

A given person may have links with the labour market at the end of November in one or more of the following ways: as a self-employed person, assisting spouse, employee in employment, on leave from employment, unemployed, on bridging benefit *[overgangsydelse]* or receiving early retirement pay*. The gross figure (stock of persons) for employees and the unemployed is divided according to whether they are on the labour market full time or part time.

First of all, the gross stocks are delimited and a decision is taken on the group (stock) to which each individual should belong as at the end of November.

If an individual could be placed in more than one stock, the order of priority for the ERE statistics is as follows:

1. Persons on leave
2. Unemployed
3. Persons receiving early retirement pay, persons on bridging benefit
4. Employers
5. Employees on the labour market full time
6. VAT-payers
7. The self-employed who are insured against unemployment
8. Employees on the labour market part time or where there is no information on whether they are part time or full time
9. Other self-employed persons
10. Assisting spouses.

The two most important employment relationships are then chosen. The ERE statistics estimates at the end of November include the employment relationship which is a person's most important or next most important link with the labour market. The RAS includes only the most important link.

The final ERE statistics at the end of November include 4, 5, 6, 8 and 10 above.

* *Efterløn*, sometimes translated as "post-employment wage". It is paid, under certain conditions, to employees who retire voluntarily (and not for health reasons) between the ages of 60 and 66.

Below is a brief run-through of the reasons why a person is linked to one or more of the relevant populations as at the end of November.

In the case of employees, the first decision is whether or not they were actually in employment at the end of November, but this decision is taken only for employees who meet the following requirements:

- During the course of the year, they must have received *in total wages or salaries corresponding to at least 80 hours* at the guaranteed minimum wage, i.e. DKK 6 302;
- They must *not be registered with Arbejdsformidlingen [Job Centre] as totally unemployed* during the last week of November*;*
- They must be registered as *resident in Denmark* at the end of the year.

In order to investigate whether a person was in employment at the end of November, information from the unemployment register is assessed. If the persons in question have *unemployment insurance*, they are included.

Information from the Salary Information Register is also analysed. If there is *a mark in the date fields* for employment at the end of November, or if the *amount of the ATP contribution* corresponds to the amount for a full-time/full-year employee, the employment relationship is also included.

Finally, a few employment relationships are included which are drawn by lot from those in respect of which there is some doubt about validity at the end of November.

For persons who were employed in more than one place at the end of November, only the *most important (primary)* and *next most important (secondary)* employment relationships are included in the gross population.

However, for a person to be included as an employee in employment in his/her secondary link with the labour market, the requirements are firstly that it is explicitly stated that the person was employed on the reference date in the relationship which is considered as the secondary one, and secondly that in that relationship a wage/salary above the limit of DKK 6 302 was earned. If these conditions are not met, the secondary employment relationship at the end of November is not counted.

The self-employed population is divided into two groups:

- employers, defined as those who own or are joint owners of an enterprise employing at least one person in the gross population of employees in employment. If a firm has more than one branch, the owners and any assisting spouses are assigned to the main department;
- those who pay VAT, defined as persons who in their *primary* link with the labour market are owners or joint owners of an enterprise which had been set up and had not been removed from the register at the end of November, and who had reported for the VAT register in at least one quarter during the year.

Assisting spouses are those who earn the key share of their income during the year as assisting spouses and are married to a person included in the gross population of the self-employed. The definition of assisting spouses is linked solely to cases where income is transferred to the spouse. If there is a wage/salary agreement instead, the spouse is considered to be an employee in employment on a par with any other employees.

In 1994, information on persons on leave or on bridging benefit was included in the ERE statistics for the first time.

Information on persons on leave comes from the quarterly AMFORA [labour market policy measures] statistics. Persons on leave are not included in employment relationships in the ERE.

Persons on bridging benefit are treated in the same way as those receiving early retirement pay and are thus not included as being primarily in employment in the ERE statistics.

This measure should be seen as an expression of the *total volume of labour* performed by the employees in the enterprise, regardless of whether they work full time or part time or whether they have been employed for the whole year or only part of it.

The number in full-time employment is calculated as follows: *the sum of payments to the ATP scheme for all employees* at a given workplace *divided by the annual ATP contribution for a full-time employee*.

Persons under 16 or over 66 years are not included in the ATP scheme, but, on the basis of the amount of their wage or salary, a *notional ATP amount* is calculated and this is added to the workplace's other ATP payments.

