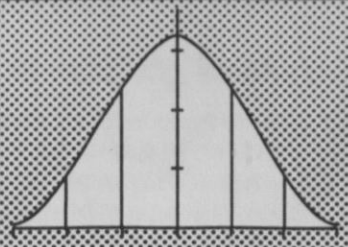
Approaches Methods Developments



Federal Statistical Office

Information of the German Federal Statistical Office

Number 1/98

Contents

Federal Statistical Office, Gustav-Stresemann-Ring 11, D-65189 Wiesbaden

Published by:
Federal Statistical Office, Wiesbaden

Information:
Statistisches Bundesamt
General Information Service
D-65180 Wiesbaden
- Phone: + 49 - 611 / 75 24 05
- Fax: + 49 - 611 / 75 33 30
- E-Mail: auskunftsdienst@stba.bund400.de
- Internet: http://www.statistik-bund.de

Further information on this publication may be obtained from:
Division I B
Phone: + 49 - 611 / 75 20 77

Published in May 1998

Printed on 100 % recycled paper.

Methods of federal statistics – Further development

Setting up new groups of reporting enterprises for trade statistics

The reporting system of official trade statistics is based on monthly, yearly, and multi-annual sample surveys in the areas of wholesale and retail trade (as a result of the 3rd law on the adjustment of statistics, enterprises operating in the field of commission trade are no longer surveyed). The sampling frame determining the sample, i.e. the reporting enterprises – which are obliged to provide information – is the Census of Distributive Trade and the Hotel and Restaurant Industry (*Handels- und Gaststättenzählung – HGZ*), which generally is conducted every ten years as a complete enumeration. The latest enumeration conducted in 1993 provided reliable data material for the first time also for the new Länder and Berlin-East, which permitted to develop a uniform sampling plan for Germany. It was designed as a one-stage stratified random sample. The objective was to obtain federal results with sufficient accuracy for the levels of classes (four-digit items) and selected subclasses (five-digit items) according to the national Classification of Economic Activities, 1993 edition (*Klassifikation der Wirtschaftszweige, Ausgabe 1993 – WZ 1993*); another intention was to obtain reliable results for the Länder at least at group level (three-digit items). The most important characteristic to be shown is turnover, which is considered as an economic indicator of the development of final consumption of households.

In the past, sample surveys in the field of wholesale trade were based on a cut-off threshold of DM 1 mn (in retail trade: DM 250,000). Enterprises below that threshold were not covered, i.e. they were not represented by the sample. This involved two shortcomings: First, data on this area, which is important at least in terms of number of enterprises and employees, were collected only at large intervals as part of the surveys mentioned before. Second – and this is a serious methodological shortcoming –, enterprises whose turnovers had exceeded that threshold in the course of time could not be recognized and, consequently, could not be included in the sample. Therefore, cut-off thresholds were no longer used for the new sampling of reporting enterprises which had to provide data for the yearly and multi-annual statistics. This approach is in line with the provisions of the EC regulation on structural business statistics.

For the monthly statistics, however, the former cut-off thresholds were doubled (to DM 2 mn in wholesale trade and DM 500,000 in retail trade). This permitted to reduce considerably the sample size defined in the sample design (cf. table below), without increasing the random sampling errors. Another advantage of abandoning the cut-off thresholds for the other surveys, as mentioned before, is that the above-described undercoverage involved in the application of cut-off thresholds can no longer occur. The reason is that enterprises exceeding the threshold in the course of time can be integrated from the groups of enterprises reporting for the annual survey into the subset of enterprises reporting monthly.

Sample surveys in distributive trade
Sample sizes and sampling fractions according to the sample design

Area	Sample size		Sampling fraction in %[1)]	
	yearly survey	monthly survey	yearly survey	monthly survey
Wholesale trade	12 780	9 332	10.4	20.3
Retail trade	32 260	22 597	7.7	14.5

1) Referring to the sampling population concerned.

The loss of information caused by cut-off thresholds is very small. Although, according to the results of the 1993 *HGZ*, the number of wholesale and retail enterprises that were below the cut-off limit accounted for over 60 %, their share in total turnover was just around 5 %. Thus small enterprises have very little influence on the relatives and rates of change for turnover and number of employees shown in the monthly statistics.

