# EUROPEAN INNOVATION MONITORING SYSTEM (EIMS)

**EIMS PUBLICATION N° 11**

## Evaluation of the Community Innovation Survey (CIS) - Phase I

BY


DANIELE ARCHIBUGI
PATRICK COHENDET
ARNE KRISTENSEN
KARL-AUGUST SCHÄFFER

IKE GROUP
DEPARTMENT OF BUSINESS STUDIES
AALBORG
DENMARK


d-00626 M

# EUROPEAN INNOVATION MONITORING SYSTEM (EIMS)

EIMS PUBLICATION N° 11

## Evaluation of the Community Innovation Survey (CIS) - Phase I

BY

DANIELE ARCHIBUGI
PATRICK COHENDET
ARNE KRISTENSEN
KARL-AUGUST SCHÄFFER

IKE GROUP
DEPARTMENT OF BUSINESS STUDIES
AALBORG
DENMARK

## Preface

Science, technology and innovation are areas in which policy makers inside the Community have an urgent need for better information. The EIMS is currently implementing three major data collection efforts which will improve the data available on technological innovation and diffusion.

One of these projects named The Community Innovation Survey (CIS) is a first step to build up a comparable database on innovation outputs and inputs at the firm level in the Community through a joint action between EIMS, Eurostat and Member States. The database covers about 40 000 records in 12 member countries plus Norway covering about 200 variables for each firm.

Almost in parallel to the building up of the database the implementation phase of CIS has been evaluated. The objective has been to review how the project was designed and implemented at the Community level as well as in member countries carrying out CIS. A special emphasis has been devoted to sampling issues in order to assess the comparability of the data as well as providing best practice advise in implementation of large scale innovation surveys.

The CIS action is in itself an innovative one. The data is new. It is the first time a harmonized innovation survey has been implemented and, in terms of its scope, coverage and international dimension, it will be a genuinely unique policy resource as well as providing considerable learning in building new indicators. Thus, the evaluation was designed, not as a traditional evaluation, but rather as providing experience and information in the ongoing efforts of producing reliable and useful technology and innovation indicators. We therefore think that this first evaluation deserves a wide dissemination and discussion.

As shown in the report three conditions seem to be absolutely essential in order to produce reliable and comparable innovation indicators. Firstly, there has to be a willingness among the Member States to adopt and harmonize a common questionnaire. Secondly, the evaluation shows the importance of using a common sampling methodology covering issues such as sample frames, sampling methods, stratification variables etc. Thirdly, coordination at the European level is essential if comparable results across countries are going to be created. A common legal base legal base for doing innovation surveys is essential if there is to be an effective coordination and timing of such surveys.

This report is one of several forthcoming projects using and evaluating the CIS database by DG XIII. EIMS has already started 14 different large scale projects on various aspects of the CIS database, all of which will be finalized during 1995.

Enrico Deiaco
Project Manager, EIMS

Robin Miege
Head of Unit, EIMS

# Foreword

This evaluation report is the output from the first phase of the evaluation of the Community Innovation Survey (CIS). The CIS was initiated by DGXIII/SPRINT/EIMS and EUROSTAT in 1991 as a pilot action seeking to develop a best-practice for collection of innovation data and at the same time seeking to make a database on innovation activities throughout the European Community. The evaluation was initiated in November 1993 when most of the national surveys contributing to the CIS had been closed or were in their final stage.

The evaluation was headed by Arne Kristensen, IKE, Aalborg University and was arranged as a joint project involving Daniele Archibugi, CNR, Rome; Patrick Cohendet, BETA, Strasbourg and Karl-August Schäffer, University of Cologne.

Even though the project was of a joint nature all of the members of the evaluation team had specific areas for which they were in charge. Thus, Daniele Archibugi has been in charge of Chapter 2 'State of the art' and Chapter 5 'Implementation issues'. Patrick Cohendet has been in charge of Chapter 1 'General introduction' and Chapter 3 'Methodological issues'. Karl-August Schäffer has been in charge of Chapter 6 'Statistical issues' and Arne Kristensen has edited the report and has been in charge of the 'Executive summary' and Chapter 7 'Conclusions and recommendations'. Chapter 4 'Aims and methods' has been written jointly by Daniele Archibugi and Karl-August Schäffer.

In addition to this report there is a separate annex with country reports on implementation issues and statistical issues. Karl-August Schäffer has been in charge of all the country reports on statistical issues. Arne Kristensen has been in charge of half of the country reports on implementation issues and Daniele Archibugi and Patrick Cohendet has jointly been in charge of the other half of the reports on implementation issues.

The evaluation team wishes to thank both the national contractors and the officials from the Commission for their frank and friendly co-operation in this evaluation project. Without this co-operation we would have had an impossible task.

Daniele Archibugi Patrick Cohendet Arne Kristensen Karl-August Schäffer

October 1994

# Table of contents

**Part II: Implementation and sampling**

# Executive summary

The Community Innovation Survey (CIS) was organised jointly by DGXIII/SPRINT/EIMS and EUROSTAT and it was a first attempt to collect firm-level data on input to and output on innovation throughout the European Union.

*Rationale*

In recent years the political awareness of the crucial role played by technology in economic development has been rapidly increasing. Whereas only 15 to 20 years ago technology related policy instruments were rather unknown and little utilised tools in economic policy, they have increased in importance and today they make up some of the core instruments in industrial policies in many Western countries.

However, there is a serious lack of empirical knowledge that can guide these policies and thus there is a severe risk of employing less efficient instruments to less suitable target groups - or even a risk of employing wrong policies, resulting in retardation of technological change.

Even today, data on innovation and diffusion consists mostly of indirect indicators such as R&D, patents, trade in high-tech products and technological balance of payment. Even though much information has been retrieved from this data through refinement, extension and re-classification, much still has to be explored if technology policy should be built on a sound empirical basis. What we know now can be considered to picture only the 'tip of the innovative iceberg': there is a hidden part which we must explore (cf. sections 2.2-2.4).

At the moment, direct firm-based surveys of innovation are the best possible method for shedding light on the hidden part of this 'innovative iceberg' since this type of survey can supply information on, e.g., innovative strategies, sources of innovation, barriers to innovation, innovative efforts, innovative results and diffusion of technology (cf. chapter 3).

Furthermore, as is the case with most statistics, the value of this type of information is multiplied once it is comparable across countries or over time. Thus, if quality and comparability prove to be satisfactory, this new set of international data can provide vital information for the setting up of technology policy both at the national level and at the EU level. Therefore, we believe that it was necessary and urgent to initiate such an initiative as the Community Innovation Survey and this report evaluates the surveys in all EU member states and Norway.

*Implementation of the CIS*

The process of initiating and developing the CIS was not an ideal one in which a series of different phases (creation of a conceptual framework, creation of a common questionnaire, development of guidelines for implementation and sampling, implementation across member states, creation of the database) followed each other consecutively and in a moderate pace.

Rather, due to several factors, the project was hampered. First, because of the pilot-character of the project, several difficult issues from the questionnaire had to be investigated in depth; most importantly, the possibility of measuring innovation costs. Second, because of the international character of the project the Commission had to negotiate with many national contractors, that, a priori, had very different experiences and expectations. Third, since the aim of the CIS survey was to reach comparability also with surveys in non-EU countries there were negotiations with non-EU OECD countries. Some member countries did not await this development but wanted to implement nationally developed surveys based on the Oslo manual and a draft harmonized questionnaire. These processes meant that the project became difficult to monitor and co-ordinate and that the process became partly self-driven, timing being unsatisfactory (cf. section 5.2).

Moreover, because no legal basis exists for collection of innovation data, the Commission was unable to impose any demands on member states and was limited to compiling a list of implementation and sampling issues which were presented to member states as recommendations. Thus, the Commission was highly dependent on member states' co-operation in the first CIS project. However, it is the opinion of the evaluation team that the recommendations supplied by the Commission were too general and did not cover all relevant aspects (cf. section 5.2).

It seems that there has been too little awareness among national contractors of the importance of keeping the harmonized lay-out to secure international comparability. Therefore, at the national level a variety of different aims and methods were employed, depending of the different experience and expertise of the national contractors (cf. sections 5.3-5.5, 6.2 and separate annex with country reports). Although this variety has created problems of comparability there are also useful lessons to be learned from this experience. Since the CIS is a pilot action this variety gives a unique possibility to assess the efficiency and efficacy of different methods of conducting innovation surveys.

*Outcome*

In view of the novelty of the project and the international character of the project it is the opinion of the evaluation team that the realisation of the first CIS has been successful in its first aim. From this pilot action the Commission, the national contractors and scholars in innovation measurement have learned much which can secure a high quality in a possible next CIS (or any other national or international innovation survey) (cf. section 5.8, 6.6 and 7.5-7.6). Regarding the second aim of the CIS - collection of comparable innovation data across EU member states - the level of success is more moderate. Data is not comparable across all countries and all variables (cf. section 6.5). However, it is the opinion of the evaluators that this aim - given the conditions under which the project was initiated and developed - could not be reached in this first phase of the CIS. Under the circumstances it is an achievement that a new data source

which may be used for some types of analysis of innovation in Europe (cf. later) has been created.

If we make a 'relative' assessment of the degree of success in the CIS - i.e. an assessment of the CIS compared to other types of data on technological development we may conclude that the CIS has been more successful. In one venture the CIS project has gathered and - with this report - disseminated much information on the 'field-methodology' of innovation surveys compared to what has been collected and disseminated for other types of surveys of technological development. Furthermore, there are also problems with the international comparability of these types of data even though, in several cases, they have been collected for many years. Thus, in a 'relative light' the CIS has come far in its first year.

*Overall quality of the realized CIS data*

We stated above that it should not be expected that the CIS would reach full international comparability of data in its first attempt. Therefore it is not surprising that we can conclude that, on the basis of the definitions employed in this evaluation, the results of the first CIS cannot be regarded as statistically comparable between all countries (cf. section 6.5), which implies that the analytical possibilities are restricted (cf. later).

At the time of the evaluation neither Eurostat nor member countries had completed the statistical work (margins of errors have not been calculated and analyses of non-response have not been performed). Thus, we have no exact knowledge of the quality of the realized data across countries (cf. section 6.4). This issue will be properly assessed by Eurostat in building up the EU database and until these margins of error are calculated results should be interpreted with utmost care.

*Reasons for lack of comparability*

We concluded above that, as might be expected, there are problems with comparability of data. In our view the main factors to account for this lack are:

- Some contractors modified some questions and thus questionnaires were not comparable between all countries.
- The survey frame was not satisfactory for a few countries.
- The sampling methods were not sufficiently harmonized.
- High levels of total and/or item non-response occurred in some countries.

Two other problems exist at the time of the evaluation (according to Eurostat and DGXIII these problems will be solved in the compilation of the EU database):

- Raising factors have not been calculated for all countries and used in the aggregation of data.
- Margins of error have not been calculated.

The evaluation has shown that the main factors to account for these deficiencies are:

- Lack of co-ordination.
- Lack of instructions about or expertise on sampling and implementation.
- Lack of awareness of the importance of international comparability.

Again, however, it is possible to go back one more step and assess the reasons for these problems. These can, in our opinion, be summarised into six points:

1. The international character of the project. The Commission had to negotiate with national contractors that, a priori, had different experiences and expectations. Furthermore, since a key priority for the CIS was to make the data comparable also with other OECD countries, negotiations took place with the OECD. These processes delayed the project and some member states did not await the results of this endeavour but implemented the survey before a final harmonised questionnaire was agreed on. Therefore both timing and harmonisation was retarded.

2. Lack of co-ordination power. Since no legal basis exists on collection of innovation data the Commission was compelled to make only recommendations on sampling and implementation to national contractors. They could make no demands on the services to be rendered by national contractors and they could not pick the best possible national contractors. This seriously hampered harmonisation and in some cases influenced the quality of data.

3. Lack of advice. The set of recommendations worked out by the Commission was not sufficiently detailed. Even though still voluntary in nature the recommendations could have been more itemised, providing detailed advice to some of the more inexperienced national contractors. This may have reduced the quality of data for some countries.

4. Lack of will. Even though all national contractors agreed on the importance of creating an internationally comparable data base of innovation statistics, some of the national contractors did not seem to have the will to comply with this aim of the project. Therefore they introduced various national-specific changes and this hampers the comparability of the data.

5. Lack of expertise. It seems that in a few countries national contractors did not have the full economic or statistical expertise to carry out the innovation survey in a satisfactory way. This hampers the quality of the data for these countries.

6. Lack of comparison of experiences. Too little was done to facilitate an interactive learning process where national contractors could learn from each other (best practices, errors, difficulties, etc.).

*Analytical possibilities with the CIS data*

Three different uses of the CIS data may be envisaged:

- Descriptive analysis of differences between countries.
- Analysis of innovation in selected industries across countries.
- Analysis of innovative structures within countries.

Within each of these uses a variety of projects may be performed. However, for all analysis margins of error must be taken into consideration.

Since errors may be smaller if non-innovating enterprises are left out of the analysis, i.e. if explorations are restricted to the set of innovative enterprises, such analysis may provide more manifest results. In these explorations also the data from Greece and Portugal may be used.

The CIS data should not be used for assessment of EU totals. For example assessment of total innovation costs in EU or the share of turnover used for innovation across EU cannot be made because of deficiencies of the data for some countries. For policy-related advice these data may be used to assess issues like the non-R&D costs of innovation, the sources of and barriers to innovation, R&D co-operation and innovative strategies, etc. Because comparability between countries in some cases is low the data should not be used for detailed analysis of differences between countries that result in detailed policy advice on the (re)distribution of EU resources between countries or regions, or on initiatives aiming at harmonising structural and institutional factors across countries.

*Recommendations for future innovation surveys*

The evaluation has shown that co-ordination at the European level is essential if comparable results across countries are going to be created. Innovation surveys are a new initiative and are not yet backed by a such solid experience as data on the main economic indicators. A common standard is still sought, and to achieve this requires close co-operation among the various organisations that are involved in the field. It is the opinion of the evaluation team that a satisfactory co-ordination can only be achieved through a legal basis and consequently it is recommended that a legal basis for innovation surveys is adopted. On the basis of a legal basis the Commission should

- Co-ordinate the venture, i.e. make sure that timing is appropriate
- Select the best possible national contractors (or contractor teams), i.e. contractors with relevant experience
- Create a revised pre-tested questionnaire to be implemented in all EU countries (using the lessons gained during this first round of the CIS).
- Work out detailed instructions of all aspects and levels of the implementation of the surveys on the national level. I.e. recommendations on:

  Target population, cut off point, timing, reminder procedure, follow up on responses, survey method, frame, survey unit, sample size, sampling technique, subsampling for non-response, imputation of missing data, raising factors and assessment of reliability (cf. section 7.6).

In case these issues are too detailed to be included in a legal basis it is recommended that the legal basis is made as a frame which can be filled in by the Commission. It is recommended that the basis for creating the legal basis is the experience from this first round of the CIS.

*The learning effects*

One thing that is considered very important by the evaluation team is the learning effects of repetitive innovation surveys. These include both 'internal' learning effects (in repeating this venture both the Commission, national contractors and respondents will have learned a lot from this first CIS), and, equally important, 'interactive' learning effects (a horizontal process in which national contractors learn from each other).

If the venture is repeated, the learning effects will ultimately imply that innovation surveys will provide more reliable and more comparable data. This information will be invaluable for the design of technology policy, both at the national level and at the EU level.

# 1. Introduction

The aim of this report is to evaluate the implementation phase of the Community Innovation Survey (hereafter, CIS) as a pilot action seeking to establish the basis for long term data collection and analysis on innovation.

As for any evaluation procedure, before presenting the methodology and the main steps of the assessment of the rationale, implementation and results of the CIS, a brief presentation of the history of CIS and its main objectives will be given.

*Objective of CIS*

CIS is an initiative jointly initiated and implemented by Eurostat and DG XIII (SPRINT Programme, European Innovation Monitoring System (EIMS)). It was developed between 1991 and 1993 in co-operation with independent experts and the OECD. CIS is based on a common questionnaire ('EC harmonised questionnaire' (hereafter, harmonised questionnaire)). The questionnaire was based on the 'OECD Guidelines for collecting and interpreting data on technological innovation - the Oslo manual' (hereafter, Oslo manual).

The basis for the CIS action is that innovation is a fundamental element in competitiveness and economic growth and therefore is a vital component of economic and industrial policies. As the importance of innovation is continuously increasing, the need for a better understanding of the innovation process has become increasingly urgent. The effectiveness and the coherence of the economic and industrial policies strongly depend on the reliability of the information available for decision making.

The problem that policy makers face is that, until very recently, the main data resources for innovation policies consisted of Research and Development (R&D) statistics and patent data. Although much has been achieved with these data, R&D and patent data have fundamental limitations when it comes to mapping input and output of innovation processes.