The ATP contribution varies according to the length of time worked for a single employer. The contribution rates are divided into four groups:

- under nine hours' work a week - no ATP;
- 9-17 hours' work a week, 1/3 full contribution;
- 18-26 hours' work a week, 2/3 full contribution;
- 27 hours or more per week, full contribution.

This means that those with *under nine hours'* work per week are *not included* in the estimate of the number in full-time employment.

Certain foreign nationals employed on a temporary basis *do not pay* ATP, either. They are therefore not included in the number of full-time employees.

N.B: Changes in the ATP scheme as from 1993 mean that the number in full-time employment in 1993-1995 is not comparable with the figures for previous years. The above hourly limits for the four groups are one of the changes, since the limits prior to 1993 were 10, 20 and 30 hours of work a week.

The end-November estimate covers, as already mentioned, only persons resident in Denmark at the end of the year. This delimitation does not apply to the estimate of full-time employment and wages/salaries.

Total annual wages and salaries are defined at workplace level as the sum of employees' A-incomes, from which labour market contributions have to be paid, and total contributions, if any, to capital pension schemes administered by employers.

A-income comes from heading 13 of the salary information form, and covers:

– Wages and salaries, fees and the like. It includes: holiday bonuses, wages or salaries paid during sickness or maternity leave, grants of any kind on which A-tax is payable, fees paid to members of boards, committees, etc. and payments of pensions or the equivalent to former employees.

All the amounts under this heading are gross, i.e. they include the labour market contribution. Long-service gratuities, severance pay and contributions to pension schemes (including ATP) are not included.

Information on the total contributions made by individuals to employer-administered capital pension schemes is derived separately from *Told&Skat*'s central pensions system.

Information on employment status is linked to each person. This is taken to be a reliable description of their primary employment relationship. If there is secondary employment as an employee, however, the status information is considered to be less reliable and these employment relationships are therefore recorded separately without any further details.

The figures on which the status information is based have changed over the past few years, *inter alia* as a result of the introduction of information on type of work in the statistics on wages and salaries which are based on Danmarks Statistik's new classification of occupations, DISCO-88. It is no longer possible to continue using the previous classification, which was last used in this edition of the ERE statistics.

At the same time, Danmarks Statistik's new socio-economic classification is introduced into the ERE statistics. This new classification, which will also be used for all person-related statistics during 1997, divides the population up on the basis of their link with the labour market and corresponds to the previous classification where most groups are concerned, i.e. there is a breakdown into self-employed, assisting spouses, employees (wage/salary earners), unemployed, students in further education, pensioners, etc.

The breakdown of employees is totally new, however, since the previous breakdown has been replaced by one based on skill level, which in turn is based on DISCO-88. The level of skill in question can be achieved via formal education and through training and practice.

Like the former information on status, the new socio-economic status is considered to describe primary employment relationships accurately. In the present article, primary employees are divided into the following groups:

- Senior managers in enterprises, organisations and the public sector;
- Employees in a job which requires skills at the highest level;
- Employees in a job which requires skills at the intermediate level;
- Employees in a job which requires skills at the basic level;
- Other employees.

Group 1 covers management work at the highest administrative level in enterprises and organisations - regardless of their size and of the level of skills required for the job. The work consists in making decisions, planning, managing, etc.

Group 2 covers work requiring the highest level of skills, and here persons are classified by job title, such as physicist, actuary, construction engineer, architect, doctor, midwife, teacher, lawyer, auditor, librarian, journalist, actor, musician or priest.

Group 3 covers work demanding an intermediate level of skills. In consists in technical work in production and transport and assistance with trade and administration, educational and care work in day centres and residential institutions, nursing and work as an assistant in the health sector. Examples might be laboratory assistant, programmer, photographer, master on a ship, nurse, kindergarten teacher, estate agent, customs or police officer.

The common requirement for work in *group 4* is a basic level of skills. This group covers general office work, customer service, personal services, security and rescue work, work in agriculture, horticulture etc., craft work and work which consists in operating and/or monitoring plant and other stationary machinery.

Group 5 is a remainder group, covering work not included in the other groups. An example of tasks in this main group might be cleaning work, messenger services, guard duties and packaging and transport work not involving the use of machinery.