The phase of conversion to the new groups of reporting enterprises has been completed, so that up to date monthly relatives are available again.

The breakdown of results is based on the *WZ 1993*. This means in particular that the representations of wholesale and retail trade exclude the field of "sale of motor vehicles, retail sale of automotive fuel", which is shown separately according to the classification of economic activities. However, in order to represent the results according to the former delimitation, which the users are acquainted with, the two areas of trade are additionally shown including the wholesale and retail branches of the sphere of "sale of motor vehicles, retail sale of automotive fuel".

Albrecht Krockow, tel. (+ 49 - 611) 75 - 20 60

Methodology of the delimitation, breakdown, and determination of education expenditures in Germany

The Federal Minister of Education, Science, Research and Technology commissioned the Federal Statistical Office to develop a comprehensive concept for the coverage and representation of education expenditures in Germany, which would allow to fill the gaps in current reporting, for instance, to international organizations. The starting point for the project was the study "Education at a Glance" performed since 1996 jointly by the OECD, UNESCO, and the Statistical Office of the European Communities (Eurostat). The project covers all educational areas (e.g. general education, vocational training, and the funding of education and training). The study includes both public and private-sector recipients and providers of educational measures.

The definition of the educational field and its delimitation against related areas (e.g. culture, research, science) is based on the international classifications ISCED (International Standard Classification of Education) and NACE (Statistical Classification of Economic Activities in the European Community). Covering the educational sector in monetary terms is based on the systematization of payment flows – which are relevant to education – between and within transactors that are relevant to education.

What is used as a basis for determining education expenditures of the public sector are data from the annual statistics of expenditure and revenue of public budgets. However, such data have to be adjusted for many non-education components (e.g. treatment of university hospital patients) and – to a much larger extent – supplemented by individual education-relevant aspects from other areas (e.g. civil service education and training, transfer payments of the Federal Institute for Employment, allowances paid by the government to public employees, fictitious pension payments for officials currently employed). What is much more difficult is determining the expenditures of private-sector educational institutions and education expenditures of households. Since finance-statistical data are not directly available – except for private institutions of higher education –, it is necessary to use a number of other official and non-official statistics and to perform appropriate estimations. Data gaps both in the private educational sector (e.g. private schools and private nursery schools) and in the field of public education supply (e.g. "in-company" training of apprentices at central, regional and local authorities as part of the dual vocational training system) are filled by means of special surveys. In order to permit comparison and linkage of the finance data of private education providers (commercial accounting) with those of public education providers (governmental accounting) – which exist partly parallel and partly complementarily to each other –, a specific delimitation of expenditures was established. Expenditures are represented both according to the "expenditures by implementing institution" concept and the "expenditures by source" concept. For better comparability of sectorial data, a distinction is made – similar to the UNESCO/OECD/Eurostat nomenclature – between expenditures for the educational process (first degree education expenditures) and expenditures for subsistence (second degree education expenditures).

In particular, in the course of the work, methods are developed for allocating education expenditures to the individual levels of education and for an improved regional comparability of educational finance statistics of the German Länder. Also, results of a special survey among private schools are used to develop an adequate method for the determination of expenditures made there and for updating such data.

The progress of the project is regularly presented and discussed in the working group on "educational finance" of the Federation-Länder Commission for Educational Planning and Research Promotion; members of that working group are representatives of the commissioning ministry, the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany, the Federal Ministry of Finance, the business and science sectors, the municipal central associations, and the social partners. The project will be terminated on 31 March 2000.

Patrick Lünnemann, tel. (+ 49 - 611) 75 - 41 46

Improvement of the concept of the environmental-economic valuation approach "abatement costs" and first steps of empirical implementation

The objective of Environmental-Economic Accounting of the Federal Statistical Office is to provide statistical information enabling the spheres of politics, science, and business, as well as other users of statistics to judge in how far our society with its manifold activities affecting the environment is moving towards "sustainable development". This requires not only data on the use of the environment in physical units, such as tonnes, joules, or hectares, but also monetary valuation. A major element here is the calculation of "abatement costs". They indicate the costs that (hypothetically) would arise if specific environmental burdens caused by economic activities were gradually abated by well-aimed measures. Generally, abatement measures include, first, abatement and reduction technologies, second, structural measures in the sense of reducing production processes which have an impact on the environment and, third, behaviour-oriented measures.