Thus, the objective of CIS is "to collect firm-level data on inputs to, and outputs of, the innovation process across a wide range of industries and across Member States and regions, and to use this data in high-quality analyses, which among others, will contribute to the future development of policies for innovation and the diffusion of new technologies at Community, Member States and regional level " (EIMS/Eurostat, 1994).

*Specific characteristics of CIS*

With reference to this global objective, CIS represents specific features that must be reflected in the attitude of the evaluation. In many ways CIS is an innovative initiative in itself and it should be interpreted (and evaluated) not as a well-achieved action, but as step in an ongoing process.

As an innovative initiative, CIS represents three main characteristics. First, the data is new. There has never before been internationally comparable data on non-R&D resources devoted to innovation and the output of the innovation processes. Second, it is the first time that a harmonised business survey has been implemented in all EU Member States. Third, the harmonised survey will not only give policy makers and analysts information on the sectoral level, but also give them a detailed picture of innovation activities at the level of European enterprises.

As a process, the CIS initiative is an open ended project which started in 1991 involving different phases. In a first phase of the project a model questionnaire for innovation surveys in Member States was developed by Eurostat, based on the Oslo manual and input from a Eurostat working party. In 1991/92 this questionnaire was pre-tested in five countries (the so-called Kleinknecht project - see Kleinknecht, 1993). On the basis of this pre-test and a new Eurostat working party a revised questionnaire was developed by Eurostat. During the spring of 1992 this questionnaire was further developed by an expert group and in April 1992 a draft harmonised questionnaire was presented at the OECD NESTI (National Experts on Science and Technology Indicators) meeting. The final harmonized questionnaire was agreed in June 1992 by Officials from EC and OECD and leading scholars in innovation surveys.

In 1992/93 the national surveys were implemented and now the international data base is being constructed in Eurostat. The next stage will be devoted to a series of analytical projects, both at the EU level and at the national level.

Not all EU countries took part in the first phase of the CIS and also non-EU countries were involved to secure comparability with non-EU OECD countries. Thus, the process described above was sometimes self driven and difficult to control. The evaluation of the CIS must take into account this particularity of the CIS action.

*Methodology and structure of the evaluation*

The evaluation of the CIS initiative aims to assess the rationale, instruments, procedures and results of the venture, and to propose suggestions and recommendations for future innovation surveys. It must be emphasised that the first implementation of CIS was a pilot action. Since it is possible that in the future innovation surveys will be conducted on a regular basis, it is important to draw lessons from the pilot phase that will assist future implementation of such surveys. Furthermore, it is hoped that the evaluation will assist analysts in understanding the quality and therefore usefulness of the realised data.

The methodology that has been chosen to investigate how the CIS respects the objectives that have been exposed above (with regard to the specific characteristics of CIS on which we insisted) follows three steps:

First, the contextual and theoretical background of the CIS initiative has been investigated:

- Why is it of particular importance today, for policy-makers as well as for analysts, to obtain reliable data on innovation?
- What are the main surveys on innovation already realised in the EU countries or elsewhere, with what results and what lessons learned?
- What are the specific difficulties related to the measurement of innovation?
- Is the conceptual background of the harmonised questionnaire (the Oslo manual) adapted to the explicit objective of the CIS initiative?

In order to answer these questions that aim to describe the context of the CIS action, a survey of the recent initiatives to measure innovation has been proposed, along with a theoretical discussion on the Oslo manual (Part I: Chapters 2 and 3).

Second, the actual implementation of the CIS has been investigated:

- How was the survey implemented in the EU member countries?
- What is the quality of the realised samples?
- How reliable is the aggregated data and results?
- What best practice recommendations can be developed?

To answer these and other questions the evaluation team undertook interviews with the national contractors that carried out the statistical survey in their respective countries. Furthermore, an analysis of realized data was performed (Part II: Chapters 4-6).

Third, a synthesis of the main results of the evaluation is proposed. Its aim is to conclude on the overall quality of the data collected and to give recommendations for best practice in future implementations of innovation surveys at the EU level. This summary results from discussions within the evaluation team, as well as discussions with the different participants in the CIS initiative (Part III: Chapter 7).

In a separate Annex we present country reports on sampling and implementation and the harmonised questionnaire in the final version.

# Part I

# Conceptual background

# 2. Measuring innovation: Rationale and state of the art

This Chapter provides a presentation of the rationale and the state of the art in innovation surveys. The rationale insists on the importance of the regular collection of data on technological innovation, and on the recent efforts made to provide reliable measures in this domain. The basis of innovation surveys is the recognition that innovation and diffusion are fundamental elements in competitiveness and economic growth, and hence are vital components of economic and industrial policies. However, the increased attention on the importance of innovation has also revealed limitations on available data. It has become increasingly clear that the data which is generally used for policy analysis and thus policy advice - particularly R&D statistics on the input side and patents data on the output side - is too narrow in scope and does not provide accurate measures to understand and describe the innovation process. That is the reason why many researchers on innovation have attempted to develop new forms of data to permit a descriptive and detailed account of the elements of innovation process.

The state of the art, which is presented at the end of the Chapter, focuses on the main topics on which the proliferation of innovation surveys have provided valuable information and discusses the main problems encountered by innovation surveys already carried out.

## 2.1 Measuring innovation: its application for analysis and policy

The role of technological change in economic and social life has been increasingly emphasised over the last decade. In particular, it has become clear that innovation plays a pivotal role in the performance of industrial companies. Contemporary economic systems have become more and more knowledge-intensive and show a technological dynamic much larger than past societies. Economics and the other social sciences should therefore provide better explanations of the determinants and impact of innovative activities. These explanations need to be backed by empirical evidence which, in turn, depends on the availability of quantitative data.

There are strong policy implications related to the availability of reliable quantitative data for describing innovative activities. Policy makers need reliable data which help them find relevant answers to such questions as: "What are the main non-R&D inputs to innovation?", "How do

patterns of innovation expenditure vary within and across industries?", "How do low R&D industries acquire new technologies?", "What factors do firms perceive as the main obstacles to successful innovation?", etc. A clear understanding to these questions could greatly improve the setting up of appropriate policies to stimulate the competitiveness of European industry.

In recent years, governments have paid more attention to the innovativeness of the business sector. This has led to a substantial increase in the amount of funds and in the number of schemes to promote innovation. National governments have not been alone in stressing the crucial role of innovation for economic growth and welfare: international organisations such as the OECD, UNIDO and UNESCO have equally emphasised it. Over the last decade, EU has also substantially increased the funds available to member countries for innovative projects. To assess these initiatives and to increase their efficiency and their efficacy reliable indicators are needed.

However, the increased recognition of the role of innovation has also revealed strong limitations existing data. It has become increasingly clear that the data which is generally used for policy analysis (particularly R&D statistics on the input side, and patent data on the output side) is too narrow in scope and does not provide adequate measures to understand and describe the innovation process. Thus, although it is generally recognised that innovation is the engine of a variety of economic phenomena, including growth, productivity, competitiveness and employment, the actual measurement of the impact of innovation still relies on proxy measures such as R&D and patenting. The lack of accurate measures for innovation has severe consequences for our understanding of economic dynamics and for the design of appropriate policies. There is obviously a high risk of designing wrong policies based on wrong indicators. Moreover there is a high risk to use inadequate existing data just "because it is there", and not because it is particularly good. Thus, one can see policies focusing on limited areas: those which are precisely highlighted by existing data.

The reasons for the limitations of existing data on innovation are manifold. Two main reasons at least can be mentioned:

First, in spite of the crucial role played by technology in everyday life the methods used to measure it appear still rudimentary if compared to those for measuring economic variables such as production, investment, trade or employment. While information on these variables is already available with satisfactory accuracy, data on technological change are not yet sufficiently detailed in terms of comprehensives, disaggregation and comparability because the regular collection of data on technological innovation is relatively recent. Data on science and technology have been collected by the majority of advanced countries for less than 30 years. Economists and science policy analysts have been ingenious in using proxy measures to account for innovation. Statistics on the resources devoted to R&D are now available and are as reliable as any other economic indicator. A large body of literature contains ad hoc data for crucial aspects of innovation. However, comprehensive data on innovative activities are still in their infancy.

Second, technological innovation is probably the single most heterogeneous economic category. A jet engine, a microprocessor, no less than a corkscrew or a hair-pin might be technologically innovative. Needless to say that their technological, economic and social significance will be very different, and that they will not necessarily overlap. Because of the very nature of innovative activities, it is not easy to deal with them using quantitative methods.

## 2.2 Comparing different technological indicators

There are a number of indicators to account for science and technology activities (for a review see, for example, Van Raan, 1988; Science and Public Policy, 1992). They include: resources devoted to R&D, patenting, the balance of payments for technology, trade of high technology products, bibliometric indicators, surveys of the introduction and of the diffusion of innovations.

These indicators are periodically analysed by national and international organisations. The US National Science Foundation regularly produces a *Science and Engineering Indicators.* For three decades, the OECD has collected and standardised data on resources devoted to R&D and other indicators, which are published regularly in the *Science and Technology Indicators Report.* Recently, EU (DGXII) has initiated the work on a European Science and Technology Indicators Report. The indicators used in these reports can be divided into two main groups.

Some of the indicators were originally collected and developed for accountancy reasons. Indicators such as patents, the balance of payments for technology, or trade of high-technology products fall into this category since none of these statistical sources were originally created to explore the nature of technological change. Patent data was collected by patent offices for administrative and legal reasons, data on the balance of payments for technology for currency control and data on high-technology products by customs. Students in the field of technological change have considerably refined, extended and re-classified the original data. In several cases this job has proven to be very demanding. In spite of all these efforts, some of the shortcomings related to the very nature of these data have not yet been overcome.

Other indicators have been collected with the specific aim of investigating the nature of science, technology and innovation from the very beginning. This is the case for R&D statistics, bibliometric indicators and, more recently, innovation surveys. Each of these indicators has its own advantages and disadvantages. And, more importantly, each of them inform on different aspects and stages of the innovative process. Table 2.1 lists the main strengths and weaknesses of the most commonly used indicators : Research and Development activities, patenting, balance of payments for technology, trade of high-tech products, bibliometrics and innovation surveys.

*Table 2.1 Measures of technological activity*

| | Strengths | Weaknesses |
|---|---|---|
| Research and Development | - Regular data collection<br>- Sectoral uniformity across industries<br>- Internationally comparable | - Excludes design, software and production engineering<br>- Underestimates innovation in small firms<br>- Underestimates innovation in services<br>- Monetary adjustments required for international comparability |
| Patenting | - Regular data collection<br>- Detailed break-down for technological fields<br>- Internationally comparable | - Not all inventions are patented<br>- Not all inventions are patentable<br>- Does not inform on services<br>- Differences in the propensity to patent across sectors |
| Technological Balance of Payments | - Regular data collection<br>- Detailed break-down for technological fields<br>- Internationally comparable | - Does not inform on non-transferred technology<br>- Measure only a small part of technological activities<br>- Data biased by financial transactions |
| Trade of High-Tech Products | - Regular data collection<br>- Direct measure of performance<br>- Internationally comparable | - Does not consider innovation in traditional sectors<br>- Does not inform on domestic innovation<br>- Problems of selecting the pertinent products |
| Bibliometrics | - Detailed break-down for technological disciplines<br>- Internationally comparable<br>- Direct measure of scientific output | - Databases include a sub-set of world publications<br>- Differences in the propensity to publish across disciplines<br>- Language barriers |
| Innovation Surveys | - Direct measure of innovativeness<br>- Potentially it includes all activities related to innovation<br>- It is applicable to manufacturing and services | - Problems in comparability over time and across countries<br>- Lack of periodicity in data collection<br>- Problems of sample definition<br>- Data can be biased by subjective judgements |

None of these indicators can claim to represent all the aspects of technological change. For example, R&D activities represent one source only of innovation, even if particularly important, but there are several other concurrent factors in the generation of innovations such as good marketing or in-house incentive schemes, etc.. The same applies for the balance of payments for technology since some firms innovate even if they do not report any international transaction for technology. Furthermore, in spite of the importance of high technology, the

majority of resources for innovation are absorbed by medium and low technology products. Finally, innovation surveys do not consider the activities carried out in universities and other public research centres which do not have a commercial application, etc.

If we consider as a detailed example the case of the patents indicators, the main advantages of these indicators may be summarised as followed:

- Patents represent the outcome of the inventive process, and more specifically of those inventions which are expected to have a business impact. They are a particular appropriate indicator to capture the proprietary and competitive dimension of technological change.
- Obtaining patent protection is time consuming and costly. It is likely that applications are presented for those innovations which, on average, are expected to provide benefits which compensate these costs.
- Patents are broken down by technical fields, providing information not only on the rate of inventive activity, but also on its direction.
- Patents statistics are available in large numbers and for very long time series.

On the other hand, patents also have several disadvantages:

- Not all inventions are patented. Sometimes, firms protect their innovations with alternative methods notably industrial secrecy (on the trade-off between patenting and industrial secrecy, see Wyatt et al., 1985; Levin et al., 1987).
- Not all inventions are technically patentable. Until recently this has been the case for software, which has an increasingly important role in the current technological advance and which, after a long controversy, is now protected in the majority of countries by copyright.
- The propensity to patent (that is, the number of patents registered for each unit of inventive and innovative activity, see Scherer, 1983) greatly varies across technological areas and industries. While in certain fields, such as pharmaceuticals, a large part of the inventions are codified in patent applications in other fields, such as nuclear physics, only a handful of patents can to be found.
- Firms have a different propensity to patent in each national market, according to their expectations for exploiting their inventions commercially. The size of national markets and the level of integration in international trade affects the number of foreign patent applications received by each country. Moreover, national patent offices receive a large number of applications by domestic inventors and firms, thus they are biased towards the domestic inventive activity (Archibugi and Pianta, 1992).
- In spite of the international patent agreements among the majority of industrial countries, each national patent office has its own institutional characteristics; the attractiveness for applicants of any patent institution depends on the nature, costs, length and effectiveness of the protection accorded.

These considerations do not diminish the significance of the indicators mentioned above, first and foremost because better indicators are not yet available. However, the indicators can be considered to picture just the tip of the innovative iceberg but there is a hidden part which we

need to explore. Direct surveys on innovation are at the moment the best possible method to acquire information on the hidden part of the innovative iceberg in industry.

## 2.3 Two alternative approaches to innovation surveys

Two main approaches have been used to monitor innovation with direct surveys. The first has taken as the unit of analysis the individual innovation. Additional information was often collected and recorded, such as the size and main product line of the firms responsible for its introduction. This approach has been implemented, with different methodologies, by the SPRU database (Townsend, 1981) and by the US Small Business Administration database (see Acs and Audretsch, 1989). Information on commercial innovations have also been collected from technical and scientific journals and magazines in a variety of countries (for an overview of this approach, see Kleinknecht, 1993). We will call this the *object approach.*

In the second approach the unit of analysis is the firm. Both innovating and non-innovating firms are considered. This approach allows the interviewer to gather information on a large variety of aspects related to innovation, including its sources, determinants and obstacles. This is the approach used by the majority of innovation surveys. Among them, the Italian innovation surveys (Cesaratto et al., 1991), the Ifo survey (Scholz, 1992), the Innovation survey of the Nordic countries (Kristensen, 1993), etc. Using the definitions of the Oslo manual, the CIS initiative has explicitly adopted this second approach. We will call it the *subject approach.*

The two approaches present different problems of standardisation, and consequently of comparability across countries and over time. Table 2.2 on the next page highlights their strengths and weaknesses, with reference to different criteria as unit of analysis, method of collecting information, international comparability, etc.

### *2.3.1 The object approach*

The main problem encountered by the object approach is that it is difficult to identify a common yardstick to classify innovations. As stated above, innovation is a very heterogeneous phenomenon and it is very difficult to give a comparable value to each of them. Not surprisingly, very different definitions have been applied by each survey. The SPRU innovation survey, for example, has taken as measurement unit *significant* innovations introduced in the UK over a 40 year period, which resulted in a sample of about 4,800 observations. The Small Business Administration has monitored innovations *commercialised* in the US in one year only, and the sample is of about 8,000 observations. In several respects, innovation counts are similar to patent statistics, although they generally provide more information for a lower number of observations. For example, about 34,000 patents were granted to American inventors in 1982, but the available patent databases provide only a basic information on each patent granted.

*Table 2.2 Innovation surveys. A comparison of the object and the subject approach*

| | Object approach | Subject approach |
|---|---|---|
| Unit of analysis | The innovation | The innovator |
| Method of collecting information | Collected from selected items such as new products, innovation inventories, patents, bibliometric directories | Collected at the level of the economic organisation such as firms, universities, research centres either by questionnaires or interviews |
| International comparability | Low, except for patent based indicators | Potentially high for qualitative data. Lower for 'soft' data |
| Comparability with R&D | Low since R&D data are not collected by projects but from organisations | High since R&D is part of innovation costs |
| Comparability with national accounts | Low because it is difficult or even impossible to relate it to the whole economic universe | Potentially high on quantitative data as innovation surveys can be related to the economic universe |
| Time series comparability | Generally high | Potentially high if data is collected periodically and is standardised |
| Coverage of innovations from the non-business sector | Generally high | Potentially high if surveys cover also the non-business sector |

Regardless of the degree of detail which can be reached by innovation counts, it is almost impossible for any survey to gather information on all the innovations introduced. This creates a basic problem for comparability since it is not possible to relate the sample of innovations monitored to any population. This will seriously hamper the possibility of making cross-country comparisons even if an international database were to be developed.