Information on the socio-economic status of persons employed by central government, counties and municipalities comes from the register of wages/salaries and staffing statistics for public employees, and for seamen from *Dansk Internationalt Skibsregister*. For other employees, the information comes either from statistics on earnings in the private sector, from a conversion of information on unemployment funds in the register of unemployment statistics or a conversion of the information on the education of the persons in question in the educational classification module.

Information on the highest level of education completed is taken from the basic figures in educational statistics and more detailed information on education can be found in the ERE basic data. It can be supplied via special data runs.

This information is grouped into various major categories:

No occupational training covers general basic education and basic vocational education such as the EFG [basic vocational education] basic year, courses leading to commercial examinations and basic courses in agricultural colleges and colleges for merchant seamen.

If there is no information on education, the persons concerned are allocated to this category.

Vocational training covers completed EFG education, apprenticeship training and basic education leading to examinations in, for example, commercial colleges, technical colleges, agricultural colleges, domestic science colleges.

Short-term further education covers, *inter alia*, teachers in kindergartens/youth workers, child welfare workers, basic training in languages for business, further education leading to examinations in technical colleges etc., training in hospital laboratory work etc.

Medium-length further education covers, *inter alia*, training for primary school teachers, domestic science teachers, journalists, librarians, welfare officers, the HA, HD, ED, and EA commercial/language examinations, technical engineers, *Danmarks Ingeniørakademi*, midwives, physio/occupational therapists and nursing training.

Finally, *long-term further education* covers separately all university education, bachelor degrees from schools of business education, *Landbohøjskolen* [Danish Veterinary and Agricultural College], dental colleges, *Danmarks Tekniske Højskole*, colleges of architecture, colleges of music, officer schools etc.

Information on ownership gives the legal ownership of the firm to which the workplace belongs.

Since 1994, when a new classification was introduced, it has been possible to draw a dividing line between the sectors Corporations and Households plus NPISHs on the one hand and the general government sector on the other, in accordance with the international agreements laid down in the SNA (System of National Accounts).

General government, which consists of central, county and municipal authorities and social funds, is defined as those institutions which produce public services. Public services are understood to be those which are not sold on the open market but are made available to the general public either free of charge or against payment which does not depend on costs.

Publicly owned corporations and public quasi-corporations producing and selling goods and services at market prices are not included in the general government sector . If they are grouped together with general government, the result is the *public sector* , meaning the total area over which government has active control and influence.

The delimitations are based on a classification of workplace units in Danmarks Statistik's business register. This means that the breakdown into sub-sectors within general government is not exactly the same as the delimitation used for statistics on public finance, which is based on a classification of expenditure items in government accounts.

Branch allocation is decided on the basis of the branch code in the business register which applies to the specific workplace at which the employee or the self-employed person is employed. This applies to both private and public workplaces. In the industry grouping annex, the breakdown by branch is defined by reference to the six-digit branch number in Danmarks Statistik's *Dansk Branchekode 1993 (DB93).*

The activities of shipping companies connected with long voyages are allocated in the geographical breakdown to the "abroad" category.

The age of the persons in employment is calculated as at 30 November 1995. The basic data include detailed information on age which can be supplied by special data runs.

There are no periodisation problems with use of the ERE. Incomes are reported on a calendar year basis, since this is the basis with which the income tax system works.

4.7.2 Methods of calculation

Since the ERE statistics cover all producer units in the economy, the statistics do not have to be grossed up for national accounts purposes. However, they do not cover all components of the compensation of employees as defined in the ESA, and various conceptual corrections therefore have to be made. Firstly, imputed employer contributions to social security schemes are not included, and therefore have to be added in. Secondly, not all types of income in kind have to be reported on the information forms. Consequently, the information on the forms is in this case replaced by the national accounts value of fringe benefits etc. Thirdly, employer premiums for industrial accident insurance are not included in the income in the form of wages and salaries which is reported to the tax authorities, but have to be calculated and added in from other sources. Finally, for obvious reasons, the information forms do not include wages and salaries from hidden activity, which therefore have to be calculated separately and added in.

For government non-market activity, the ERE-based calculation is replaced by the compensation of employees as calculated from government accounts, as part of the calculation of the output value of general government. This is to ensure that non-market activity remains consistent within the system. The same applies to industries where public corporations traditionally predominate and to agriculture, construction and a few other industries.