Checking and further developing the valuation approach "abatement costs" by the Federal Statistical Office was the empirical core of the research project "Methodological problems in the calculation of environmentally adjusted national income figures" of the European Commission, which has been terminated. In an iterative process, the individual conceptual steps of the approach were checked for their empirical feasibility and, where necessary, adjusted to statistical requirements. For this purpose, abatement cost curves were calculated for emissions of selected nitrogen compounds which have an impact on the environment (NO_x and N_2O into the air). The calculations were based on the emissions and on the reduction potentials actually available in the economic sectors in the former territory of the Federal Republic of Germany in 1990.

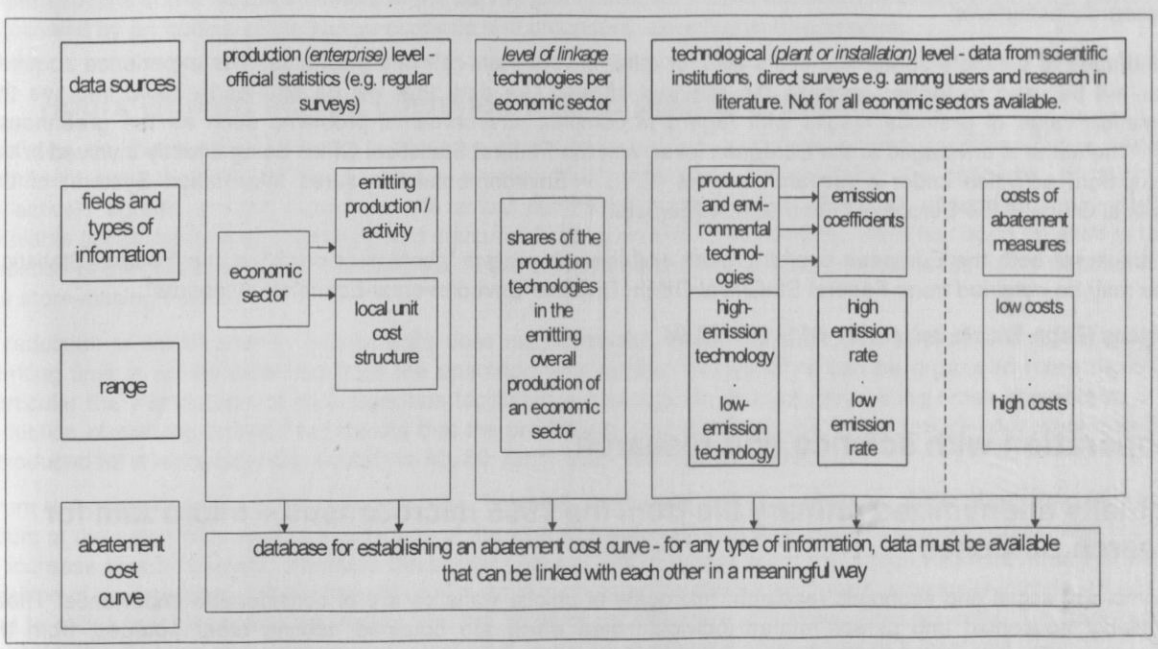
Major restrictions of the study of the Federal Statistical Office were the following:

- The limitation to technological abatement and reduction potentials in the sense of end-of-pipe technologies and integrated environmental protection technologies.
- Only the direct costs of any technological abatement measure were calculated.

Indirect costs – induced by changes in intermediate consumption flows and shifts in demand as a result of introducing a technological measure – and the effects of structural measures can be analyzed only by means of econometric models referring to the overall economy. In Germany, such model calculations should be performed by scientific institutes because this is not part of the task of the German Federal Statistical Office.

Generally, a vast number of data in several fields of information are required to calculate direct abatement costs of technological measures. The following chart shows the main components:

Types of information required and their linkage as part of calculating abatement cost curves



Based on the three main fields of information shown in the chart, three steps of calculation are necessary:

1. Calculating current emissions with a detailed breakdown by production processes generating the emissions and by technologies used, unless the emissions are collected directly in a quantitative manner.
2. Linking the emission-relevant quantitative frame from step 1 with the abatement and reduction technologies currently available. As a result, the hypothetical emissions for any case are obtained.
3. Calculating the costs of the hypothetical reduction of emissions by estimating the costs of applying the individual abatement/reduction technologies.