Furthermore, innovation counts according to the object approach are not easy to relate to national accounts data. Data on employment, sales, value added and R&D are generally collected at the enterprise level and they can be connected to classifications of innovations only indirectly.

The main advantage of the object approach is its ability to obtain direct information on the innovation introduced. This approach can, for example, include innovations developed not only in the in the business sectors, but also by government agencies, universities or other non-profit organisations. It also allows time series comparisons, as shown by a body of literature on the business cycle.

### *2.3.2 The subject approach*

The subject approach has more or less the opposite strengths and weaknesses of the object approach. Since the statistical unit is the enterprise, accurate directories of the entire population are available in all industrial countries. A sample of enterprises monitored can therefore be related to the entire population. This in turn allows international comparisons to be drawn. Since the unit of observation is the enterprise, the information of innovation can also be directly related to other data from the national account.

However, in practice, the subject approach has not yet been able to provide firm comparisons over time because questionnaires have only occasionally been kept identical. Moreover, non-business organisations have seldom been monitored, and this has meant that the innovations introduced by the public sectors, universities, etc. are not monitored.

The differences mentioned above between the object and the subject approaches show that the target population of the former is the universe of innovations only (since non-innovations cannot be a statistical unit). The latter, on the contrary, include both innovating and non-innovating firms.

## 2.4 Innovation surveys: the state of the art

Over the last decade, a dramatic proliferation of innovation surveys has occurred. Industrial associations, universities, public and private research centres and statistical offices have produced their own innovation surveys. Reviews of the literature have been attempted by De Bresson (1986), Archibugi (1988), Hansen (1992), Smith (1992), Kaminski (1993) and Kleinknecht (1993). Furthermore, a special issue of *Science Technology Industry Review* (1992) presents results from surveys carried out in France, Germany, the Netherlands, Italy, Sweden and Norway. However, not even these reviews combined will provide a comprehensive list of the innovation surveys carried out so far. Table 2.3, "Characteristics of recent innovation surveys" (next page), attempts to present the most important surveys over the last two decades.

Innovation surveys have provided invaluable insights into the nature of technological change and empirical evidence for policy makers. Their proliferation indicates how important the gathering of information on innovation has been seen. The majority of these initiatives were tailored to specific needs and for specific research programmes. Almost all of them have been carried out at the national level, although a few had a larger scope.

*Table 2.3 Characteristics of recent innovation surveys*

| Name | Country | Nature | Number of observations | Method | Purpose |
|---|---|---|---|---|---|
| SPRU | UK | Object | 4800 innovations | Selected on the basis of expert opinions | Significance of innovation<br>Inter-industry differences<br>Sources of innovation |
| ISRDS-CNR-ISTAT I | Italy | Subject | 24700 firms (16700 innovating and 8000 non-innovating firms) | Census analyses | Sources of innovation<br>Obstacles to innovation<br>Process and product innovations |
| ISRDS-CNR-ISTAT II | Italy | Subject | 8200 innovating firms | Innovative firms selected from previous survey | Innovation costs<br>Sources of innovation<br>Process and product innovations |
| YALE survey | USA | Subject | 600 questionnaires | Sent to a sample of managers | Sources of innovation<br>Appropriability of technology |
| US Small Business Administration | USA | Object | 8000 commercialised innovations | Collected from technical track and engineering journals | Size of innovating firms<br>Market structure |
| Scherer matrix | USA | Object/ subject | 440 large companies | Patents classified by sector of production and use matched with R&D expenditure | Inter-industry technology flows |
| First French survey (Piatier) | France | Subject | 5300 observations | 5000 questionnaires and 300 interviews | Number, type and novelty of innovations<br>Innovation costs<br>Obstacles to innovation<br>R&D forecasts |
| Second French survey (OST) | France | Subject | 15000 observations | Questionnaires | Innovation output<br>Sources of innovation<br>Sectoral patterns |
| Nordic survey | Nordic countries | Subject | 650 innovating firms | Questionnaires<br>Sample: R&D performing firms | Innovation output<br>Sources of innovation<br>Innovation costs<br>Obstacles to innovation |
| IFO surveys | Germany | Subject | 3000 observations | Questionnaires<br>Panel data | Investigation on innovative activities and firm strategies |
| Dutch surveys | Netherlands | Subject | | Questionnaires | R&D activities and patenting<br>IT application and development<br>Training |
| NSF survey | USA | Subject | 600 observations | Questionnaires | Innovation output<br>Appropriability problems<br>Innovation function |

In the sections below, the surveys carried out before the Oslo manual was available will be considered with a view to draw a few lessons for the assessment of the CIS venture. The Oslo manual and the CIS have in fact given a new impetus to develop new and more comprehensive databases and have substantially improved the landscape of innovation surveys. By now surveys on the basis of the Oslo manual have been made, are under way, or planned in almost all OECD countries outside EU (for example, Finland, Canada, Australia, Switzerland, USA) and several other countries have also implemented innovation surveys, e.g. P.R. China and the former Czechoslovakia.

We will focus below on five main topics on which innovation surveys have provided valuable information namely: 1) the non-R&D costs of innovation 2) the role of small and large firms in innovation 3) the sources of innovation 4) the barriers to innovate; and 5) taxonomies of innovating firms. There are certainly other fields which could be also considered as relevant (for example, the role of high, medium, low-tech branches or firms), but the five topics described in details in the following paragraphs are by far the most commonly used factors.

### *2.4.1 The non-R&D costs of innovation*

The main advantage of innovation surveys is to uncover activities which are not captured by traditional indicators such as R&D and patenting. This information can be gathered around two main questions. First, how many firms are innovative? Second, what is the amount of resources they devote to innovation and what is the type of activity at stake?

The first question was addressed by several surveys, including the first Italian survey (CNR-ISTAT 1). One of its main findings was the very high number of innovating firms monitored: 16,700 out of the 24,000 participating firms declared they had innovated. Only 16 % of them, i.e. 2,700, declared that they had also performed some sort of R&D.

But the finding that as many as 2,700 firms were somehow involved in R&D activities was unexpected. In fact, the Italian annual survey on R&D monitors about 1,000 firms only. The Italian results were consistent with other surveys carried out elsewhere. In particular, it was shown that a similar underestimation of R&D active firms occurred in the Netherlands (see Kleinknecht and Reijnen, 1991 for the Netherlands, and Archibugi et al., 1987 for Italy). These studies have indicated that several small firms did not participate in the R&D survey either because they did not receive the questionnaire or because it was too complicated for them to respond. In quantitative terms, the unrecorded R&D appears to be small.

Two lessons can be drawn from this experience. First, the number of firms engaged in R&D activities is larger than previously expected. Second, innovating firms are much more numerous than R&D performing firms.

Another crucial issue addressed by several innovation surveys is the resources devoted to innovation. Another relevant class of information is its composition among R&D, design, investment, marketing, etc. Already in the early 1970s it was pointed out that R&D accounts only for a part of total innovation costs. More evidence has been collected in recent years for countries such as Italy, Nordic countries, and Japan. The Italian survey showed that R&D accounts for about 21 % only of the total innovation costs, while according to the Nordic survey they account for about half. This has tremendous implications for policy making. We focus more on non-R&D aspects of innovation in enterprises.

However, these results are highly dependent on the definitions of both R&D and innovation. What should be included? How is innovation defined? While some standard R&D definitions

were provided long ago, only recently have similar attempts been made for innovation. The differences that emerged between countries could be due either to nation-specific factors or to some differences in the definitions provided in the questionnaires.

### *2.4.2 The role of small and large firms in innovation*

A more comprehensive measure of innovation will also allow the testing of some of the classical hypotheses of the economics of technological change, namely the relationship between industrial concentration, market structure and innovation. These topics have often been addressed from a policy perspective. The assessment of the potential of small and large firms is required to guide innovation policy and to distribute funding for industrial innovation.

The relationship between market structure and innovation will be of particular importance for the European Union. What effects will European integration have on industrial innovation? Will the increasing concentration which is expected to occur within Europe increase or decrease the amount of resources devoted to innovation? Which policies and schemes are more likely to promote innovation in small or large firms?

The available technological indicators, and most notably R&D and patents, are not suited to test the relationship between innovation and market structure since they are biased in favour of large companies and tend to underestimate the innovations of small and medium sized firms. Evidence based on innovation surveys has rehabilitated the innovative potential of small firms, especially in traditional industries (see, for example, Pavitt et al., 1985 and Acs and Audretsch, 1988). An analysis based on innovation costs suggests that small firms are less innovative than large ones, although their distance from the latter is much larger in terms of R&D resources (Archibugi et al., 1994). The database developed by the CIS will for the first time provide empirical evidence to address these questions on the basis of more comprehensive information.

### *2.4.3 The sources of innovation*

A large theoretical and empirical literature has indicated that the sources of innovation vary considerably across firms of different industries and size (Nelson and Winter, 1982; Pavitt, 1984; Levin et al., 1987; Von Hippel, 1988 and Dosi, 1988). In particular, it has been shown that traditional industries use external sources to back their innovations. This body of information has enormously enlarged our knowledge of the devices which allow firms to be (or not to be) innovative, and has helped policy makers to tailor innovation policies to industry-specific needs.

The various surveys have consistently shown that inter-industry differences are similar across countries and over time. A systematic comparison of the findings is, however, problematic because the methods used to gather information do not allow direct comparisons. Kaminski (1993), for example, made a series of comparisons on sources of innovation in Italy, France and Germany based on industry rankings rather than on absolute values. He met several

problems due to different industrial classifications, different sampling procedures, different ways of presenting results from apparently similar questions, different writing methods, etc.

The research carried out on the sources of innovation has followed very different methodologies. The majority of surveys have asked their respondents to provide information based on the experience of their own firm. The Yale survey on appropriability (Levin, 1987) has instead asked the respondents to provide information on their whole line of business. Von Hippel (1988) has studied the sources of a selected number of significant innovations.

### *2.4.4 The barriers to innovation*

The concepts of innovation failure and of factors hampering innovations are slightly different. The former implies that an innovation has been attempted and possibly even developed but that its outcome was a failure. Factors hampering innovations, on the contrary, concern non-innovating and innovating firms alike. The Sappho project has studied the conditions for innovations to be successful (Freeman, 1982). Piatier (1984) has considered the barriers to innovation. Several other surveys have asked to firms to indicate the obstacles which have made it more difficult or impossible to innovate. The approaches employed so far are not complementary and will not allow the drawing of general conclusions.

### *2.4.5 Taxonomies of innovative firms*

One of the most significant outcomes of the recent literature on technological change has been the attempt to classify firms or industries according to technological criteria rather than according to the two standard criteria of main product line and size. Pavitt (1984) presented a taxonomy of innovating firms based on several criteria. Pavitt's taxonomy has proved tremendously robust and within one decade has been widely applied in industrial economics. Some have tested this taxonomy at the industry level (Archibugi et al., 1991; Kristensen, 1993), others at the firm level (Cesaratto and Mangano, 1992). However, only occasionally non-innovating firms were included. New data at the firm level are required to further develop these taxonomies and to test their validity in more than one country. Such taxonomies will allow the development of policy instruments which are tailored for a specific target groups (which in turns will reduce waste or increase the effectiveness of innovation policies).

## 2.5 Conclusions

We have mentioned above only a few of the topics which have been addressed by innovation surveys. As already started, the Oslo manual and the CIS venture have made it possible to overcome several of the shortcomings encountered so far. At the moment, the main problems of the innovation surveys already carried out can be summarized as follows:

- Individual surveys, even within the same country, have often overlapped. New surveys have ignored the problems encountered and the lessons to be gained by previous experiences. This has led to several duplications and eventually to a waste of resources.

Firms have often been overloaded with questionnaires and interviews with similar scope. Often, this has reduced their willingness to participate in these surveys.

- Only occasionally a systematic comparison of the results obtained in different countries was possible. This is one of their major shortcomings compared to other indicators, most notably R&D and patents. Even when national questionnaires were broadly comparable, a few differences in the wording, implementation or methodology has not allowed international comparisons to be made. It is uncertain which of the measured differences should be credited to country-specific or to survey-specific factors.
- The lack of a standardised approach has not allowed, even within countries which have conducted more than one survey, to make comparisons over time.

The recent standardisation of innovation surveys proposed by the Oslo manual and the CIS can for the first time overcome some of these shortcomings. Innovation surveys will offer more than the already available indicators if they will satisfy the requirements of *standardisation, international comparability, and periodicity*. The information provided will have a much larger scope if it will be *compatible* with national economic accounts as well as with other technological indicators, including R&D. In assessing the CIS initiative, we will therefore pay special attention to these requirements.

# 3. Conceptual background

## 3.1 Introduction

The official version of the Oslo manual, 'OECD Proposed Guidelines for Collecting and Interpreting Technological Innovation Data - Oslo Manual' was published by OECD in 1992. It is expected that the manual, after the first round of surveys based on it, will be revised and further developed in the light of experience gained - in particular the experience from the CIS project. Therefore, the Oslo manual should not be considered as a bible, but as an impetus to develop new and more comprehensive data bases on innovation. What will be investigated in this chapter is the theoretical coherence of the manual and its feasibility as conceptual background for innovation surveys.

At the end of the chapter we present a detailed discussion of the connection between the Oslo manual and the harmonised questionnaire. In this discussion the Oslo manual will be taken in its actual form without being questioned. The point at stake is the coherence between the conceptual background expressed in the manual and the actual questions that have been included in the final version of the questionnaire ('EC Harmonised Innovation Surveys 92/93 - Final Questionnaire'. See Annex B).

## 3.2 The coherence between the Oslo manual and innovation theory

The Oslo manual clearly states that its theoretical basis is the "chain-link" model proposed by Kline and Rosenberg (1986). This model conceptualises innovation in terms of interaction between market opportunities and the firm's knowledge base and capabilities, with feedback between all parts of the process. A key element in determining the success (or failure) of innovation is the extent to which firms manage to maintain effective links between phases.

The Kline–Rosenberg model is in one of the best models that conceptualise the improvements made over the last two decades in the ways to represent and analyse the process of innovation. It is now fully recognised that innovation is basically a learning process. It is neither an exogenous promethean gift nor a multi-purpose knowledge base that can be oriented according to relative price changes. This neo-classical view of the 1960s accommodated well the models

of homothetic growth. With the crisis of industrial restructuring and the emergence of new technologies that came later we were incited, if not obliged, to look closer at the inner characteristics of technology and they, by no means, matched such a view. Technological innovation is a process which occurs differently across industries and over time (Pavitt, 1984); it is at the same time localised, partly tacit and to a large extent history-rooted and with strong irreversibility character that makes it strongly path-dependent and of very limited transmissibility. Let us say a few words about each of these characteristics.

The "local" nature actually encompasses two interrelated dimensions. The first is the recognition that technological innovation takes place within a particular structure, a specific context of industrial products and production processes. More fundamentally the local character expresses a sort of "fixed point" that results from a cumulative process of successive adjustments of the technology to its context. In such a process a specific configuration of technological, sectoral and scale-related arguments gradually match each other to create a specific "personality" or pattern[1].

The tacit aspect is linked to the fact that production technology is about doing things, not only about knowing things in the form of an abstract (scientific) principle. Operating know-how is something very different and is much less transmissible (Dosi, 1988). It is often dependent upon a division of labour which very often contains implicit components that make the production setting unique. It incorporates working routines and craftsmanship components and sometimes, even with the voluntary transfer of the original blueprints, it is not possible to imitate the ensemble of gesture and implicit working practice in such a way that the product quality is satisfactory.

The third and most powerful argument of the technological learning process is its historical nature. The view of technological development as an evolutionary process is now well established due to the joint effort of N. Rosenberg's (1976, 1982) historical research and the implementation of the evolutionary approach by Nelson and Winter's continuing effort (1982, for example). Irreversibility of the technological learning process was further documented and analysed in the work of P. David (1975, 1988) and of B. Arthur (1985) on path-dependency. In addition to the localised aspect there is a clearly time-dependent behaviour in learning. If for some random reason a given technological trajectory is selected, it will undergo successive improvements that will make it perform better and it will, by this very fact be selected for

---

[1] The consequences of this local character are far from trivial and in 1969 Atkinson and Stiglitz presented it as selective displacement of a point or a segment rather than a shift of the overall production function. And as Sahal (1981) put it "... insofar as the learning tends to be localised, the very concept of a production function representing a near-infinity array of techniques sharing a common state of knowledge may have to be abandoned".

future applications. If the current choice had to be made previously (before the first choice was made) an altogether different technique could have been chosen[2].