For the calculation of the compensation of employees, the source is the ERE statistics and not tax statistics based on income tax returns (self-assessment), simply because, although both are fiscal sources connected with the income tax system, the ERE statistics based on the information forms are very much more detailed than the published tax statistics. Although it is in principle conceivable that households would report income from hidden activity which, in contravention of the law, is not reported by employers, this is considered an unrealistic possibility. The ERE statistics must therefore be considered to be much the best source for the calculation of the compensation of employees from the point of view of employees.

Total annual wages and salaries are defined at workplace (local kind-of-activity unit) level as the sum of employees' A-income, from which labour market contributions have to be paid, and any overall contributions to capital pension schemes administered by employers.

A-income comes from heading 13 of the salary information form. This field covers:

- wages and salaries, fees and the like. It includes: holiday bonuses, wages or salaries paid during sickness or maternity leave, grants of any kind on which A-tax is payable, fees paid to members of boards, committees, etc. and payments which are equivalent to pensions to former employees. (A-tax is PAYE (pay-as-you-earn) tax, i.e. a tax which the employer is obliged to withhold).

All the amounts under this heading are gross, i.e. they include the labour market contribution deducted at source. Contributions to social security schemes (ATP, the labour market supplementary pension scheme), long-service gratuities and severance pay plus contributions to pension schemes are not included. With one exception, these are recorded in other fields on the information form. The exception is total contributions to capital pension schemes administered by employers, which are different from other pension contributions in that they are not normally a feature of a contractual relationship and no information is given on the form. It is instead obtained separately from *Told&Skat*'s Central Pensions System, which is based on mandatory returns from financial institutions. As already stated, contributions to employer-administered capital pension schemes are added to the values in field 13 when ERE total wages and salaries are calculated, whereas contributions to other pension schemes and to social security schemes (ATP) are not included in total wages and salaries as reported in the ERE statistics.

The ERE statistics do not include income in field 36 "B-income, from which labour market contributions have to be paid" in their calculation of total wages and salaries. We cannot rule out the possibility that certain income in the form of fees, which should be treated in the national accounts as compensation of employees, is posted under this heading and is therefore not included in the calculation of total wages and salaries in the register statistics. The national accounts implicitly correct for this in all cases which are GNI-relevant by basing the estimate of total wages and salaries directly on exhaustive accounting information (non-market output).

In the national accounts, the compensation of employees is divided up over 130 industries. The source for the vast majority of these industries is the ERE statistics, but for some of them, in particular agriculture etc. and all government non-market activity, accounting information is used

instead of the register information. For agriculture etc. and general government, the compensation of employees was already calculated in national accounts terms for the production of primary statistics.

The connection between total wages and salaries in the ERE statistics and the compensation of employees in the national accounts is illustrated in the table below. The item "employer contributions to pension schemes etc." includes both actual and imputed pension contributions. In Denmark, the imputed contributions refer only to civil service pensions.

Table 102 Connection between the ERE and the national accounts

	Wage/salary component	**DKK million**
	Compensation of employees in the national accounts	534 094
	Total wages and salaries in the ERE statistics	490 966
	of which employer-administered capital pensions	*7 209*
-	Tax value of fringe benefits	2 862
+	Long-service gratuities and severance pay	1 905
+	Gifts in kind	12
+	Wages from hidden activity (1 660 + 1 235)	2 895
+	National accounts calculation of fringe benefits	6 788
+	Employer contributions to the ATP	4 906
+	Employer contributions to pension schemes etc.	26 284
+	Industrial accident insurance	1 631
=	Corrected total wages and salaries for ERE statistics	532 525
	Percentage discrepancy	0.29%

The national accounts figures for the compensation of employees are higher than the ERE figures after the latter have been corrected for conceptual differences because the ERE figures for a few - but important - industries are replaced by accounting information. The main such case is general government, which accounts for almost one-third of total wages and salaries in the economy. It appears likely that the higher values in the general government accounts than in the conceptually corrected ERE statistics are due primarily to the treatment of certain fees etc. in the ERE statistics. Presumably, in the case general government entities, these fees are not in all cases included on the information forms under point 13 because legally speaking there is not an employment relationship. Instead, they come under taxable B-income. Broadly speaking, this problem is assumed to apply only to general government, where the remuneration for certain assignments such as lecturing, involving only few hours' work, consists of non-invoiced fees which are considered to be wages/salaries.

When equivalent services are supplied to market producers, there will in most cases be an invoiced sale of services, and so the same problem does not arise.