There is not only the problem of obtaining in a roughly consistent form the required data both at the local unit/enterprise level and the technological level for all activities generating emissions. Empirical work has shown that another crucial problem is the conceptual requirement that data be linked with each other although they were collected according to different classifications and following different motives. For the end-of-pipe technologies, it was possible to link the physical (emission-relevant) information with cost data. However, for the complex of "Integrated Technologies", manifold problems emerged, in particular with regard to determining the environment-

related share of costs. Any data on the cost structure of the various production processes that are available and suited for statistical evaluation refer to the level of local units/enterprises only. As is shown in the chart, however, the emissions are determined and aggregated at the technological level by types of plants or installations (broken down by technological characteristics); this is also true of determining emission abatement potentials. Without further work-intensive conversion, the underlying classifications are not compatible with the statistical classifications by economic activities and sectors of homogeneous production. Consequently, the results available on nitrogen abatement cost curves have so far consisted mainly of data on the application of the various end-of-pipe technologies.

In connection with these empirical activities of the Federal Statistical Office, the European research group has worked out the following conceptual results and recommendations regarding the further development in this area:

- For selected pollutants, environmental problems, and generators, it is possible for official statistics to determine with a sufficient degree of accuracy direct abatement costs of technological measures.
- It is not possible with justifiable efforts to empirically determine complete abatement costs which would provide a monetary estimate of all environmental damages caused in a national economy over a given period. It is recommended instead to interpret abatement costs as partial monetary indicators which help to quantify the current distance from individual objectives of sustainability.
- To obtain abatement costs for the overall economy and to simulate a sustainable national economy, it is necessary to link data on direct abatement costs of technological measures with macroeconomic models and to supplement such data with additional information on structural changes. Naturally, however, the results of model calculations strongly depend on the assumptions made before. This means that normally such model calculations are not really part of the tasks of official statistics and should therefore better be performed by research institutions.

It is intended to estimate abatement costs also for other environment-relevant pollutants. The experience acquired so far will be used to further optimize the required efforts. The data that will be obtained should improve the information value of previous results with regard to complex environmental problems such as the greenhouse effect. The same is envisaged at the European level, with the Federal Statistical Office being directly involved in the development activities under a research project (EPIS – Environmental Pressures Information System) of the Statistical Office of the European Communities (Eurostat).

The results of both the European overall project and the sub-project "abatement costs" of the Federal Statistical Office may be obtained from: Federal Statistical Office, Division "Environmental-Economic Accounts".

Wolfgang Riege-Wcislo, tel. (+ 49 - 611) 75 - 27 94

Cooperation with science and research

Factually anonymized primary file from the 1995 microcensus – microdata for research purposes

For empirical social and economic research, microdata of official statistics are of considerable importance. These are mainly household and person-related individual data which are obtained, among other sources, from the microcensus. The latter is conducted on an annual basis and covers one percent of the households in Germany. It is a multi-purpose sample survey providing statistical data with a detailed material breakdown on labour market trends as well as the economic and social situation of the population.

Since the Federal Statistics Law (*Bundesstatistikgesetz – BStatG*) was amended in 1987, scientific institutions have had a special status. They have not only been allowed to use absolutely anonymized microdata but have also been given the opportunity to obtain "factually" anonymized data. Art. 16, para. 6 of the BStatG stipulates that, for the purpose of scientific projects, the Federal Statistical Office and the statistical offices of the Länder may transfer individual data to institutions of higher education or other institutions entrusted with tasks of independent scientific research if an allocation of the individual data to respondents or persons concerned is possible only with an excessive amount of time, expenses and manpower. Such data may be used only for the purposes for which they were transmitted and they have to be deleted as soon as the scientific project has been terminated (Art. 16, para. 8 of the BStatG).

Producing factually anonymized primary files has proved to be an efficient form of data supply. This means that the data material has to be anonymized once only. With such a file, the Federal Statistical Office has the opportunity to provide rather quickly those variables or sets of variables which serve a concrete research project of individual scientists.

In the microcensus, factually anonymized primary files were established for the years 1989, 1991, 1993, and 1995. The primary file for 1995 comprises 70 % of all data records, i.e. anonymized individual data on about 571,000 persons or about 258,000 households. In addition to detailed socio-demographic and socio-economic information, it also contains data on the health situation as well as on private old-age provision and company old-age pension schemes.