All three major characteristics of the learning process mentioned above imply, in one way or another, severe restrictions on the ability to transmit the information contained in the novelty. D. Foray went a step further in the argument (1989). He claims that the standard theory actually mixes up different types of diffusion effects. The main one consists of the dissemination of the information about the outcome of an R&D activity. Such dissemination occurs without serious adjustment costs while the actual implementation of the innovation usually requires considerable adjustment costs. If the former diffusion costs are low it is because the firm invests a lot in maintaining a considerable R&D capacity that can quickly absorb such information. To put it differently, the adoption costs which are also a measure of appropriability become an inverse function of the investment in R&D. As Pavitt has argued his view is particularly valid for basic research, where, although the results are publicly available, it takes constant investment effort in theoretical research to be able to interpret and use abstract results. Even this really public good can only be appropriated with significant costs which are required to maintain the learning process alive.

To a very large extent the Oslo manual is coherent with this conceptual background. However, we would like to insist on some theoretical points that could have been extended it more details in the Oslo manual in particular, and in innovation surveys and measurements in general:

a) The "local" nature of the innovation process. At the level of enterprises this means the full recognition that technological innovation takes place within a specific industrial structure, a specific industrial context that contributes to shape the innovation activities in a specific manner. Practical ways to identify these specific structures and to assess their impact on the innovation process are thus required. How for instance business structures (multidivisional, hierarchical, etc.) can enhance or hamper the innovation process, how modes of enterprises organisation (structure of projects for realising innovation, how inter-industrial modes of organisation interface with the innovation process (networks of enterprises, group of subcontractors, etc.). These are questions on which the Oslo manual could have provided more details.

b) The distinction between the information contained in innovation (which can be patented, coded, etc.) and its tacit aspect (in operating know-how for instance) is certainly a critical distinction which is particularly difficult to investigate by interviews. The first aspect of

---

2 This is of course a major cause for a restricted transmission of knowledge. Since we can only keep alive a relatively small number of options – our "practice memory" capacity being very limited – purely theoretical options have a very small chance of being reactivated. After the oil shocks, no one really considered simply going to coal-based organic chemistry. This would have created new hydrocarbons from liquefied coal and then used petroleum-based techniques that have been constantly improved.

innovation being obviously easier to apprehend than the second. However, learning (in any form) and, in general, cumulative aspects of innovation should certainly deserve a more in-depth investigation in a future version of the manual.

c) Irreversibility and path-dependency are also key characteristics of the innovation process. Those irreversibilities that lead to the standardisation process, are phenomena which deserve more attention. In particular, the *speed* of the diffusion process is an important parameter (which is not constant in general) which is relatively ignored by the Oslo manual.

d) If the interactive nature of the innovation process is fully recognised by the Oslo manual, some of the interactive features are not, or only to a small extent, apprehended by the manual. This is the case, for instance, for the *frequency* of interactions between different components of the enterprises.

## 3.3 Discussion of the theoretical limits of the Oslo manual

The Oslo manual is very clear about the limits of the theoretical background which is actually proposed. These limits are easily explained by the complexity of the innovation process that necessitates drastic choices in order to get to a version that can lead to practical application. However we shall explore again these limits in order to identify the possibilities of further developments and improvements in innovation surveys in general, and in particular which could be possibly incorporated in further developments of the Oslo manual

### 3.3.1 The choice of the "subject" approach to innovative surveys

As explained in Chapter 2 there are two different approaches that can be used to monitor innovation with direct surveys.

The first one, the object approach takes the individual innovation as unit of analysis. Once the list of successful (or unsuccessful) innovations has been established, once can survey enterprises which introduced them and explore various factors which have influenced the innovations (this approach has been chosen by SPRU for instance).

The second one, the subject approach, takes the firm as unit of analysis and explores the innovative behaviour and activities of the enterprise as a whole. The idea is to explore the factors influencing the innovative behaviour of the enterprises (strategies, incentives, results and barriers to innovation).

The Oslo manual has clearly chosen the second one, mainly for reasons of comparability and international standardisation which is fully understandable. However, this choice implies some limitations to the Oslo manual that should be detailed.

The Oslo manual subject approach makes it extremely difficult to track the process of diffusion, specially to track flows of innovation and technological change from one industry to

another and hence to track the spillovers of productivity raising activity related to the evolutionary nature of innovation as a learning process is the ability to stimulate change within and across the frontiers of the firm. This time the appropriation matter relies not only on the same innovation, but rather on the chain of induced effects. Spillover effects attracted considerable analytical attention over the last decade and attempts were made to relate their magnitude to technological infrastructure (Jaffe, 1986), or to the capital and productive structure.[3] We will come back to this later. To better grasp the spillover effects one usually distinguishes between productivity spillovers and inter- or intra-industry spillovers.

Productivity spillovers express the fact that, in addition to the surplus gains for the innovating industry, these (typically materials or process) innovations tend to reduce input prices in downstream industries and therefore increase their own surplus. These, however, are not always considered as pure externalities since they are transmitted through the market mechanism.

The industry spillovers stem from the informational nature of the innovation, considered as a commodity that can be easily appropriated. The innovating firms can expect other firms, either from the same industry or from altogether different sectors, to appropriate relevant information about the innovation properties with relatively negligible costs (problems of rivalry). These kinds of spillovers are made easier with procedures and practices such as skilled labour mobility, systematic screening of data bases of relevant professional literature, reverse engineering and sometimes technical intelligence gathering. However, although we should recognise these "negative" spillovers, we should notice at the same time that a rapidly increasing number of firms expect globally positive feedback effects from the different modes of technological transfer. It is certainly the case for technological partnership agreements, for the "Technopole" – that is, new technologies industrial breeding areas – and for the voluntary dissemination of R&D results to potential competitors in order to sustain strategic markets. Inter-industry transfers are even less problematic and very often, as in the case of advanced materials, one cannot really implement innovation unless a significant scaling-up allows for price reduction that makes the whole diffusion possible. Indeed, diffusion is often a controlled process where the reciprocity rule prevails in a long repeated game. In addition, very often the development of an innovative "milieu" proves to be essential and information exchanges are

---

[3] One usually finds the claim that the social rate of return exceeds the private one, but in no way can we realistically conclude that the spillover effects escape altogether the innovating firms and are to be considered as being a negative incentive for RD investment. A recent increase in the rate of technological partnership formation especially in advanced technologies does seem to indicate that spillovers are mastered to a large extent by reciprocity agreements and the partners do find that the benefits of sharing and sometimes even diffusing knowledge overcome the welfare or profit loses. Actually many economists now believe that these R&D externalities should be taken into consideration in the new normalisation of real growth. Although their macro magnitude is hard to assess – hence the difficulty to say to what extent they offer an escape from decreasing returns – it is quite possible that the return to the originating firms becomes higher than their losses even though they may capture only a small part of the social overall return.

the very way of participation. In all these instances spillovers are of a "positive externality" type and there is absolutely no reason to conclude that under-investment in innovative activity would prevail.

All these innovation phenomena related to diffusion cannot in principle be taken into account by choosing a subject approach.

The subject approach cannot track how innovation processes evolve at different speeds across time, with irreversibilities, hysteresis, thresholds, etc. These characteristics are precisely those which are at the basis of the definition of technological trajectories. Even if questions raised for the Oslo manual are concerned with "significant" innovation (radical innovation which implies a real threshold) or incremental innovation, it must be recalled that this distinction is the one viewed by the individual respondent of the enterprise. But a significant innovation viewed by an individual is not in general a significant innovation for the economy. What could be viewed ex ante as a major innovation by society. Society can realise only ex post that some of the innovations can be considered as significant ones.

In this context, we must be extremely careful in understanding what we really can compare when we analyse two successive surveys of innovation realised with reference to the Oslo manual. We cannot compare the evolution of innovation, nor track the process of diffusion of innovations through the industry. We compare changes of behaviours by firms, changes of innovative contexts through time, changes of innovative capabilities and impact of innovative on economic performances.

The subject approach is not appropriated to deal with "technological fields" (biotechnologies, information sciences, chemistry, etc.). However, some studies try to encompass this type of information in a subject approach by asking firms what type of technologies they have introduced.

### 3.3.2 How to go further with the subject approach

Since the subject approach is a real opportunity to go "into the black box", it seems that some points could be investigated in more details in the Oslo manual with reference to this approach:

As mentioned above, more emphasis should be given to the relationship between business structures and innovative behaviour in order to understand the importance of organisational innovations. It is proven that some of the organisation choices (matrix or multidivisional form) of enterprises facilitate the innovative process. It is also currently admitted that enterprises are choosing specific internal temporal forms of team organisations to monitor innovation. In an innovation survey of large manufacturing companies realised by INSEAD companies were asked about the influence on innovations resulting from organisational measures as overlap product-process engineering, planned joint activities with vendors, manufacturing personnel with design experience etc. This is also the case with the *project* structure, which is rarely taken into account by the Oslo manual. Beyond these project structures many fundamental questions for the enterprises are at stake: how to cope with the classical division of competence

when dealing with projects, does the enterprise need a specific project team, etc. If we consider inter-enterprise relationship, more emphasis should be allocated to the influence on innovation activities resulting from the participation of the enterprise to *networks*, either formal ones (subcontractors) or informal ones.

Related to the above discussion, a more specific attention to the problem of *selecting* innovations. What are the mechanisms at stake? The Oslo manual pinpoints some of them (imitation, following of a leader, etc.) but should insist on this fundamental question particularly in order to know how the selection mechanisms operate within the firms (on what criteria does the firm select innovation projects?).

Also, it is important to know how the *internal* transfer of technology is realized within a given enterprise. Some specific measures, some specific structures, could favour this process, which is certainly a key one at least in medium and big enterprises.

Since the focus is on the behaviour of innovative firms, the question of the role and perception of governments (viewed by industrialists) to intervene on the points of innovation, to stimulate it, to set up efficient service organisms in order to help the transfer of technology or the financing of innovation is particularly relevant. The Oslo manual could devote more attention to this important question specially along two points that are neglected in the questionnaire. On the one side the role and interaction with so-called public structures of transfer of technology, on the other side the role of public research and innovative programs (EC research programs, for instance). One of the questions at stake is the perception by firms of the necessity of governments to intervene in the process of innovation, because of a potential feeling that they cannot appropriate (at least to some extent) the benefits of innovation (hypothesis put forward by K. Arrow in 1962).

We should also consider very seriously the problem of SMEs[4]. It is not obvious that we should encompass in the same innovation survey large enterprises and SMEs. (In fact, on a practical basis, most smaller SMEs are not taken into account by national surveys that adopt cut off point and generally exclude enterprises under 20 employees). The behaviour towards innovation differs dramatically between these two types of enterprises. For example, belonging to networks of contractors, the support of services of transfer of technology, the local environment, etc. are important factors for most of SMEs, while they can be considered as negligible for large units. On the other hand, for large units the multi-product type of activities, the nature of the organisation (hierarchical, multidivisional, etc.), the market share, are variables of major interest that do not play a similar role for SMEs. Thus a separate survey could be envisaged for SMEs.

---

[4] Small and Medium Size Enterprises

## 3.4 The Oslo manual and innovation surveys

As mentioned above the Oslo manual has proven its robustness vis-à-vis utilisation for innovation surveys. The main blocks of questions (sources of information for innovation, objectives of innovation, acquisition/transfer of technology, R&D activity, factors hampering innovation, innovation costs, etc.) are clearly set-up in the questionnaire and clearly understood by most of the respondents. However, what we would like to insist on is the use of the Oslo manual when connection with other surveys than the EC harmonized survey on innovation is at stake. It seems that the Oslo manual could be improved in a way to better interact with other type of questionnaires.

Concerning the questionnaires on R&D ("Frascati type") the very problem when comparing the two surveys is the fact that the Frascati manual relies on the lineal model of innovation, while the Oslo manual relies on the interactive model of innovation. Thus, there could be theoretical incompatibilities between both manuals and, as a consequence, between the two related questionnaires. However, it seems that in its conceptual design, the Oslo manual could manage the interaction with the Frascati manual more systematically and insist more on the innovative aspects related to R&D activities and the innovative aspects that are not related to R&D activities. For instance, in the Kline & Rosenberg spirit, it could be important to define the relationship between the different activities of the firm (marketing, production, etc.) and the R&D department or the R&D activities. In fact, 4 categories of situations to investigate can be defined in a systematic way:

*Table 3.1: Innovative and R&D activities*

| | |
|---|---|
| Innovative activities<br>With R&D activities | Innovative activities<br>Without R&D activities |
| Non innovative activities<br>With R&D activities | Non innovative activities<br>Without R&D activities |

Any firm observed should belong to one of these four categories (when R&D activities are present, it is also important to distinguish between internal and external R&D). Of course, the size of the four categories are not of the same order of magnitude. In particular, the category "non-innovative activity, with R&D activities" could appear as purely hypothetical. However one can conceive companies making pure R&D (in biotechnology, for instance) without realising real innovation, in the economic sense (those companies can be assimilated to private laboratories). A systematic presentation of this in the Oslo manual seems to be important because the EC questionnaire, as we have seen earlier, does not always make a clear distinction which makes it possible to isolate the R&D component.

### 3.4.1 Connection with other surveys

The experience of the last few years, and the tendency observed at the level of national contractors both show evidence that many types of surveys related to some extent at least to the innovation surveys are in an emerging phase. Among these potential or actual surveys, we can quote:

–Surveys on appropriability (the Yale surveys).

–Surveys on organisation.

–Surveys on regional aspects of R&D.

–Surveys on the problems of normalisation, quality and certification.

–Surveys on technologies.

–Surveys on SMEs.

All these fields are developing rapidly, and a global coherence requires that there should be some way to avoid overlapping and to favour – at least to an acceptable level – harmonisation between all these surveys and the possible future CIS surveys.

## 3.5 The coherence between the Oslo manual and the EC harmonised questionnaire

In this section, we shall discuss the coherence between the Oslo manual in its actual form and the . harmonised questionnaire. What is questioned is the quality of the "translation" of the manual into the questionnaire. The discussion is made with a positive attitude for, in many points, the questionnaire just follows the recommendation of the manual (example: the sources of innovative ideas selected in the questionnaire are to very large extent those detailed in the manual in § 130). However, a series of remarks could help to clarify the "translation" in the spirit of the manual. We shall begin with general remarks, then we shall end with detailed remarks according to the presentation of the harmonised questionnaire.

This section, which is essentially technical, requires that the reader has both the EC harmonised questionnaire (see annex) and the Oslo manual . The discussion will follow the order of topics that appear in the questionnaire.

### 3.5.1 General remarks concerning the coherence between the Oslo manual and the EC questionnaire

*General aim of the questionnaire*

In its paragraph 87, the Oslo manual clearly states that among the two main approaches to collecting data on innovations – either working on significant innovations, or exploring the innovative behaviour of firms – it recommends the latter because this approach is "more amenable to international standardisation" (§ 88). Thus, the manual concerns more the subject

of the process of innovation (firms), than the objects (technological groups, technological trajectories, etc.). If this is the real philosophy of the manual, then the way the questionnaire is introduced (Introduction: "The questionnaire is concerned with technological innovation"... Definitions: "New products", "New processes", ...) could be slightly misleading for the respondents vis-à-vis the spirit of the Oslo manual because it insists mainly on the object. Some words about the specific focus of the manual would be welcomed in the Introduction to orient the respondent. They must know that the focus of the survey is to understand the innovative behaviour of firms. Then the definitions of technological innovation can be given.

*Importance of innovative activities*

The Oslo manual pinpoints that one of the most important purposes of a questionnaire on innovation of this type is to provide a measure of the innovative activities of firms (which firms are innovative, and which are not innovative). In the Oslo manual from § 107 to § 119 a detailed (not exhaustive) list of innovative activities is provided (research and experimental development, tooling up and industrial engineering, etc., design). The questionnaire does not take this list into account and tackles this topic with only three questions (questions 1 to 3 in the EC questionnaire). It seems that, with reference to the Oslo manual, more emphasis should be given to the importance of innovative activities in the questionnaire. A way to improve the questionnaire and to determine the type of innovative activities of the firm could be to clearly distinguish : "innovation of product", "innovation of process", "innovation of organisation", "innovation of design", "innovation of packaging", etc. The purpose of this detailed distinction is to avoid limiting the breakdown of innovative activities to solely product and process innovations, and thus to offer to firms a larger spectrum to determine their innovative behaviour. If such a spectrum is not offered, there is the risk that fewer firms than what is actually the case will declare themselves innovative.