For GNI purposes, the crucial point is that the general government accounts should include all the compensation of employees. In the national accounts, it is the public accounts values for the compensation of employees, which are higher than the ERE values, which are included in the calculation of output and value added for general government.

The initial income-based GDP estimate for 1995 is as follows, where MTM is the target total module derived from the intermediate system based on business accounts, as illustrated previously in Tables 4 and 7:

Table 103 Initial estimates of income-based GDP

		DKK million
	Initial estimate of gross value added from the MTM	868 298
-	Compensation of employees according to the MTM	527 678
-	Wages in the hidden economy other than for private households with paid employees	1 660
-	Long-service gratuities and severance pay	1 905
-	Other taxes less subsidies on production	
=	Initial estimates for gross operating surplus from the MTM	337 816
+	Compensation of employees calculated from the employees' point of view	534 094
+	Taxes on production and imports	173 270
-	Subsidies	35 618
=	Initial estimates of income-based GDP	1 009 562

The initial estimate of GDP via the income approach is corrected for "Wages in the hidden economy other than for private households with paid employees" and "Long-service gratuities and severance pay" because these amounts are not coded as the compensation of employees in business accounts or in the target total module MTM. In the MTM, wages in the hidden economy are implicitly coded as gross operating surplus, with the wages share of hidden activity calculated afterwards in connection with the calculations of employment based on the ERE statistics. The exception is industry 950000, Private households with paid employees, where there is a national accounts calculation and where the output value in the MTM is, of course, coded as compensation of employees. Long-service gratuities will normally be counted by the businesses concerned as wages/salaries whilst severance pay is assumed to count as extraordinary expenditure. Since by far the largest share of the total amount for long-service gratuities and severance pay as posted in the salary information register must be assumed to relate to severance pay connected with redundancies, it was decided to make a correction equal to the total amount for long-service gratuities and severance pay the value of which comes from tax figures.

The income-based GDP calculation of the compensation of employees must be considered to be as independent of the other two methods of estimation as it is possible to make it. When compared with the output-based estimate, it is an extremely robust check, since total wages and salaries in the intermediate system and the MTM which are the wage/salary counterpart to output-based GDP are calculated from processed and grossed-up business accounts, whereas the income-based estimate is calculated from a *total census* of the *individual employees' income in the form of wages and salaries* as reported to the tax authorities. One could hardly wish for a statistically more robust comparison of data on wages and salaries.

It is obvious that the calculation of the other two components of income-based GDP, namely gross operating surplus and mixed income and taxes on production and imports less subsidies, cannot, of course, be independent of the output-based calculation, since the figures rely on the same business accounts/ government accounts.

It is only with a more primitive statistical system, where there are no calculations of integrated accounts by industry and sector based on exhaustive primary accounting statistics, that it might conceivably be possible to produce independent output- and income-based estimates of that share of GDP which corresponds to these primary income components. Obviously, if the output-based estimate is constructed to any noticeable extent on a foundation other than business accounts, using indirect methods of estimation and projections from a base year, for example, there might be grounds for claiming that the output- or income-based calculations of GDP may be more or less independent as regards the gross operating surplus and mixed income share. In a more highly-developed statistical system, this independence goes by the board.

With the compensation of employees accounting for some 53% of GDP in 1995, the output and income approaches may be said to be independent only as regards this percentage at most. Since the consistency required of the costs-based estimate of the value of government non-market output, non-market output in NPISHs and non-market output relating to employees in private households by definition rules out any independent estimates in this field, and non-market activity accounts for a good third of total wages and salaries, only around 35% of GDP can genuinely be said to be estimated independently.

The income- and expenditure-based calculations may be considered to be totally independent of one another apart from non-market activity and imputed transactions or transactions calculated by convention (the rental value of owner-occupied housing, fringe benefits, etc.), where by definition they give the same result. Since these components together account for around 27% of GDP, the two calculations (income and expenditure) of the other 73% or so can be said to be truly independent.

4.8 Other taxes on production

Table 104 summarises other taxes on production in the national accounts for 1995.