As part of a pilot project funded by the Federal Ministry of Education, Science, Research and Technology and developed by the Federal Statistical Office in cooperation with representatives of the academic community, the factually anonymized primary file of the 1995 microcensus may be obtained against a small fee, which was not possible in previous years. A relevant administrative agreement was concluded between the Ministry and the Federal Statistical Office in December 1996.

For further information on how to obtain that primary file please contact the Federal Statistical Office, Division VIII C (D-65180 Wiesbaden, tel. (+ 49 - 611) 75 - 25 47, fax (+ 49 - 611) 72 40 00).

Sabine Köhler, tel. (+ 49 - 611) 75 - 26 93

Methodology of the survey of production days performed in manufacturing

Preliminary remarks

The survey of the number of working or production days performed in manufacturing between May and July 1994 was conducted as a survey for specific purposes on the basis of Art. 7 of the Federal Statistics Law. The present paper explains some typical concepts of the survey and describes in brief the survey development. This description is followed by an outline of the survey contents and procedure, sampling and estimation.

1 Hours worked versus operation time or machine running time

Statistics on the number of hours worked by employees are part of various official statistics. Those statistics document both the total of hours worked, a value which is required for analysing the economic situation, the collectively agreed working hours and the actual number of hours worked. However, information has not been available about the time of machinery and equipment operation and, consequently, there has been no answer to the question of the operation times of local units. In the services sector, the question concerns the hours of business or the store-closing hours.

A reduction of working hours in local units does not necessarily entail a shorter time of production. The individual working time is rather detached from the operation time so that the workflow can be organised more flexibly. In particular the introduction of shift work has facilitated a prolongation of machine running times irrespective of the reduction of working hours. That means that the production time of a local unit where the 35-hour week has been introduced for employees may amount to 40, 50, 60 or even 100 hours a week.

From the perspective of business economics, however, local units are in a dilemma. If the shortening of working hours at the same time means a reduction of the time of machinery and equipment operation, productivity will have to increase proportionately, otherwise the capital cost per unit of output will climb, which might adversely affect the position of a given local unit in competition. To avoid such negative effects, shift work is typically introduced with the aim to utilise machinery and equipment (i.e. the stock of capital) more efficiently. Consequently, it is necessary to consider the duration of machinery and equipment operation in assessing the output of a local unit.

Another aim of the above survey was to refine the method of adjustment for working-day variations[1)] in the context of seasonal adjustment in short-term economic statistics.

Since the direct relation between production and the working days indicated in a calendar has been increasingly blurred, for instance, by more flexible working hours and one- or several-day closures of businesses, calendar adjustment procedures which are solely based on the week-structure of a month (e.g. the number of working days, assuming a five-day working week) have posed ever more problems. The frequency of working days has become increasingly unsuitable as an explanatory variable in regression analyses to estimate calendar effects. Hence calendar-related systematic effects should be necessarily considered to a greater extent in order to improve the analysis.

The Federal Statistical Office has broken new ground in the above area. Scientific approaches are rare – at an international level, too.

1) See Strohm, W.: "Zur Frage der Kalenderbereinigung von Zeitreihen", in: Wirtschaft und Statistik 6/1986, p. 421 – 428 and Herbel, N.: "Zur Neuberechnung der Produktions- und Produktivitätsindizes im Produzierenden Gewerbe auf Basis 1985", in: Wirtschaft und Statistik 3/1988, p. 182 – 187.

2 Survey development

The Federal Statistical Office considered the compilation of statistics of production days for the first time at the end of 1988. A specialist committee on statistics in production industries at that time started to discuss the question to what extent the transition to flexible working hours could be covered by the statistical instruments available. A working group was set up to develop operational concepts of coverage, definitions and delimitations and to prepare the legal bases required for a potential statistical survey in that area.

The issue met with very good response. The constituent meeting of the working group in January 1989 was attended by representatives of the Federal Ministry of Economics, the Deutsche Bundesbank, the Federation of German Industries, the Confederation of German Employer Associations, the Land Office for Data Processing and Statistics of North Rhine-Westphalia and the Federal Statistical Office. Later, representatives of the statistical offices of the Länder of Baden-Württemberg and Saxony-Anhalt joined the working group as well.