Moreover, the corresponding block of questions in the questionnaire ("General Information about Innovation Activities") implies a clear distinction between those firms that are innovative (those that at least answer "yes" to one of the three questions) and those that are non-innovative (those that answer all "no"). But such a clear-cut distinction could be exaggerated if we admit that a firm which has just a vague intent to develop a new product or process and which has answered "no" to questions 1 and 2 will be considered as innovative and be treated completely differently. A detailed block of questions for this critical item should be recommended to avoid such an ambiguity (to determine the "intensity" of the innovative activity of the firm)

*Relationship between R&D and Innovation*

References to R&D data in line with the Frascati manual are numerous in the Oslo manual (§ 22 to 51 and § 192 to 202 particularly). If it is clearly indicated that R&D is only a small part of the innovation process, the complementarity between the Fascati approach and the Oslo approach is fully recognised. In order to facilitate the connection between the two related questionnaires, it seems that the EC harmonised innovation survey could:

– Explain briefly in its introduction, the differences and complementarities between the concepts of R&D and of innovation.

– Be structured in a way that makes it possible to benefit from the two approaches, and that prepares clearly the exploitation of the complementarities. For example, in Chapter II of the questionnaire ("Sources of Information for Innovation") it is not possible to isolate the pure R&D components and to make at this level a comparison with some of the results coming from the Frascati questionnaire. This question could have been a good opportunity to understand for this topic the complex relationship between R&D and innovation.

In the harmonised questionnaire, questions about R&D activity are concentrated in Chapter V, and it seems that this chapter is not adequately designed to allow the exploitation of the complementarities with questions coming from the Frascati manual. The whole EC questionnaire should be structured in such a way as to allow the best complementarity with the (revised) Frascati manual.

In general, the questionnaire should be designed in a way that allows the best complementarities with other surveys related to innovation (for example the Yale survey on appropriation and surveys on innovation in organisations).

The focus of the questionnaire explains that the questions related to the diffusion of innovation, to the flows of innovation from one industry to the other, to the measures of the spillovers are extremely limited in the final questionnaire despite the recognition of the importance of diffusion by the Oslo manual (§57 to 61) and (§179 to 190). The Oslo manual explains clearly why it is difficult to develop indicators which give a more detailed picture of user sectors and the need for surveys of technology. But there should be more questions in the questionnaire about diffusion (example: speed of diffusion within the enterprise, etc.).

### 3.5.2 Detailed remarks concerning the coherence between the Oslo manual and the harmonised questionnaire

We shall expose the detailed remarks by following the EC questionnaire. We are aware that some of the remarks suggesting the introduction of more detailed explanations or questions in the questionnaire are in conflict with the need to maintain the questionnaire as short as possible. However, in several cases we believe that more explanation or more detailed questions are necessary.

*Introduction*

Remark 1: The introduction starts with the sentence: "This questionnaire is concerned with technological innovation". With reference to the above discussion, it is suggested that the questionnaire should start with the real objective of the questionnaire which is concerned with the innovative behaviour of business enterprises. We should be careful to formulate the questionnaire so as not to discourage non-innovating firms not to fill the questionnaire. This

care is needed to stop firms who may be innovative, but do not consider themselves so, from believing that the survey do not concern them.

Remark 2: In the Oslo manual §90, it is indicated that "innovations therefore involve a series of scientific, technological, organisational, financial, and commercial activities". This presentation seems more informative for the respondent than the one given in the introduction of the questionnaire "the innovation process is the combination of activities – such as design, research market investigation, tooling up, and *so on*". It seems particularly important to mention that financial and organisational activities may have an impact on innovation.

*Definitions*

Remark 3: In the distinction of the two types of product innovation ("significant" and "incremental") it must be clear that the difference between what is "new" and "old" is the one perceived by the respondent himself. We refer here to the well known discussion in the literature about the fact that an individual could have a perception of achieving a significant innovation which is only incremental for the economy. It also must be clear that the reference is made to "ex post" attitudes than to "ex ante" attitudes. This clarification is needed to clearly understand questions 1 and 2 rather (where the respondent is confronted to "ex post" interpretation) and question 3 (where he is confronted to ex ante interpretation, which is extremely difficult and vague).

Remark 4: For complex products such as cars or television sets, respondents could have difficulties in defining what is "new" or "old". Indications given in the Oslo manual § 102 could be helpful.

Remark 5: For new processes, one could precise (as it done for new product) "what do we not include as an innovation". With regard to this point, explanation given in §105 of the Oslo manual is extremely helpful, specially the example concerning the "just in time" system (which is treated as innovation), which could be ambiguous for many respondents without any explanation.

*General information*

Remark 6: The Oslo manual suggests (§248) that questions on the average life span of the firm's product groups could be asked at this level to assist evaluation on the data on its innovation output. This question could be useful to distinguish between different types of firms.

Remark 7: The questions about economic activities refer to years 92 and 90. If the expected frequency of the innovation survey is three years (Oslo manual § 266), a three year interval could be adopted for this item as well.

Remark 8: The questions about "General information about innovation activities" have already been discussed in 3.5.1).

It also seems that question 3 (ex ante strategical vision of innovation) is difficult and not really discussed in the Oslo manual. This question raises many difficulties. For instance, a "yes" could be answered by enterprises preparing a new innovation model of car to be on the market next year, and enterprises having a vague intention to introduce some changes in their products in the coming years. Moreover, from a logical point of view, with reference to questions 1 and 2, there should be a distinction between products and processes, therefore if it is maintained question 3 should be split into two parts.

*Sources of information for innovation*

Remark 9: As mentioned earlier, the breakdown of internal sources seems to be insufficient: the importance of the "in-house R&D" for understanding the link between R&D and innovation would at least justify introducing this source in the block "internal sources".

Remark 10: The role of public or private centres of technological transfer should be explicitly mentioned in the sources of innovation.

Remark 11: Even if we can expect some enterprises to consider in the block "other external sources" the participation to in industrial or research networks, the importance of this variable should justify a special item in the list of sources of information.

*Acquisition/transfer of technology*

Remark 12: This block is related to section V.7.3. of the Oslo manual, and appears to be more detailed than the manual. However in the list of the forms of Acquisition or Transfer, the role of inter-enterprise agreements (on exchange of technology) and the role of specific (private or public) organisms of transfer of technology should deserve a specific mention (even if we may expect respondents to give them in the line "Other").

Remark 13: Questions about the internal transfer of technology (within firms) could be relevant.

Remark 14: The Oslo manual does not recommend explicitly for this block of questions any time reference. The EC questionnaire has chosen to refer to what happened the year before (1992). But justification for this choice is not obvious. May be a three year period could (as for many other questions) be envisaged. As mentioned by some contractors, there is a problem because most of the questions of this block refer to 1992 only, whereas question 9 are asked for 1990-92.

*R&D activity*

Remark 15: In the Oslo manual, it is suggested that the innovation survey should include a question on the existence of a central R&D unit and the distribution of intramural R&D

between it and other departments of the company. This suggestion should be taken into account.

Remark 16: As mentioned in the Oslo manual (§ 201) questions on participation to European and other international and national R&D programs are of particular interest for understanding R&D attitudes of firms.

Remark 17: In question 10 it is never obvious if we refer to both internal and external R&D or not.

*Factors hampering innovation*

Remark 18: The Oslo manual proposed to isolate the factor R&D: "R&D expenditure too small" within the block of enterprise factors. The EC questionnaire did not follow this suggestion. It seems to be a pity because, once again, this variable is a key one to assess the complementarity between R&D and innovation.

For the rest of the question, the harmonised questionnaire follows to a large extent the Oslo manual.

*Costs of innovation*

Remark 19: As mentioned in the Oslo manual, "measuring the total cost of innovation activities in enterprises and industries is one of the major aims of innovation surveys" (§.215). The Oslo manual clearly suggests that they are two ways to approach the cost of innovation:

– To measure the total expenditure on innovations (including activities that do not in the end lead to innovations, including for instance expenditure on projects that have aborted).

– To measure expenditure for innovations introduced in a given period (excluding expenditure on projects that are aborted and on general R&D expenditure not connected to any specific product or process application.

The manual recommends the first approach (which is coherent with the conceptual approach of the manual). But in the questionnaire, this very fundamental distinction is not clear, and some respondents could be misguided. It seems, at least, that in the information block which explains question 13, some details about the approach must be given.

Remark 20: As suggested in the Oslo manual a breakdown between "Intramural R&D expenditure" and "extra-mural R&D expenditure" is needed. The questionnaire does not make any distinction between these two critical variables.

Remark 21: Training expenditure in connection with the introduction of new products and processes should be separated items in question 13 (cf. Oslo manual §227).

Remark 22: It is not clear if the question is related to product innovation costs or to product and process innovation costs. It seems that the second approach is coherent with the Oslo manual, but some national contractors understood it differently and took the first approach.

## 3.6 Conclusions and recommendations

In conclusion, the questionnaire appears as reasonably adapted to the manual, particularly in some key topics as the determination of the barriers to innovation, where the questionnaire closely follows the recommendations of the Oslo manual. However with the experience gained by the CIS project, one can propose the following recommendations:

- The questionnaire should clearly announce that it is built on the *subject approach,* and hence that it is focused on the innovative behaviour of firms. This could even be clearly reflected on the title of the survey which could be "Survey on the Innovative Behaviour of Firms", instead of "Survey on Innovation".

- Questions to identify the "innovativeness" of firms should be restated. In the present questionnaire, the division between innovative and non-innovative firms is depending on the sole central question, "Are you innovative, yes or no?" - This induces several possibilities of bias:
  - First, there could be some firms especially among big firms, that will answer "no" because they have not realised any major innovations during the period, even if they may have made several minor innovations.
  - Second, there could be some firms, especially among small firms that will answer "yes" because they want to emphasise a minor innovation that they have realised, even if in general they are not innovative at all.

  The frontier between innovative and non-innovative firms could then be blurred, with a risk of bias. To avoid this potential bias, one possibility may be to include a multinomial (instead of a bi-nominal) question on the intensity of innovation: "are you: non innovative; slightly innovative; moderately innovative ...".

- Among the sources of innovation, it is proposed to isolate clearly the R&D component in order to facilitate the interpretation of the importance of R&D and non-R&D input. Although R&D is just one of the many sources of innovation, there are many theoretical issues that give special attention to the link between R&D and innovation. In particular among the R&D expenditures, it is proposed to distinguish systematically between intra-mural R&D and extra-mural R&D to analyse the issues related.

- The questions on the costs of innovation should be restated to facilitate the understanding. Among the variables, a breakdown between intra-mural R&D, and extra-mural R&D, should be made.

- In the case of SMEs, it is suggested to formulate specific questions, if not to a specific questionnaire.

However, if in some aspects the questionnaire should be made more adapted to the Oslo manual, there are also some aspects of the Oslo manual that can be improved at the light of the implementation of the CIS initiative. The suggestions emerging from this experience can be found in the Part II of the evaluation report.

# Part II

# Implementation and sampling

# 4. Aims and method of the evaluation of the CIS

## 4.1 Introduction

Part II of this report is devoted to analyse in detail the experience of the Community Innovation Survey. We will review how the project has been designed and implemented at the Community level and within the individual member countries of the European Union. The CIS is the largest initiative undertaken so far to acquire information on innovation in industry, since all member countries and Norway have participated in a common survey. The core of this survey is the Oslo Manual and the European harmonised questionnaire, which have already been discussed in Part I of this report.

The harmonised questionnaire, however, has been used far beyond the European Union. All OECD countries have implemented, or are going to implement, similar surveys (see OECD/ NESTI, 1994). Several ex-planned economies have also performed surveys which are, at least partially, based on the Oslo Manual. There is, however, a specific reason which justifies an evaluation exercise limited to the member countries of the European Union and Norway. Within the CIS framework 13 countries participated in the design of a common statistical survey with the explicit purpose to gather internationally comparable results.

Although it would also have been relevant for the understanding of the problems of data collection on innovation to explore the surveys carried out outside the boundaries of the European Community, we have not gathered them for two reasons. First, other surveys were neither financed nor promoted by the European Commission. Second, the majority of them are still at their initial stages[1].

As in any new survey an extensive process of trial and error has occurred. The fact that similar, although not identical, surveys were carried out in such a high number of countries has increased the number of experiences. As we shall see, the background and the expertise of the

---

[1] An overview of these surveys is offered in OECD/NESTI (1994).

national performers of the survey were rather different. A variety of approaches have been followed to implement this survey. This has two implications for our evaluation. On the one hand, we had to explore to what extent these differences hampered the possibility to compare the data obtained in the various national surveys. On the other hand, the large variety of experiences we saw gave the possibility to assess which of them are more suited to provide information on innovation.

The issues concerning the design and implementation of the survey are discussed in Chapter 5. The statistical issues concerning sampling and realised data are considered in Chapter 6. In the next section we describe the aims and the method used for this evaluation. In section 4.3 we define some of the key terms which have been used for this evaluation.

## 4.2. Aims and method of the evaluation

According to 'The update of the Community Innovation Survey', our evaluation should be directed to "evaluating the data collection methods and the methodological aspect of the implementation procedure in order to identify the best practice of innovation survey implementation" (EUROSTAT, 1994). The aims pursued in this evaluation can be synthesised as follows:

1. to assess how the survey was carried out in member countries;
2. to pick up the best practice experiences in the methodology and implementation with a view to extend them across all member countries;
3. to check the possibility to compare the data collected across different countries.

Our evaluation started when the surveys were almost completed in all countries. Even where some tasks were still in progress, the method and design of the survey had already been defined. Therefore, our evaluation is mainly *ex-post.*. However, we also provided some advise on the actions to be taken before the closing of the survey. At least in part, we have also made an *on-going evaluation.*.

Any evaluation needs to collect information on the project which is to be evaluated. This information can either be publicly available or collected from the participants in the research project. The majority of the information we have used to assess the CIS was collected from the participants. Whenever possible, however, we have also used written documentation, including statistical materials.

We interviewed the national contractors and the personnel of the European Commission involved in the project. Since the interviews were our key source, an effort was made to get the maximum amount of information out of them. The information needed was split into two main parts, implementation and sampling. We prepared a set of interview guidelines on each of these two aspects which were sent in advance to the contractors. These guidelines provided the basis

for the interviews held. Several contractors answered in writing to all or to parts of the guidelines. The detailed information about each country is provided in the country reports in Annex A.

The aim of the interviews was to collect information on three different aspects:

First, we wanted to know how the survey was carried out in terms of procedures and sampling. This was needed in order to assess differences across countries and to assess the quality of the data.

Second, we wanted to collect the opinions of the contractors on the usefulness of the exercise and on the design of the questionnaire. The contractors carried out the survey and used the questionnaire in field work. As all users, they are familiar with the strengths and weaknesses of this venture.

Third, we also wanted to collect information from the contractors on the attitudes of the respondents to the survey. The contractors had the closest contact with the people who should provide information about innovation, i.e. the respondents, and they are the most suited persons to describe how they reacted to this specific survey.

Table 4.1 (page 42) provides an overview of:

1. the national contractors;
2. the sub-contractors or the collaborators to the survey;
3. the dates of the meetings;
4. the persons who participated in the interview;
5. the evaluators who participated in the interview.

With one notable exception (the Wallonee region), we managed to meet all contractors[2]. The interviews allowed us to collect detailed information on how the survey was carried out in each country. The contractors have been very helpful, and we wish to thank all the interviewed colleagues for their friendly collaboration. In some cases, we have not yet received the additional information required. We planned to interview the national contractors for a working day. The interviews lasted from a minimum of 4 hours to a maximum of 7 hours with an average of about 6 hours.

Information from DGXIII (SPRINT/EIMS) and Eurostat was also acquired by means of an additional interview. This meeting was prepared on the basis of *ad hoc* guidelines and lasted approximately 3 hours. We also interviewed Keith Smith who acted as an consultant to DGXIII and thus played a crucial role in the design and management of the CIS.

---

[2] Neither the contractor nor the sub-contractor of the Wallonee region participated to the interview held in Brussels on the 10th of February 1994. We contacted both of them several times and we were promised a written report to answer the questions of the interview guidelines.