Table 104 Other taxes on production, 1995

Type of tax	DKK million
Employer contributions to *Arbejdsgivernes Elevrefusion* (AER)	2 262
Road fund licence on vehicles used in production	840
Property taxes	10 325
Payroll taxes	2 502
Taxes linked with checking, supervision, licences, etc.	86
Other taxes on production, total	16 015

It shows that there are only a few types of tax which are classified as other taxes on production. The AER contribution, which finances apprenticeships and traineeships, does not give the individual employer or employee any rights and is therefore a tax and not, for example, a contribution to social security schemes. That share of total vehicle registration taxes which relates to vehicles used in production is calculated from a breakdown by owner of the total number of vehicles registered. In the national accounts, road fund licences on consumers' vehicles are "direct taxes", i.e. taxes on income and wealth etc. Obviously, property taxes are not linked to products. Payroll taxes are, as their name indicates, a tax on the wages and salaries paid by financial institutions, to offset the fact that most financial services are exempt from VAT. It can be seen that only taxes for checking etc. amounting to DKK 86 million are counted as taxes under paragraph 4.23 e) in the ESA 95. All other payments by

producers in connection with government checks and licences are considered to be purchases of services.

All the above taxes are clearly other taxes on production. There is no borderline case of any importance in quantity terms. All the taxes are purely national and not EU schemes. The total tax revenue is assigned to the domestic general government sector.

Other taxes on production are divided up over the national accounts' 130 industries as follows:

Most of the other taxes on production consist of *property taxes* (approximately 65%). On the basis of the property type "dwellings and agricultural property", a good 60% of property taxes can be divided between the national accounts industries "agriculture" and "dwellings". The remaining property taxes, which come from taxation on non-residential property, are divided up using a specific branch key.

Payroll taxes are paid, as stated above, by various enterprises which are not liable for VAT. The Payroll Tax Act (Consolidated Act No 666, 1994) and the VAT Act (Consolidated Act No 375, 1994) govern which national accounts branches pay the tax. The breakdown of the tax over the branches paying it is based on the total wage/salary bill.

The *AER contribution* is paid for each full-time employee in the enterprise and for this reason the national accounts industry breakdown is based on the national accounts distribution of the number of employees.

The motor vehicle registration tax on vehicles used in production is divided up in a special calculation system. In a sub-system known as the "vehicle distribution system", the road fund licence amounts are divided up by type of vehicle and national accounts industry using the register of motor vehicles as the source.

Taxes on *checking and supervision* comprise the:

tax on payment cards;
tax to *Arbejdsmiljøfonden* [Work Environment Fund].

These very small amounts are divided up using specific branch keys.

Taxes connected with *licences* include:

the tax on pharmacies;
taxes under the cultural fund, and the
tax for the operation of the school ship "*Danmark*".

These very small amounts are broken down on the basis of specific branch keys.

4.9 Other subsidies on production

Subsidies on production which are not linked to products come under both EU and national schemes. Table 105 summarises these other subsidies:

Table 105 Other subsidies on production, 1995

Type of subsidy	DKK million
Total EU schemes	1 650
Interest rate subsidies	293
Aid per hectare and set-aside	730
Other EU schemes	627
Danish schemes, total	15 128
Subsidies to pharmacies	40
Interest subsidies and contributions, housing	4 234
Municipal subsidies to private sports halls	359
Municipal subsidies to theatres, orchestras, cinemas, etc.	381
Central government subsidies to regional orchestras	81
Danmarks Erhvervsfond [Trade and Industry Fund], export-promoting arrangements	112
Danmarks Erhvervsfond, losses on guarantee and surety schemes	305
Subsidies for product development	264
Start-up grants replacing offers of employment	822
Retraining grants	1
Arbejdsgivernes Elevrefusion (AER)	2 175
Compensation for employers' contributions to the ATP	1 486
AMBI reduction	588
Municipal grants for the employment of the unemployed	349
Expenditure under the County Land Tax Act	154
Interest and interest rescheduling subsidies to agriculture	8
Central government subsidies to private railways	84
Municipal subsidies for the running of buses and other transport activities	1 528
Home helps	259
Other subsidies on production to private enterprises	1 896
Other subsidies on production, total	16 777

The only borderline case is the subsidy referred to as "AMBI reduction". This covers repayments to producers of a tax known as the AMBI, which was unlawfully collected. The amounts paid back are not to be offset against taxes in the years in which they were originally collected but are considered to be subsidies in the year in which the repayment is made. The economic reasoning behind this is that the behaviour of the economic decision-makers was not affected by the repayment in the year in which the tax was unlawfully collected, since it was not until later, under a court ruling, that the tax was declared unlawful.