Original considerations referred to the collection of information about the number of production days and times of machinery and equipment operation. Questionnaires were elaborated for that purpose. To examine the degree of acceptance of a survey of that kind, the ZUMA Institute was commissioned to test the response to the above questionnaires.[2)]

The above pretest showed clearly the difficulties arising in the context of such a survey. A decisive problem will be to cover the time of operation as mere production time, i.e. without preparation, set-up and repair periods. Business accounting does not facilitate a separation of production time from total working time or operation time.

The above-said applies in particular to local units with continuous manufacturing such as businesses of the chemical industry. Here, the procedure to be applied to determine the time of production would make it necessary to document both operation times and utilisation down to the individual plant. In the opinion of the German Employers' Association of the Chemical Industry, this would cause an unreasonable burden on respondents. Some large enterprises particularly of the chemical industry, which submitted their comments on the planned survey, arrived at a similar conclusion. For the above reasons, the chemical industry even refused to participate in the planned test survey.

Because of the apparent problems of coverage and the burden on respondents, the working group drew the conclusion to examine in the context of the test survey only the feasibility of recording the individual production days mainly for the purpose of economic analyses. The idea of recording times of machinery and equipment operation was dropped, though covering them in a follow-up survey was not ruled out.

To take into account, at least to a limited extent, the differing degrees of capacity utilisation of local units, production days should be classified in two categories, namely utilisation up to 50 % and utilisation of more than 50 %.

The working group on recording production days has covered the state of the art in its considerations.[3)] Various institutes such as the IFO Economic Research Institute, which cooperated in this area with the Institute for Employment Research, the Institute for Social Opportunities Research (ISO) and the German Institute for Economic Research (DIW) have dealt with that issue of working times and times of machinery and equipment operation. In addition, approaches of foreign institutions like the Italian Statistical Office have been taken into consideration, too.

3 Survey variables and procedure

The questionnaire included the following variables: production days with regular employment and production (degree of capacity utilisation: more than 50 %) and production days with considerably reduced employment and production (degree of capacity utilisation: 50 % maximum). Production days are days on which products are manufactured for being either sold or further processed. Days over the whole length of which products are not manufactured because of setting-up or repair works are regarded neither as general production days nor as production days with a degree of utilisation of up to 50 %. In principle, any day of the week, i.e. Saturday and Sunday as well, may be a production day.

The transition to flexible working hours has made it possible to vary the number of shifts on individual working days or the duration of shifts and the number of staff. For the purpose of analysing the economic situation, however, it was necessary to restrict the information about production days to the above-mentioned two categories, namely the

2) For more details see Ehling, M. / Porst, R. / Wein, E.: "Fragebogen-Pretest zur Erfassung von Produktionstagen und Betriebszeiten", in: "Pretest und Weiterentwicklung von Fragebogen", Volume 9 of Spektrum Bundesstatistik, 1996.

3) A detailed contribution regarding this issue was published in the final project report (as of January 1998).

days with regular production (capacity utilisation of more than 50 %) and days with a capacity utilisation of 50 % or less.

The reference period covered the months of May, June and July 1994 and included one- or several-day closures of businesses and holidays as well. The economic breakdown was based on the Industrial Classification of Economic Activities, 1979 edition, Version for Statistics of Production Industries (SYPRO).

Out of the total of units covered by the sample survey, which had been selected from the group of local units obliged to provide information for the monthly production rapid report, 5,673 units answered the questions posed to them. In cases where non-response regarding individual monthly information on production days with a capacity utilisation of more than 50% could not be settled by queries, the figures required were estimated in order to prevent a systematic underestimation of results. To this end, the average of the data reported in the branch of economic activity concerned was taken as an estimate for the missing information.

In the following, the methods and procedures of sampling and estimation will be described in more detail.

4 Methods of sampling and estimation

As mentioned in the previous chapter, the aim of the survey is to examine the feasibility of recording monthly production days in the production industries for the purpose of economic analyses. The final objective is to obtain estimated results in a breakdown by selected branches of economic activity combined with size classes of employees.

4.1 Sampling frame and population

The File of enterprises and local units in production industries (September 1992) maintained by statistical offices served as sampling frame. However, local units founded after the reference date of the above file could not be covered as an adequate frame was not available.