*Table 4.1 Overview of contractors and interviews*

| Countries | Contractor | Sub-contractors or collaborators | Date of the meeting | Interviewees | Evaluators |
|---|---|---|---|---|---|
| Belgium-Flemish | Institute for Science and Technology | University of Gent | Brussels, 10.2.94 | J. Larosse, I. Fleurent | Archibugi, Schäffer |
| Belgium - Wallon | Wallonee Ministry | Federal office for scientific, technical.. | | | |
| Belgium-Brussels | Science Policy Office, Prime Minister's Service | Université Catholique de Louvain | Brussels, 10.2.94 | N. Pinsart, I. Fleurent | Archibugi, Schäffer |
| Germany | Ministry for Research and Technology | ZEW and INFAS | Bonn, 18.1.94 | P.Hassenbach,G.Licht, D.Harhoff E.Nerlinger,J. Felder,M.Smid | Archibugi, Kristensen, Schäffer |
| Denmark | Statistics Denmark | | Copenhagen, 18.3.94 | H. Jørgensen, A. Slyngborg | Kristensen, Schäffer |
| Spain | National Statistical Office | | Madrid, 31.1.94 | C. Angulo Martin, M.A. Garcia | Archibugi, Kristensen, Schäffer |
| France | Statistical Office of the Ministry of Industry and Technology | Ministry of Higher Ed. and Research | Paris, 7.2.94 | J.-P. François, P. Temple, S. Lhuillery | Archibugi, Cohendet, Schäffer |
| Greece | Secretary General for Research and Technology, Ministry of Industry | | Athens, 1.4.94 | A. Hatziparadissis | Archibugi, Schäffer |
| Italy | National Statistical Office | ISRDS-CNR | Rome, 1.2.94 | A. Del Santo, A. Silvani | Archibugi, Kristensen, Schäffer |
| Ireland | Forfas, Science and Technology Agency | | Dublin, 7.4.94 | L. Harding | Kristensen, Schäffer |
| Luxembourg | Central Statistical Office | INSEE | Walferdange, 9.2.94 | A. Tibesar | Archibugi, Schäffer |
| The Netherlands | Ministry for Economic Affairs | SEO | Amsterdam, 19.1.94 | A. Kleinknecht | Archibugi, Cohendet Kristensen, Schäffer, |
| Portugal 1st survey | National Board for Scientific and Technological Research | | Lisbon, 15.12.93 | J.B. Bonfim | Archibugi, Cohendet |
| Portugal 2nd survey | National Board for Scientific and Technological Research | | Lisbon, 29/30.3.94 | J.B. Bonfim | Kristensen, Schäffer |
| United Kingdom CSO | Central Statistical Office | | London, 8.4.94 | L. Kay, J. Golland | Kristensen, Schäffer |
| United Kingdom CBI | Confederation of British Industry | | London, 8.4.94 | F. Steele, L. Luchlan, | Kristensen, Schäffer |
| Norway | Central Statistical Office | STEP Group | Oslo, 21.3.94 | F. Foyn, K. Smith | Kristensen, Schäffer |
| Commission of the European Union | Keith Smith, Consultant, D.G. XIII | | Luxembourg, 11.7.94 | K. Smith | Archibugi, Cohendet, Kristensen, Schäffer |
| Commission of the European Union | Eurostat, D.G. XIII | | Luxembourg, 26.5.94 | E. Deiaco, W. Grünewald, P. Nymand-Andersen, M. Doudeyns | Archibugi, Cohendet, Kristensen, Schäffer |

Other sources of information have also proven very useful and they include:

1. the documents produced by Eurostat and DGXIII for the national contractors and for the European Working Party on R&D and innovation;
2. the reports sent by the national contractors to Eurostat;
3. the preliminary statistical data produced;
4. other material handed out by the national contractors in the course of the interview;
5. additional written information required at the interviews.

When discrepancies emerged between the information provided during the interviews and evidence form other sources, we double-checked with the national contractors and with the EC functionaries. As it should be expected in a project which has involved several parties, we have found that the same issues were viewed differently by the national contractors on the one hand and the EC functionaries on the other. When important differences emerged, we have reported the views of both the parties and, if needed, our independent judgement was added.

The next section defines some concepts and terms we have used to assess the performance of the survey.

## 4.3 Definitions of the concepts used in the evaluation

To carry out our evaluation, a few terms of reference were needed. The 'Update of the Community Innovation Survey' mentions the terms "*reliability*, *validity* and *comparability*" to guide our evaluation (EUROSTAT, 1994). The definitions we have used for the first two terms, derived from Chadwick et al. (1984), are the following:

> "*Reliability* refers to consistency or stability of the measurement of a variable using a given operational definition."[3]
> "*Validity* is the degree to which an operational definition actually measured what it was supposed to measure."

The term *comparability* required a specific definition. We have distinguished three stages of comparability:

1. comparability of questionnaires;
2. comparability of data records;
3. comparability of aggregated data.

---

[3] We have also considered that the "Proposal for a Council Regulation (EC) on Community action in the field of statistics" has defined the term "reliability" as follows: "Community statistics must reflect as faithfully as possible the reality which they are designed to analyse. The users shall be informed of the sources, methods and procedures utilised" (see *Official Journal of the European Communities*, No C 106, 14.4.94, p. 25).

The *weak comparability* of two questionnaires is given if the questions asked in both questionnaires correspond to each other exactly.

The *strict comparability of two questionnaires* is given if:

1. the questionnaires are comparable in the weak sense;
2. the questionnaires are presented to the respondents in the same way (for example, by an interviewer or by mail);
3. the order of the questions is the same in both questionnaires;
4. no question exists in one of the questionnaires that is not contained in the other questionnaire.

The *comparability of two data records* can be assumed if:

1. the measurement of both data records are based on the same operational definitions, i.e. the questionnaires are comparable at least in the weak sense;
2. the measurements can be regarded to be "reliable" and "valid".

The comparability of data records is not a sufficient condition for the comparability of aggregated data. This is true, in particular, if the data from surveys are based on the principle "pars pro toto" (the parts on behalf of the total), because at least one of the following facts applies:

1. the frame of the survey is incomplete because it is restricted purposely to a part of the target population as defined in the Oslo Manual, or because there are urgent reasons to use an imperfect frame;
2. a sample from the frame has been drawn;
3. units (e.g. enterprises) that are included in the survey do not respond to the questionnaire completely or partially so the survey is impaired by unit non-response and/or item non-response.

The mass of units for which data are available is a part of the target population. The inference from the sample to the population cannot be safe and sound if it were not possible to state for any unit in the sample the number of units which it represents in the population. This number is called raising factor and it should be taken into account during the process of aggregating data to results. Analogue results (e.g. for one branch in different member countries) are comparable only if they have been evaluated according to this principle.

Therefore, the *weak comparability of aggregated data* requires that:

1. the data records are comparable among themselves;
2. for each data record the correct raising factor is stated;
3. these raising factors are taken duly into account during the aggregation of data records.

These postulates are necessary, but by not sufficient to reach full comparability of results. The assessment of the difference between two aggregated data (e.g. innovation costs of enterprises in the same branch in two countries) has to take into account both the random and the systematic errors of the results. A difference that is not bigger than its error must not be regarded to be significant.

Hence, the *strict comparability of aggregated data* demands that:

1. they are comparable in the weak sense;
2. margins of error are estimated and used for the assessment of results.

According to these considerations the demand for comparability of results based on a sample comprises substantially more than just the comparability of the questionnaires used for the survey or the comparability of data records.

The concepts defined in this section will be used in the next chapters. The problems related to implementation of the survey are discussed in Chapter 5. In particular, it will deal with the design of the survey. The statistical issues concerning sampling are discussed in Chapter 6. Finally, Chapter 7 will provide an overview of the this evaluation and it will summarise our recommendations.

# 5 Implementation of the CIS

## 5.1 Introduction

This chapter is devoted to a detailed assessment of how the CIS was implemented both at the European level and in member states. This assessment will be the basis for the discussion in Chapter 7 on "best practice" implementation.

There are two issues which are specific to the CIS. First, it is a survey which involves a large number of different organisations at both national, European and global levels: 13 national contractors (all the EU member countries and Norway), two services of the Commission (the SPRINT/EIMS programme of DG. XIII and Eurostat) as well as the OECD have been involved. Second, innovation statistics have a rather limited background in terms of systematic data collection. Before the CIS there was only one attempt to make internationally comparable innovation data (the survey in the Nordic countries already discussed in Chapter 2).

The surveys carried out in all member countries had as a common background the Oslo Manual and the EC. harmonised questionnaire prepared by Eurostat and OECD. However, substantial differences emerged in the design of the questionnaire, the procedures and the methodology of the survey. As already stressed in section 4.3, each of these aspects seriously affect the reliability, validity and comparability of the results of the survey.

This chapter will discuss the procedures and the methodology used in the implementation of the surveys. This chapter is not, however, only an evaluation of the survey implementation. To a large extent it is also a description of the different experiences carried out at the national and European level. The organisations involved in the CIS initiative have different expertise and this has led to the application of a variety of methods and approaches. Although this variety has often created problems for the comparability of the survey results, there are also very useful lessons to be drawn from the experiences from the national level. Therefore, we have tried to transmit to our readers some of the insights which we have gained during our field research. Innovation statistics is a new field and it is not surprising that some relevant experiences have not yet been disseminated among all the practitioners. This chapter is also an attempt to fill this gap by describing the experiences which could potentially be applied in Europe and elsewhere.

Scant attention has be paid to the formulation of national questionnaires in this chapter. The differences among the national questionnaires and the EC. harmonised questionnaire have already been described in detail by Eurostat (1994). This exercise has also classified each question for each country according to its level of *weak comparability* (as defined in section 4.3 of this report).

The next section is devoted to an assessment of the implementation at the European level. Sections 5.3 to 5.5 are devoted to the implementation at the national level: section 5.3 describes and comments on the national organisation of the survey. Section 5.4 reports on the effect of the questionnaire on the field. Was it clearly understood? Which problems were encountered? Section 5.5 assesses the participation in the survey. We have taken into account the factors which might have affected the quantity and the quality of the responses and their reliability. The perspectives of the CIS initiative are discussed in section 5.6, while sections 5.7 and 5.8 are, respectively, a summary and a list of recommendations.

## 5.2 Implementation at the EU level

Unlike the majority of statistical surveys carried out at the national level, the first CIS did not follow an ideal process in which a series of consecutive steps from (1) creation of a coherent conceptual framework to (2) creation of a harmonized questionnaire, (3) setting up of a harmonized and adequate survey design and (4) implementation of the survey in all EU countries simultaneously resulted in (5) creation of a data base with strictly comparable data (as defined in Chapter 4).

On the contrary, the CIS project was partly a self-driven process in the sense that once the project was initiated, many parties (such as the OECD, member states and even individual leading scholars in the field of innovation) were interested in participating and influencing the project. The growing interest in the project was itself an indicator of the importance and need of innovation statistics. But, on the other hand, since an increasing number of organisations were involved, it became more difficult to control and to co-ordinate the CIS project.

Furthermore, although it set was as a priority to reach comparable results within the European Union, it was equally important to obtain data which could potentially be compared with non-EU OECD countries. The Commission awaited international negotiations on a harmonised questionnaire, and thus the EU project was retarded. Some member states did not await these negotiations but, for various reasons, wanted to implement nationally developed innovation surveys before a common standard was agreed. Although in all cases these surveys were based on the Oslo manual and on draft versions of the harmonized questionnaire, this also increased the diversity.

This process entailed that the Commission never had the opportunity to undertake a co-ordinated project covering all EU countries but in some cases lagged behind in planning and co-ordination. However, some critical points on the Commission's role in this process should be

raised in order to allow for assessment of the implementation process and recommendations for future implementation of innovation surveys.

The set of recommendations which the Commission provided to the national contractors gave general recommendations on how the CIS should be implemented at the national level. There were three types of problem with this set of recommendations.

- Firstly, they were *general* i.e. they did not offer detailed information on the subjects covered. For example, it was recommended that the survey unit should be the enterprise but no detailed definition of enterprise was offered (see section 5.3.4 below).

- Secondly, the recommendations did not give advice on all the issues raised. For example the issues of sample frame, sample size and sample design were raised, but no recommendations were given.

- Thirdly, the recommendations covered only some of the relevant issues. For example no advice was offered regarding subsampling for non-response or on classification system to be used.

The Commission tried to compensate for these deficient recommendations through a series of working parties in which most of these issues were discussed and in some cases developed. It is our belief, however, that it would have been better to develop more detailed written recommendations, as written material usually has greater effect than non-committal discussions.

There was a substantive reason why the Commission was not keen to issue more detailed recommendations, namely that member countries were not bound to follow them. The Commission may issue *recommendations* but not give *instructions.* The Commission is not in a position to demand services from member states without a legal basis, and a legal basis does not exist for the collection of innovation data. The lack of legal basis implies that the Commission cannot impose a harmonised questionnaire to member states or select a specific contractor. Thus, in the case of the CIS, the Commission was dependent on member states' co-operation.

However, as will be discussed later in detail, we noted that on several occasions national contractors lacked the expertise on specific issues and that they would have substantially benefited from more detailed guidelines.

## 5.3 Implementation at the national level

### *5.3.1 The organisers of the survey*

National contractors were selected at the national level since the Commission does not have the authority to select the national contractors. In half of the cases, the contractors were the national statistical offices. Exceptions were Belgium, Germany, Greece, Ireland, the Netherlands,

Portugal and the United Kingdom. In these countries the surveys were organized at the Ministerial level or by other state or semi-state agencies. In Germany, the Netherlands and Belgium the Ministries sub-contracted parts or the whole task to academic research centres. A special case is represented by the United Kingdom. Although a feasibility study was carried out by the Central Statistical Office, the main survey was performed by the Confederation of British Industry.

When the survey was performed by, or in collaboration with, the central statistical offices, it was possible to use the standard directory of enterprises normally used for economic surveys. In other cases there were problems in acquiring the directory of statistical offices and the contractors were forced to rely on less satisfactory directories. This problem was particularly important in the United Kingdom and in the Netherlands.

Several contractors were already experienced organisers of R&D surveys. This is the case of Greece, Ireland, Italy, Norway, Portugal and Spain. In other countries, such as France, Italy, Luxembourg, the Netherlands, Portugal and the United Kingdom the contractors benefited from the experience gained from previous innovation or technology related surveys. The lack of previous experience on either R&D or technology surveys has proven to be a serious handicap in carrying out the CIS survey properly. However, the organisers of R&D surveys in some cases did not entirely perceive the differences between the R&D and the innovation surveys.

In the majority of countries, the survey was adequately financed once it was approved by the national authorities. A notable exception was Spain, where the survey was cancelled by a cut in the national budget a week after the questionnaires were mailed. Subcontractors in Germany and the Netherlands reported also to be under some financial stress. This has prevented them from carrying out additional tasks that emerged in the course of the survey (such as the analysis for non-respondents in Germany). In general, it seems that when the survey was carried out by, or in close co-operation with, government agencies, the resources were found also for unexpected tasks.

### *5.3.2 Co-operation with other institutions*

Innovation surveys require the combination of two different types of expertise: expertise on statistical issues and on innovation and industrial economics. At the European level, these two different types of expertise were supplied by the collaboration of EUROSTAT and the SPRINT/EIMS programme of DGXIII.

The problem of combining teams with experience in these two fields was also perceived at the national level. In several cases the contractors collaborated with other public or private research centres in the implementation of the survey. For example, the French Ministry of Industry performed the survey in close co-operation with the Ministry of Higher Education and Research. In Germany, the Ministry for Research and Technology subcontracted the task to an economic research institute (Zew) and to an institute specialised in sample surveys (Infas). In Greece, a team of four professors of economics was an integral part of the survey team.

In some countries (including France, Greece, Italy and Spain) the performers co-operated with *ad hoc* national working parties in the design of the questionnaire and in the organisation of the survey. The members of the working parties were chosen from the public administration, industry and academia. The working parties provided advise in order to clarify key concepts of the survey, such as the definition of innovation, product and process innovations, etc. They also had an important role in awakening the interest of industry and of other potential users for the results of the survey, which has indirectly helped to increase firms' participation. The drawback of external involvement was that these national working parties often pressured to modify the Eurostat questionnaire or to add up new questions.

The co-operation between the performers and representatives of public administration, industry and academia is essential for the management of the survey. The experience of *ad hoc* working parties has proven to be very useful and the whole CIS could benefit from the information and the feed-back provided by national working parties. However, it is essential that the activities of national working parties are co-ordinated at the European level.

The combination of know how in statistical and economic issues has proven to be very important to implement satisfactorily the survey.

### *5.3.3 The nature of the survey*

The methodology adopted by the national contractors to carry out the survey differ on three different aspects:

a) census/sample;

b) mandatory/voluntary;

c) direct interview/mailed questionnaire.

Table 5.1 (page 52) contains a synopsis of these aspects for the countries considered. Below these differences are considered in detail.

#### *Census/Sample*

In the majority of EU countries (the exceptions were Ireland, Italy and Luxembourg) the survey was carried out on a sample basis[1]. Many countries, including Belgium, Denmark, France, Germany and the Netherlands, have sampled all enterprises above a certain size. However, a common criterion was not agreed.

---

[1] A special case is represented by Spain, where the survey was planned to be on a census base before it was cancelled.

*Table 5.1 Methods of survey implementation*

| | Survey | Participation | Method |
|---|---|---|---|
| Belgium | Sample | Voluntary | Postal |
| Germany | Sample | Voluntary | Postal |
| Denmark | Sample | Voluntary | Postal |
| Spain | Census - cancelled | Mandatory/Voluntary | Postal |
| France | Sample | Voluntary | Postal |
| Greece | Sample of likely innovators | Voluntary | Interview |
| Italy | Census | Mandatory | Postal |
| Ireland | Census | Voluntary | Post. & Inter. |
| Luxembourg | Census | Voluntary | Interview |
| The Netherlands | Sample | Voluntary | Postal |
| Portugal | Sample of likely innovators | Mandatory | Postal |
| United Kingdom CSO | Feasibility | Voluntary | Postal |
| United Kingdom CBI | Sample | Voluntary | Postal |
| Norway | Sample | Voluntary | Postal |

Very different sampling methodologies have been applied (see Chapter 6 for a detailed discussion). In some countries the sample was not representative of the target population as defined in the Oslo Manual, even if leaving aside the service industries. This was the case for Greece and Portugal, where the contractors created a special sample aimed to reach all innovative enterprises of the country.