The subsidy known as "*Arbejdsgivernes Elevrefusion*" is the counterpart to the other tax on production known as the AER, which was shown in Table 104. All employers contribute to a pool

which finances apprenticeship and trainee places in connection with vocational training. Those employers who employ apprentices and trainees receive a subsidy from the pool.

The allocation by industry of other taxes on production is a two-stage process. It begins with the most detailed information on general government from *Databasen for Integrerede Offentlige Regnskaber (DIOR)* and first of all assigns a wide range of specific subsidies *directly* to industries. Around 60% of all other subsidies on production are allocated to an industry in this way.

Next, the *remaining* subsidies are classified into eight different categories, each of them broken down using a specific industry key.

4.10 Gross operating surplus and mixed income

4.10.1 Statistical sources

Since the income component "gross operating surplus and mixed income" is estimated in a fully-fledged calculation system as an integrated estimate of output and operating surplus based on business accounts, the sources and methods in this case will, of course, be exactly the same as for the output-based estimate. For further details, therefore, reference should be made to Chapter 3, in particular sections 3.1.2, 3.2 and 3.3.

For this component, an alternative source might be the income tax returns for corporations and the self-employed used for corporation tax/personal income tax. Since the accounting information in the income tax returns is much briefer and much further from national accounts concepts than the accounts underlying the intermediate system and the target total module, these are obviously a much less satisfactory statistical source than the accounting statistics described in Sections 3.1.2, 3.2 and 3.3.

4.10.2 Methods of calculation

For the GDP component "gross operating surplus and mixed income", the statistical sources are the same as for the production approach and the initial estimate for this component is derived from the functional target total module as follows:

output (1015) – intermediate consumption (2010) – other taxes on production (3110) + other subsidies on production (3210) – compensation of employees (4010).

The periodisation of the accounting statistics used for the functional target total module was described in Section 3.1.2.3.6

The final balanced value of gross operating surplus and mixed income is obtained when the initial estimates for output and intermediate consumption are replaced by the corresponding figures in the balanced supply and use tables, and corresponding initial estimates for other taxes and subsidies on production plus the compensation of employees are replaced by the final balanced values in the national accounts industry tables.

4.11 Mixed income

The initial estimate of gross operating surplus and mixed income prior to balancing is calculated as an aggregate in the target total module (MTM) on the basis of information in the intermediate system, as described in Section 4.10.2. As a result of the consequent double coding of the accounting information by both industry and institutional sector, that share of the "gross surplus" which relates to producer units belonging to institutional units in the household sector (S.14) can be obtained directly from the target total module. The "gross surplus" in the household sector is referred to in the ESA 95 as "mixed income", since it is a mixture of return on capital and a reward for labour in the form of the labour input of the self-employed in their own enterprises.

4.12 Consumption of fixed capital

In general, the estimate of the consumption of fixed capital is not relevant to GDP or GNI, since these concepts are, of course, *gross*, i.e. production or income aggregates before deduction of the fixed capital consumed.

There is, however, one very important exception to this main rule, namely non-market activity, where by convention output is calculated from the costs point of view, and where the consumption of fixed capital is one of the components of costs. Non-market activity occurs in Sector S.13, general government and Sector S.15, non-profit institutions serving households. The latter is private non-market output. The vast majority of non-market output comes from government.

As regards the minor share of non-market output in S.15, the consumption of fixed capital is, as stated in Section 3.1.2.4.3, calculated at 49.4% of total wages and salaries. This percentage is based on an estimate of capital stock in the sector carried out in 1995, where the latest final figures referred to 1992. This capital stock estimate consisted of a mixture of direct estimates of stocks and PIM (perpetual inventory method) calculations. Since the link between the consumption of fixed capital and total wages and salaries may be assumed to be relatively stable in this field, it was decided to project the 1992 total wages and salaries benchmark in the current calculations of this relatively modest amount. A new benchmark is now available for years up to 2000 and the percentage figure in the calculation will be revised in 2002, in the light of this new information.