All local units in Germany were covered which first were obliged to provide information for the production rapid report[4] and secondly belonged to the branches of economic activity listed in Table 1:

Table 1: Branches of economic activity covered by the survey

SYPRO code	Branch of economic activity
25	Quarrying, extraction and working up of stone and earths
3011 - 3015, 3030	Cold drawing of steel, steel-wire drawing, cold rolling mills
3021 - 3025	Manufacture of drop and light smith hammer forgings, heavy pressed, drawn and stamped parts, secondary transformation of metals
31	Manufacture of structural metal products, rolling stock
32	Mechanical engineering
33	Manufacture of road vehicles, repair of motor vehicles
36	Electrical engineering, repair of electrical household goods
37	Manufacture of precision and optical instruments, clocks and watches
38	Manufacture of tools and finished metal goods (excl. electrical equipment)
50	Manufacture of office machinery and data processing equipment
54	Manufacture of wood products
63	Textile industry
64	Clothing industry
68	Food and drink industries
69	Tobacco industry

4) Pursuant to the Law on Statistics in Production Industries.

4.2 Stratification and allocation of the sample size

The survey was based on stratified random sampling. In accordance with the specified breakdown of results, the local units of the sampling frame were stratified by the branches indicated in Table 1 and by seven size classes of employees (see Table 2) and, in regional terms, by the former territory of the Federal Republic incl. Berlin-East and the new Länder. A stratification by old and new Länder was necessary because of the great structural differences between the local units in the old and new Länder at the time of sample planning. The total number of strata specified amounted to 174.

Table 2: Stratification by the number of employees

Size class of employees	Number of employees from ... to ...
1	1 – 19
2	20 – 49
3	50 – 99
4	100 – 499
5	500 – 999
6	1,000 – 4,999
7	5,000 and over

The allocation of the sample size was based on the principle of comparable precision (precision grading) and the Neyman-Tshuprov principle of optimal allocation[5] in line with the two objectives of stratification, namely the breakdown of the results and increase in the precision of a total result consisting of several strata. The allocation variable selected was the number of employees.

First, the total sample size was allocated to the individual strata (which represent combined branches of economic activity and employee size classes) in accordance with a given precision level with the aim to achieve more precise results for larger strata. This means a grading of the relative standard error μ_g of the allocation variable is provided depending on the total values X_g of this variable per stratum g:

$$(1) \qquad \mu_g = \frac{C}{X_g^{\alpha}}$$

where

g: index of the stratum,

C : constant value representing the level of error,

α : exponent controlling the degree of error grading.

Exponent α can be arbitrarily set between 0 and 0.5. If $\alpha = 0$ (extreme case), the relative standard error will be the same for all strata, while the other extreme will cause a particularly strong precision grading. In the survey discussed here, α was set at 0.3 since grading was intended to be rather strong.

As a next step, the resulting sample size n_g per stratum g was allocated to the two regional strata gj (former territory incl. Berlin-East and new Länder) based on the Neyman-Tshuprov optimum formula:

$$(2) \qquad n_{gj} = \frac{N_{gj}s_{gj}}{\sum_j N_{gj}s_{gj}} n_g$$

where

5) See Krug, W./Nourney, M./Schmidt, J.: "Wirtschafts- und Sozialstatistik: Gewinnung von Daten", 4th edition, München, Wien 1996, p. 118 – 122 and p. 114 – 118.

j : index of the regional stratum of stratum g,

n_{gj} : sample size of stratum gj,

N_{gj} : number of local units in stratum gj,

s_{gj} : estimated standard deviation of the variable Number of employees in stratum gj.

With a given sample size, the above allocation minimises the sample random error of the cell result composed of the regional strata with regard to the variable *Number of employees*.

4.3 Sampling

Before sampling, the local units of the sampling frame were arranged within the strata by Länder and number of employees per local unit. Arranging makes it possible to increase precision. Systematic random sampling including a random start was applied to select the local units.

4.4 Estimation

The sample results were in two steps raised to the population of local units of the economic groups selected. As a first step, the results reported were raised to the sampling frame with the inverse values of the selection probabilities. The estimator was the same for all local units in a stratum. The allocation of a local unit to a certain stratum depends on the stratum to which the unit belonged during sampling even if, in the meantime, it has been allocated to another branch of economic activity or another size class of employees.