An international comparison among sample surveys is possible only if the sample can be related to the target population. However, we have found out that there was little awareness among the contractors of the statistical problems related to the international comparability of the CIS data. On the one hand, national contractors claimed that they decided autonomously on the procedures of sampling since they were not advised by the Commission to follow a common methodology. On the other hand, the Commission has argued that specific guidelines on

sampling were not released because national contractors were not bound (and possibly not willing) to apply them.

Several workshops were held at the European level involving both the Commission and the national contractors. However, these issues were not solved.

*Voluntary/Mandatory*

In all countries but Italy and Portugal the participation in the survey was voluntary. This reflects in part the national legislation, in part the fact that in several countries the survey was not implemented by the national statistical offices.

The voluntary nature of the survey has reduced the response rate in some countries, especially in the United Kingdom and Germany. In these countries, mandatory surveys generally have a substantially higher response rate. However, the voluntary nature of the survey is not the only factor affecting the response rate. Countries where the survey was voluntary, such as France and Luxembourg, reached a very high response rate. In the case of France, although the survey was voluntary, it helped that the organiser, the Statistical Office of the Ministry of Industry (SESSI), has the authority to undertake mandatory surveys.

*Postal/Interview*

In the majority of countries the survey was carried out by mail. Direct interviews were held in Greece and Luxembourg only. In both the countries the interviewers were trained before starting their field work. In Ireland the questionnaire was planned to be postal, but 400 direct interviews were also held at a subsequent stage to increase the response rate. The quality of the information obtained by direct interview is high, and item non-response is low. In a new initiative such as this survey, the interview approach is particularly helpful in order to explain the key concepts and the rationale of the exercise.

Direct interviews provide better documentation than mail questionnaires, but they require more resources. There is a trade-off between quantity of the enterprises monitored and quality of the responses, the exchange rate is however unknown.

### *5.3.4 Reporting units*

In any survey the units of analysis should be comparable. In the CIS, they should be comparable both within each individual survey and across different national surveys. For this purpose, the following definitions were used:

• *legally defined enterprise* is a unit which has a legal status in a given country. It might have one or several establishments, one or several business units. In several cases, it corresponds to the unit registered for tax purposes. According to this definition, establishments or business units located outside the borders of the nation should not be included.

• *economically defined enterprise* is classified according to the ownership or control. It includes all establishments or business units which are owned or controlled by the enterprise, located in the same or in a different country than the enterprise's headquarters. Often, large economically defined enterprises are subdivided, even within one country, into several legally defined enterprises.

• *business unit* is a part of the enterprise, although several enterprises are composed by a single business unit. A business unit may have one or more establishments.

• *establishment* is a geographically specific production unit. Several enterprises, especially among those of smaller size, have a single establishment only.

In the majority of countries the sample unit was the legally defined enterprise (see column 1 of table 5.2). This is in line with the standard Eurostat definition. In some cases, however, a

*Table 5.2 Coverage of CIS surveys*

| | Sample Unit | Coverage | Cut-off point* | Census above* | Pre-test |
|---|---|---|---|---|---|
| Belgium | Establishment | Manufacturing | > 9 | > 199 | No |
| Germany | Enterprise. Large firms subdivided | Manuf. and services | > 4 | > 199 | 50 interviews |
| Denmark | Legally defined enterprise | Manufacturing | > 19 | > 199 | Only the 1st Eurostat |
| Spain | Legally defined enterprise | Manufacturing | > 19 | > 19 | No |
| France | Legally defined enterprise | Manufacturing | > 19 | > 1000 | 30 interviews |
| Greece | Legally defined enterprise | Manuf. and services | No | N.A. | 7 interviews |
| Italy | Legally defined enterprise | Manuf. and services | > 19 | > 19 | 25 int. plus 15 mailed |
| Ireland | Legally defined enterprise | Manufacturing | > 9 | | 44 int. plus 44 mailed |
| Luxembourg | Legally defined enterprise | Manufacturing | No | > 0 | No |
| The Netherlands | Principal establishment | Manuf. and services | > 9 | > 99 man. > 199 ser. | 8 interviews |
| Portugal | Enterprise | Manuf. & part of serv. | No | | No |
| United Kingdom CSO | Holding company | Manufacturing | N.A. | N.A. | 44 mailed questionn. |
| United Kingdom CBI | Autonomous enterprise | Manuf. and services | > 25 | | Only the 1st Eurostat |
| Norway | Legally defined enterprise | Manufacturing | > 5 | > 99 | Only the 1st Eurostat |

* Number of employees.

different statistical unit was used, especially in those countries which could not use the directories of the national statistical offices.

In Belgium the survey was directed to establishments rather than enterprises. In Germany a very few large enterprises were subdivided into business units. In the Netherlands, the survey was carried out at the principal establishment level, but some participants responded centrally for all the establishments of the same economically defined enterprise. Several contractors reported cases in which the responses were not coherent: in parts of the questionnaire the respondents referred to the establishment, in others to the economically defined enterprises.

The entity *legally defined enterprise* is substantially different in small and large countries. In fact, the share of enterprises which are controlled by foreign companies is greater in small countries than in large countries. This implies that in small countries the legally defined enterprise is more likely to correspond to the establishment. This case is particularly evident in the smallest country of the European Union, Luxembourg where half of the enterprises monitored were found to be owned by foreign companies.

The CIS database does not allow to consolidate establishments and business units into their economically defined enterprises. This is due to, at least, two different reasons. First, although in the questionnaire respondents are asked to state if the reporting unit belongs to a group (i.e. to an economically defined enterprise), they are not asked to identify it (see Section I of the questionnaire). Second, since in the majority of countries the survey was on a sample base, parent enterprises of the sampled units are not necessarily included in the sample.

The lack of a identical unit of analysis is quite common in European industrial data and cannot be easily solved. However, it should be borne in mind that the lack of homogeneity in the unit of analysis across countries makes it problematic to use these data to address issues connected to the industrial organisation of the European industry such as market structure, concentration, and innovation. It will not be possible to consolidate establishments, business units or legally defined enterprises into economically defined enterprises, in spite of the fact that the questionnaire reports information on the ownership of the responding enterprise.

### *5.3.5 Coverage of the survey*

#### *Coverage of Manufacturing and Services*

As stated in section 2.2, one of the main strengths of innovation surveys is that they can provide information also on the service sector (which are not easily provided by other technological indicators such as patenting or trade in high-tech products). So far, the role of services in technological change has been underestimated, also for lack of empirical evidence. Quite rightly, the Oslo Manual has included the service industries into the target population of innovation surveys.

However, Eurostat and DGXIII did not make recommendations to include the service sector in the survey, leaving the choice to the national contractors. In four countries only (Germany, Greece, Italy and the Netherlands) national contractors have included service industries in the sample. In Portugal, some enterprises in services closely related to the manufacturing sector were also included. It is significant that in both Germany and the Netherlands the response rate of enterprises in services was equal or even higher than enterprises in manufacturing.

The results obtained for the service sector in Germany, Greece, Italy, the Netherlands are the first large scale attempt to explore innovations in services by direct survey. As already suggested in the Oslo Manual, we need to cover the service sector, and the CIS provides a unique opportunity to study the feasibility of innovation surveys in services.

*Cut-Off Point*

In several countries enterprises with less then 20 employees were cut-off from the survey (see column 3 of table 5.2). Some small countries, such as Belgium, Ireland and the Netherlands, applied a cut-off point at the level of 10 employees. Luxembourg has not applied any cut-off point. The issue does not apply to Greece and Portugal since, as stated above, the selection of a specific target population would have included any likely innovating enterprise.

Special cases are represented by the Norwegian and the German surveys, with a very low cut-off point of 5 and 4 employees respectively. In the German case, enterprises with less then 20 employees have returned the questionnaire slightly less than the average.

### *5.3.6 Classifications available*

The possibility to compare the survey results across countries heavily depends on the comparability of the classifications used by the national contractors. The three main classifications used in the CIS concerns industry, size, and region. Table 5.3 (page 57) reports a synopsis of the classifications used by national contractors.

Classifications can be considered at two different levels. The first is the level of the classification used for the purpose of sampling, the second the detail of the information which is available in the database. The classification used for the purpose of sampling needs to include a number of units in each stratum to allow to draw a statistically significant sample and therefore cannot be too detailed. On the contrary, there is no limit to the level of disaggregation of the classification collected with the questionnaire and which is reported in the database. In fact, some variables, such as employment, turnover or value-added, can be recorded on a continuous scale. We will consider here the latter aspect, while the former will be discussed in Chapter 6.

*Table 5.3 Classifications in the CIS database*

| | Industrial classification | Size classification | Regional classification |
|---|---|---|---|
| Belgium | NACE 70 - 4 digit | 9 employment classes | Regional |
| Germany | NACE rev. 1 - 3 digit | Number of employees | East/West, Nuts 3 |
| Denmark | 4 digit | Number of employees | Nuts 3 |
| Spain | NACE rev. 1 - 2 digit | Nr. of Employees & sales | Nuts 2 |
| France | Nap(trans.Nace 70-4 dig.) | Number of employees | No |
| Greece | NACE rev. 1 - 2 digit | Number of employees | Nuts 2 |
| Italy | NACE rev. 1 - 4 digit | Number of employees | Nuts 2 |
| Ireland | NACE rev. 1 - 4 digit | Number of employees | Nuts 2 |
| Luxembourg | NACE 70 - 4 digit | Number of employees | Regional (post codes) |
| The Netherlands | NACE rev. 1 - 3 digit | 6 employment classes | Nuts2 |
| Portugal | National - 3 digit | Number of employees | Yes. Level? |
| United Kingdom CSO | N.A. | N.A. | N.A. |
| United Kingdom CBI | NACE rev. 1 - 2-3 digit | Number of employees | Nuts 2 |
| Norway | ISIC, rev. 2 | Number of employees | Yes, in principle |

*Industrial classifications*

The majority of countries have applied the standard NACE classification or national classifications which are compatible with NACE. Norway has used the ISIC classification which, with some effort, can be translated into NACE, and Belgium, France and Luxembourg have used the NACE '70. The level of disaggregation available in the NACE version, however, is not the same: Spain and Greece have used the 2 digit level. This classification does not seem disaggregated enough to capture some fundamental inter-industry differences in innovative behaviour. For example, it merges together Chemicals and Pharmaceuticals. More generally, this will make it problematic to use these data for detailed industry case studies.

The Dutch survey found out that in 5% of the cases the real sector of activity of the enterprises monitored was different from the one recorded in the sampling directory, also because it was not sufficiently updated. This problem illustrates the need to use, whenever is possible, the directories of the national statistical offices.

The OECD has proposed to use a 2 digit level classification as a minimum common standard (see Akerblom, 1994). The level of disaggregation which was achieved in the majority of European countries was more detailed.

*Size of the enterprises*

In all countries the size of the enterprises has been measured according to employment. In all countries but Belgium and the Netherlands the measure is metric. In these two countries, data are available by 9 and 6 employment classes respectively. A common standard for the tabulation of the results according to employment classes has also been proposed by the OECD (Akerblom, 1994). But it is certainly an advantage to have information on the number of employees for each reporting unit.

It should be stressed that the possibility to make cross-country comparisons of firm size does not depend only on the measure of size selected, but also on the unit of analysis used to define enterprises. As stated above (section 5.3.4), national surveys have not adopted a unique definition of enterprise.

*Region*

Not all countries have used the lowest level of regional disaggregation, i.e. Nuts 3. The level of regional disaggregation available also depends on the definition of the unit of analysis. The official address of legally defined enterprises tend to be located in or near the main economic centres whereas establishments are located in low-cost areas.

## 5.4 The questionnaire in the field

Chapter 3 has already discussed the main architecture of the questionnaire in the context of the Oslo Manual. But a questionnaire should also be evaluated according to its action on the field. In this section the comments received by the contractors (and by the respondents) on the design of the questionnaire are discussed. We will also explore the reasons why some national contractors have decided to modify the harmonised questionnaire.

### *5.4.1 Pre-test of the questionnaire*

The pre-test has proven to be a very valuable step to improve the design of the questionnaire and to improve the quality of the whole survey. A first pre-test in five European countries (Great Britain, Norway, Denmark, Germany and the Netherlands) was sponsored by Eurostat

(for a synthesis of the results, see Kleinknecht, 1993). However, the conclusions of this pre-test are not always consistent with what emerged during the CIS survey, possibly because the questionnaire was further modified after it. The most notable case is represented by the question on the life cycle of the product, which appeared to be well understood by respondents during this pre-test while it was one of the questions with the highest item non-response in the real survey.

Some countries which did not modify the harmonised questionnaire (Belgium and Luxembourg) have not done any pre-test at the country level. Other countries which were involved in the European pre-test, such as Denmark and Norway, did not engage in additional pre-tests. Very detailed pre-tests were carried out in France, Germany, Ireland and Italy (see column 4 of table 5.2). In some cases these pre-tests dramatically changed the design of the questionnaire. The most remarkable case is France; where the organiser of the survey (SESSI) drastically simplified the harmonised questionnaire and dropped all controversial questions.

The lessons to be gained by the pre-tests carried out by national contractors proved to be very useful, even because different approaches were followed in each country. On the one hand, the pre-test provided some valuable insights on the understanding of questions. On the other, it gave to the national performers some ideas on the most appropriate method to perform the survey. Unfortunately, these lessons were kept at the national level, and often the Commission was not even informed of the lessons gained. Several countries have declared that in future ventures they would like to extend pre-testing before the launch the survey.

As an example, the results of the pre-test carried out in Italy are reported in Table 5.4 (page 60). Respondents were asked to assess, for each part of the questionnaire:

i. its level of comprehension;
ii. the precision of the answer provided;
iii. the time required to answer.

Overall the Italian pre-test indicated that the comprehension of the questionnaire was good, that the data were considered to be reasonably accurate, and that it was not too time consuming for the respondents to provide answers. It should be noted that Italy might represent a special case since the CIS is the third main innovation survey performed in the country and the majority of enterprises had already participated in former surveys. Moreover, these results were to some degree contradicted by the results of the real survey: several respondents required help from the statistical office to understand the meaning of some questions.

The pre-test showed that the questions on innovation costs and on the impact of the innovations gave major problems to the Italian respondents. Although the majority of respondents comprehended these questions, only 18% and 26%, respectively, of the respondents declared that the data were easily available. Not surprisingly, these were also the questions with the highest item non-response.

*Table 5.4 Pre-test of the questionnaire. Results from the Italin pre-test. Percent*

| Parts of the questionnaire | Satisfactory | Average | Problematic | Total |
|---|---|---|---|---|
| General information | | | | |
| a) comprehension | 86 | 14 | 0 | 100 |
| b) accuracy | 86 | 14 | 0 | 100 |
| c) time required | 78 | 22 | 0 | 100 |
| Sources of information | | | | |
| a) comprehension | 77 | 23 | 0 | 100 |
| b) accuracy | 53 | 40 | 7 | 100 |
| c) time required | 53 | 43 | 4 | 100 |
| Objects of the innovation | | | | |
| a) comprehension | 70 | 30 | 0 | 100 |
| b) accuracy | 53 | 43 | 4 | 100 |
| c) time required | 60 | 33 | 7 | 100 |
| Technology transfer | | | | |
| a) comprehension | 68 | 25 | 7 | 100 |
| b) accuracy | 54 | 39 | 7 | 100 |
| c) time required | 54 | 39 | 7 | 100 |
| R&D activities | | | | |
| a) comprehension | 76 | 24 | 0 | 100 |
| b) accuracy | 48 | 45 | 7 | 100 |
| c) time required | 41 | 48 | 11 | 100 |
| Innovation costs | | | | |
| a) comprehension | 55 | 35 | 10 | 100 |
| b) accuracy | 21 | 41 | 38 | 100 |
| c) time required | 18 | 53 | 29 | 100 |
| Impact of the innovations | | | | |
| a) comprehension | 57 | 21 | 22 | 100 |
| b) accuracy | 21 | 54 | 25 | 100 |
| c) time required | 26 | 48 | 26 | 100 |
| Sectors of use of innovations | | | | |
| a) comprehension | 69 | 17 | 14 | 100 |
| b) accuracy | 79 | 14 | 7 | 100 |
| c) time required | 78 | 22 | 0 | 100 |
| Role of public intervention | | | | |
| a) comprehension | 82 | 14 | 4 | 100 |
| b) accuracy | 54 | 39 | 7 | 100 |
| c) time required | 50 | 46 | 4 | 100 |
| Factors hampering innovation | | | | |
| a) comprehension | 76 | 21 | 3 | 100 |
| b) accuracy | 65 | 32 | 3 | 100 |
| c) time required | 61 | 39 | 0 | 100 |

Also the real questionnaire in the Netherlands asked respondents to specify if the answers provided where "rough" or "accurate" estimates, or if "no answer was possible". It would be useful to study these results since they refer to a much larger number of observations than those of the Italian pre-test. The information provided by Kleinknecht indicates that also in the Dutch survey the accuracy of the answers provided to the questions on innovation costs and sales according to the stage of the life cycle is lower than that of the answers to the other questions.