The description below therefore refers solely to the consumption of fixed capital in S.13, general government. For 1995, it was obtained by projecting from an extremely detailed and reliable calculation of the level for 1992. Exceptionally, 1995 was not calculated directly as a level, owing to the necessary time lag between publication in 1997 of the new ESA 95 national accounts for years from 1988 onwards and the point (year 2000) at which it became possible to calculate estimates of the stock of fixed capital which were consistent with those figures. This process was time-consuming because all the capital formation series in the new national accounts first had to be extrapolated back to 1966 and then at a more summary level even further back, so that they could form the basis for calculations of the stock of fixed capital goods using long capital formation series and the PIM. These new capital stock calculations were first published on 9 February 2001. For 1992, the figures for fixed capital consumption match the previous figures almost exactly at macro level, which indicates that the 1992 levels used for the calculation of S.13 for 1995 are essentially correct. In future (as from the final figures for 1999), the consumption of fixed capital in S.13 will be calculated directly as a level each year in connection with the very detailed estimates of capital stock in the general government sector.

General government capital stock consists of buildings, structures such as roads, bridges etc., machinery, transport equipment and intangible fixed assets, which for this sector is in practice software. Buildings constitute by far the largest share of government capital stock and capital formation. For 1992, the consumption of fixed capital was obtained - as required by the ESA 95 - via an estimate of the gross stock of the individual types of capital and use of the straight line depreciation method. Whereas the ESA 95 is to a certain extent flexible, the ESA 79 demanded the linear method. One important strong point in Denmark's estimate is that for buildings and transport equipment the calculations could be based on a *direct estimate of stocks* which in turn was based on register information for a benchmark year - in this case 1988 - for buildings, and every year for transport equipment. In contrast to PIM calculations, there is therefore absolutely no uncertainty as to how many square metres of buildings there actually were in S.13 in 1988. The only uncertainty concerns their lifetimes. For buildings, the PIM was used solely to project the 1988 benchmark to 1992. The following table shows the methods used for each type of capital formation:

Table 106 Methods of estimating capital stock in S.13

Type	Method	Survival curve	Assumed average lifetime	Number of products
Machinery	PIM 1947→	Winfrey S3	Varying	Approx. 350
Transport equip.	Direct estimate of stocks	Not relevant	Varying – 13 years for cars	6
Buildings	Direct estimate of stocks for 1998 projected to 1992 using the PIM		Constructed 1960 onwards, 65 years Constructed prior to 1960, gradually increasing lifetimes back in time	2
Roads and bridges	PIM 1850 →	Winfrey L3	40 years	1
Software	PIM		3 years	1

The level thus obtained for general government fixed capital consumption in 1992 is projected to 1995 at aggregate level, the only distinction being between software on the one hand and all other capital stock on the other. The projection to 1995 for all capital stock other than software is obtained by inflating the previous year's consumption of fixed capital to the current year's price level using the implicit deflator for all gross capital formation and adding 1/25 of the current year's gross fixed capital formation, corresponding to an assumed average lifetime of 25 years. For software, the 1995 calculations reckoned on a fixed lifetime of three years. In the new capital stock calculations which will be included in the national accounts as from 1999, a Winfrey S3 survival curve is used, and software is divided into two types, namely purchased and own-produced. For purchased software, the lifetime will now be taken to be four years and for own-produced six years.

The above procedure may be considered to be rather a summary one, but one must bear in mind that it has been used for a very short period of only six years (1993-1998) on the basis of what may be

considered in international terms to be an extremely solid benchmark for 1992. There is no doubt that, compared with other countries' calculations, the 1995 one is very reliable. As already stated, the projection will cease as from the final national accounts for 1999, which will be based on a direct estimate of levels every year.

The following table shows the consumption of fixed capital in general government, 1992-1999, calculated according to the ESA 95. Differences in definition from one system to the other are also shown.

Table 107 Consumption of fixed capital in S.13, ESA 95 and ESA 79

DKK million	1992	1993	1994	1995	1996	1997	1998	1999
Consumption of fixed capital, ESA 95	21 887	22 775	24 514	24 706	25 280	25 796	25 939	26 159
- Roads, bridges, etc.	6 567	6 963	7 500	7 447	7 546	6 897	6 821	6 771
- Leasing	104	65	106	101	97	103	106	110
- Software	632	911	1 100	1 303	1 574	1 687	1 759	1 782
- Military capital goods	520	548	571	588	611	629	646	664
+ Threshold value, consumables	1 066	1 145	1 230	1 321	1 420	1 510	1 581	1 652
Consumption of fixed capital, ESA 79	15 130	15 433	16 467	16 588	16 872	17 990	18 188	18 484