To improve the quality of the sample results, the estimated values were adjusted to the number of local units on the reference day. As a result of this kind of adjustment, the bias expected due to non-response can be minimised. To this end, the local units of the sample were first subdivided on the basis of combined economic groups and size classes of employees. That subdivision proved to be appropriate since, on the one hand, the required current data were available in the File of enterprises and local units in production industries (level: March 1995) and, on the other, the above adjustment was very effective because of the largely differing response rates. Another breakdown, for instance in regional terms, would have resulted in a very small class frequency so that considerable sampling variances would have been likely.

Afterwards, an adjustment factor was computed for each class k. The adjustment factor is the quotient of the known total M_k and the estimator $\hat{M}_k$ of the local units in adjustment class k; for the sake of simplicity, the stratification index will be referred to as h (instead of gj):

$$(3) \qquad \frac{M_k}{\hat{M}_k} = \frac{M_k}{\sum_h \frac{N_h}{n_h} n_{hk}}$$

where

N_h : number of local units in stratum h (stratum size according to sample plan),

n_h : size of stratum h (according to sample plan),

n_{hk} : number of local units with analysable data in the sample in stratum h and adjustment class k,

M_k : number of local units of the population in adjustment class k on the reference day.

The estimator of the population total of a variable will hence be

$$(4) \qquad \hat{X} = \sum_k \frac{M_k}{\hat{M}_k} \hat{X}_k^{free}$$

where

$\hat{X}_k^{free} = \sum_h \frac{N_h}{n_h} \cdot \sum_i x_{hk,i}$ refers to the estimator (Horwitz-Thompson estimator) in adjustment class k. $x_{hk,i}$ is the value for variable X reported by local units i in adjustment class k and stratum h.

To determine the average number of production days $\hat{\bar{X}}$, the production days of local units were weighted with the number of employees. This way the data of larger units could be considered to a greater extent. The estimated value is obtained as the proportion of employees-days (= production days • number of employees) to the number of employees based on the following formula:

$$(5) \qquad \hat{\bar{X}} = \sum_k \frac{M_k}{\hat{M}_k} \frac{\sum_h \frac{N_h}{n_h} \cdot \sum_i x_{hk,i} y_{hk,i}}{\sum_h \frac{N_h}{n_h} \cdot \sum_i y_{hk,i}}$$

where

$x_{hk,i}$: production days of local unit i which belongs both to adjustment class k and stratum h,

$y_{hk,i}$: number of employees of local unit i in adjustment class k and stratum h.

Since information about the size of local units was not available for the reference period (May to July 1994), related data were estimated based on employment data for September 1993 and 1994 of the File of enterprises and local units in production industries.

4.5 Calculation of errors

Upon estimation, the random sampling error was estimated to judge the quality of the results obtained. It should however be noted in this context that the random error only describes the error caused by sampling and can hence be regarded as the lower boundary of the total error only. The error which is not due to sampling (systematic error) can hardly or not at all be estimated without special analyses and was consequently not taken into account in error determination.

The magnitude of the random error was estimated by calculating the relative standard error. Assuming that the sampling results are distributed normally, the relative standard error describes the interval to surround the respective sampling value (confidence interval). The true value is situated within this interval with a probability of 68%. The calculation of errors performed for the relatives was simplified, i.e. it did not consider any adjustment factors. The magnitude of the error is expected to be reflected rather accurately in this way since adjustment balances biases which are due to non-response rather than reduces the random sampling error.

5 Bases of evaluation

Since only selected branches of economic activity were covered by the sample survey, the results do not apply to manufacturing as a whole (as is otherwise typical of the reporting system applied in compiling statistics on production industries), but to 140 SYPRO four-digit items (of 204 items; without mining) only. The estimates refer to approximately 25,000 units with a total of 3.7 million employees. Due to the above difficulties, important branches such as the chemical industry are not included in the survey. The cell frequencies per branch are smaller than those for the whole range of respondents in manufacturing because selection was only based on local units obliged to provide information for the production rapid report.

Carola Kühnen, tel. (+ 49 - 611) 75 - 23 78, Erwin Wartenberg, tel. (+ 49 - 611) 75 - 28 04