Overall, it was our impression that the CIS survey could have benefited more form the pre-tests carried out at both the European and national levels.

### *5.4.2 Comments on the questionnaire*

The majority of the contractors believed that the key concepts of the survey, and most notably R&D and innovation, were well defined, although almost all of them stated that the clarity of the definitions could be improved. Some problems of comprehension emerged in traditional sectors. Often innovation was not understood as "technological", but simply as a change in products. In Belgium, Italy and Spain a good number of enterprises in clothing and shoemaking asked if a fashion innovation should have been included. The same problem applies to the word "design": it was not always clear to enterprises in traditional sectors that the survey deals with *industrial* rather than with fashion design. On a lower scale, not even the concept of R&D was always understood by respondents in traditional industries. These key terms might be better defined by practical examples than by theoretical definitions.

In several countries confusion was generated by the fact that some parts of the questionnaire refer to one year only (1992), others to three years (1990-1992). Finally, it was not always clear to respondents if innovation costs afforded for failed projects should also be included.

It was sometimes stated that too many and similar concepts were used in the questionnaire. Examples mentioned include:

- process innovation overlaps with investment;
- process innovation overlaps with innovation engineering;
- product innovation overlaps with product differentiation.

*The definition of innovation costs*

Innovation costs is one of the basic, possibly the basic, indicators which should come from innovation surveys (see Smith, 1992; Hansen, 1992). However, as already shown by a number of pre-tests, this question presented conceptual as well as statistical difficulties. It was remarked that the Oslo Manual has too little detail on what should, or should not be included in innovation costs. In comparison, the Frascati Manual (OECD, 1981), has a much larger section on the components of R&D expenditure.

It was not clear to the contractors how to deal with process innovation costs. Some assumed that it should be included since, according to the Frascati Manual (OECD, 1981), expenditure for R&D equipment is included into R&D expenditure. Others suggested that a good proxy measure for embodied process innovations is already available as investment in machinery and equipment. Therefore, process innovation costs should be restricted to expenditures for intangible activities. A similar problem was raised concerning investment in machinery and equipment related to product innovation, should it be included in innovation costs? Others assumed that product innovation costs should be understood as the "immaterial" investment.

These issues were often interpreted differently by national contractors, and sometimes this led to alterations in the formulation of the questions and/or in the instructions provided to the respondents. This creates some problems to compare innovation costs across countries. The definition of innovation costs needs to be clarified and further standardised in the next round of innovation surveys as well as in future version of the Oslo manual.

*The relationship between R&D and innovation*

Respondents also found the relationship between R&D and innovation problematic. Some contractors remarked that the definition of R&D provided in the questionnaire is too narrow and does not correspond to the latest version of the Frascati Manual.

In the harmonised questionnaire, respondents are asked twice about their R&D expenditure: in the section on R&D (question 10) and in the section on innovation costs (question 13). However, it was noted that often the two figures were not consistent. Several respondents have also reported no innovation costs in spite of the fact that they reported some R&D expenditure.

### *5.4.3 National differences in the questionnaire*

A detailed analysis of the comparability across countries of individual questions has already been performed by Eurostat (see Eurostat, 1994). This evaluation has focused on why the original harmonised questionnaire was modified.

All contractors declared that the harmonised questionnaire was perceived to be too long by the respondents. In spite of this, all contractors but the French added up new questions or kept the harmonised questionnaire unchanged (Belgium and Luxembourg).

In some cases differences in the questionnaires were due to the timing. In Greece and Spain a questionnaire was planned before the latest version of the harmonised questionnaire was available. In Italy a few questions were added up to secure comparability with previous innovation surveys. Quite a few other minor modifications were introduced in several countries for a variety of reasons. Sometimes, background information on the enterprise was collected (this was particularly the case for the CBI survey in the United Kingdom). In particular, if the contractors were not the national statistical offices, it was often needed to collect background information on the enterprise since it was not available from other sources. Detailed

descriptions of the questionnaire modifications introduced in each survey are reported in the country reports (Annex 1).

It was not perceived that even small changes in the formulation of the questionnaire, such as for example, to add up or to delete an option in a multiple choice question, reduce the comparability (as defined in section 4.3) of the questionnaire.

Some of the national differences were due to the different timing in which the questionnaire was prepared, discussed, pre-tested and modified across countries.

*Translation of the questionnaire*

No major problems have emerged in the translation of the questionnaire. In Germany, Norway and Denmark it was not clear how "incremental innovation" should be translated but the same problem emerged also in Ireland. Several contractors had problems with the terms "mother", "sister" and "daughter" enterprise but this seems to be a conceptual rather than a linguistic problem since the terms were unclear also in the United Kingdom and Ireland.

## 5.5 The participation

Any survey relies on its ability to involve the respondents. Actions should be taken in order:

i. to secure a high response rate;
ii. to identify the appropriate respondent;
iii. to secure reliable answers.

These issues are more important in mailed questionnaires since there is no interviewer to provide classifications to the interviewee. They are also important for voluntary surveys since the respondent may or may not be willing to participate and can decide how much effort is needed to fill in the questionnaire. As seen, the majority of the national surveys within the CIS are both mailed and voluntary surveys.

### *5.5.1 The response rate*

While the statistical issues connected with the response rate are discussed in Chapter 6, some problems related to the way in which the implementation of the survey has affected the participation of industry are discussed in this section.

The golden rule to secure a high response rate from voluntary and mailed questionnaires is to make it as simple as possible and the CIS experience has confirmed it. The questionnaire was perceived to be too long in all countries and, according to the contractors, this is the main reason why the response rate in some cases was unsatisfactory. This had implications not only for the number of returned questionnaires but also for their reliability. Respondents were often

pressed in time to fill in the last parts or to provide accurate answers to the questions which required more time and thought.

The questionnaire used in the French survey was much simpler and shorter than the harmonised questionnaire for two reasons. First, the statistical services in France do not allow the duplicate asking of questions to the same respondent (but it is very easy to recover the information from other surveys by the enterprise's identification number). This implied that all questions related to economic and R&D activities were dropped. Second, the French organisers also decided to drop all questions which the pre-test proved to be too difficult for the respondents (including innovation costs).[2] As a result, the French questionnaire was about half the length of EC one. This is probably the main reason for the very high response rate achieved.

The contractors used several devices to induce the respondents to participate in the survey. Some contractors have attached a supporting letter. In the Netherlands, a covering letter of the Minister for Economic Affairs was attached to the questionnaire and to all reminders. In Greece, the Secretary General for Research and Development, who is also responsible for the incentives to industrial innovation, sent a covering letter. In the Flemish part of Belgium the President of the region signed a supporting letter.

The collaboration of industrial association has also proven to be useful to increase the response rate in some countries. In Norway and in the Netherlands the press awoke the interest of potential respondents by reporting information on the start off of the survey. Professional newsletters and the journals of industrial associations are suitable vehicles to alert the public about the survey.

Another method used to increase the response rate was to send several reminders. This has proven to be effective in the Netherlands and in France, where the sampled units received three reminders along with a new copy of the questionnaire. It has proven to be less effective in Germany, where as many as 5 reminders were forwarded to the sampled units: the last three reminders together increased the response rate of 3% only.

In several cases it has helped to combine written reminders with telephone calls as in Italy and Germany. The involvement of regional contact organisations, such as in Italy and in Greece, has also proven to be useful. Personal contacts between the organisers of the survey and the respondents have also shown to be very fruitful. This was particularly the case of small countries such as Norway, Greece, Portugal and Luxembourg.

In some countries, the questionnaire was customised. This is the case of Germany, where three different questionnaires were used: one for large companies, another for small companies and a

---

[2] A substantial price was however paid to simplify the questionnaire: several of the most relevant issues cannot be addressed by the French survey. Therefore, the possibility to compare France to the other European countries is reduced to some questions only.

third for the service sector. In the Netherlands and in Italy a special questionnaire was prepared for the service sector. Although there are some advantages to customise questionnaires, it is important to secure comparability among responses. At this stage, advantages and disadvantages cannot be assessed.

Some contractors have remarked that in order to obtain the participation of industrialists it is important to give them a synthesis of the results of the survey; if they see the usefulness of the information they provide they will be more willing to participate. The German pre-test has indicated that the business community is particularly interested in industry-specific studies. In France, a synthesis of the results of the 1991 innovation survey was attached to the new questionnaire.

The relative high response rate in countries which have performed the survey also in the past (such as France, Italy, Luxembourg and the Netherlands) might also suggest that the response rate increases when the survey is repeated periodically. There are two combined "learning effects": on the one hand, respondents get familiar with the content of the survey and find it easier to answer the questionnaire and to provide accurate answers. On the other, the organisers improve the methods to increase the response rate.

We have encountered very different approaches to secure high response rates. A large number of devices have been applied at the national levels. While some of them might be transferable across countries, others are typically nation-specific, institution-specific or even person-specific. We believe that if an identical method will be applied across all the European countries, the outcome would be very different national response rates. Inasmuch as the response rate is concerned, the most sensible advise is to exploit the experience of survey organisers, provided that a response rate which would allow international comparisons is obtained in all countries. Informal information on the experiences carried out in other countries can however have a very useful effect, especially if some organisers have not previous experience on statistical surveys. Although to apply a variety of methods across countries is a must rather than a choice, the systems applied in each country should be agreed with the Commission.

*Competing surveys*

One factor which could reduce the response rate is an excessive number of questionnaires received by enterprises. This was noted by several contractors. In Ireland, the respondents reported that sometimes they receive as many as 12 questionnaires per week. One enterprise declared that it received 25 questionnaires in a week. Besides the usual complaints by respondents on the "survey load", specific competing surveys have not been identified. The most notable exception is represented by R&D surveys, which should often be compiled by the same person as the innovation survey.

### *5.5.2 The respondent(s)*

To identify the appropriate person within the enterprise is a crucial task to secure a high response rate and reliable answers, in large companies especially. Several contractors have addressed, when possible, the questionnaire to a competent person. In the United Kingdom the questionnaire was sent to a pre-identified managing director. In Germany this was done for all large enterprises. Contractors have stated that when they were able to identify the right person in the enterprise it was more likely that the questionnaire was returned and filled in properly. Contractors in Belgium and Denmark have stated that in future surveys they will expand the activities carried out before the questionnaire is mailed in order to pre-identify the appropriate respondent within the enterprise.

The typical respondent in small and medium sized enterprises was the general manager (or the owner). In larger companies the respondent was either the R&D manager, the chief engineer or the innovation manager. One might fear that the large number of R&D managers which have filled in the questionnaire has biased the results in favour of the significance of R&D in innovation.

It is not entirely clear which manager is the most appropriate to fill the questionnaire in large firms. The German contractors have stated that the best option is to identify, and forward the questionnaire to, the manager who has the authority to make the other staff fill in the questionnaire. The Central Statistical Office in the United Kingdom has suggested dividing the questionnaire into two parts: a first part with all the numerical questions to be addressed to the financial division of the company and a second part on the attitudes towards innovation, and which should be complied by the managing or marketing director.

In large enterprises the questionnaire has often circulated among several persons. Of course, the larger the enterprise, the more it is needed to acquire information from different departments. Typically, the accountancy office would fill in the quantitative questions and the R&D or technical departments the qualitative questions. The pre-test carried out in Germany has indicated that an average of 2-3 persons collaborated to fill in the questionnaire.

*Time required to fill in the questionnaire*

The direct interviews in Greece and Luxembourg also provided very important feedback on the time needed to fill the questionnaire. In both the countries the minimum time required was 30 minutes, and the maximum 90 minutes. The pre-test carried out in Germany (see Felder et al., 1993) indicate that the average length of the oral interviews was 105 minutes, with a variability from 30 to 210 minutes. The estimated time needed to fill in the written questionnaire increases to approximately 120 minutes, with a range from 60 to 210 minutes. A specific question on the time needed to fill the questionnaire has been added up in the CBI survey in the United Kingdom.

### *5.5.3 Securing reliable answers*

A survey should not be judged by the response rate only, but also by the quality of the responses received. In a new initiative such as this innovation survey it should be expected that respondents have queries about the appropriate way of filling the questionnaire. It should also be expected that questionnaires are returned with several questions unanswered and that some answers are inconsistent or contradictory. Therefore, national contractors should provide all the possible help to the respondents to facilitate their task, especially when the questionnaire is mailed. Also, they should make sure they will be able to get in touch with the respondents after the questionnaire is returned, if needed. These activities should be planned and budgeted at the very beginning of the survey.

Some contractors, such as the Portuguese, have attached to the questionnaire some guidelines for filling it in and with the definitions of the terms used in single questions. Shorter instructions were already available in the questionnaire, but it has been noted that there is little point in extending the instructions on the questionnaire because the majority of the respondents do not read them if they are too long or too complicated. But it is likely that some respondents will make use of additional guidelines, especially if relevant examples are added up.

A free-call help line was established in Greece and Italy to provide assistance to fill in the questionnaire. In Italy, about 1,000 respondents used it.

Checks of plausibility and compatibility of the returned questionnaires have been made in all countries. The guidelines provided by Eurostat (see Eurostat, 1993) have been very helpful and the majority of the contractors have used them. In a few cases contractors were in the position to check some of the data provided by the enterprises with data coming from other surveys. In Norway the information obtained by the innovation survey was systematically checked with data from R&D and accounting surveys.

Both item non-response and inconsistent answers were high in several countries (for a quantification, see Chapter 6). Telephone follow-ups have been used in several countries to reduce item non-response or to clarify the answers provided. In Portugal and Norway a very large number of enterprises were contacted after the questionnaire was returned. In the Netherlands, about 30 to 40% of the respondents had to be contacted again. Telephone follow-ups are useful if it is possible to contact the compiler of the questionnaire; the returned questionnaires should therefore provide information on the name, address and telephone number of the compiler.

### *5.5.4 Reliability of responses*

The majority of the contractors consider the quality of the responses obtained either standard or good. In other words, no contractor believed that the quality of the responses provided by the respondents is particularly bad. However, this often required a lot of follow-up work to contact the respondents to fill the missing parts or to get classifications.

It was our impression that the reliability of the results in countries which had already conducted innovation surveys, as in France, Italy, the Netherlands and Luxembourg, are better. As already mentioned (and developed in section 7.6.3), there is a very important "learning effect" to improve the reliability of the data it is therefore necessary to persist to perform surveys on a regular or cyclical basis.

In a number of countries, including France, Italy and the Netherlands, a few enterprises returned two copies of the questionnaire filled by different persons independently. This happened when a reminder was sent and at the same time a completed questionnaire was returned. These "double questionnaires" sometimes show significant differences. A systematic analyses of them, however, has not yet been made.

### *5.5.5 The participation of non-innovating enterprises*

In several countries, the contractors had the impression that non-innovating enterprises had a lower propensity to participate in the survey. They were less willing to participate to the survey either because they were not familiar with the concepts of "technological innovation" and "R&D" or because they assumed that their answers were not relevant for the survey. This was one of the conclusions reached by the pre-test carried out in Germany.

The only possible way to test this hypothesis is to check if non-respondent enterprises are significantly different from respondents. The statistical analysis for non-respondents done in Ireland and in the Netherlands, as well as those which are in progress elsewhere, needs to be carefully explored.

Even if we cannot provide any firm conclusion on a different behaviour of innovating and non-innovating enterprises towards the survey, actions should be taken to ensure the participation of both groups of enterprises. The section on the factors hampering innovation concerns non-innovating enterprises as much as innovating enterprises. More importantly, a lower participation of non-innovating enterprises impede the possibility of relating the sample to the population and therefore to draw conclusions for the entire economy.

A method to ensure that also non-innovating enterprises participate in the survey is to split the survey into two stages, as already experienced by the first Italian innovation survey. The first stage consisted of a very short questionnaire to be sent to all enterprises asking a few basic questions on their innovations. A longer questionnaire was sent to innovating enterprises only. But this method also created some problems since there was a substantial drop of responses between the first and the second questionnaire.

To secure the participation of non-innovation firms, the German survey has modified its title to "Future perspectives for German industry". This specific title might be problematic since respondents would find that innovation is the central argument of the survey. The CBI survey

in the United Kingdom has removed the filter question in section 1 to prevent a drop of interest of non-innovating enterprises.[3] In fact, in several countries (and particularly in Norway) enterprises which assumed to be non-innovating "discovered" that they had somehow innovated after a telephone conversation with the interviewers.

Several contractors have stated that non-innovating enterprises should be asked to fill in also the section on technology transfer.

### *5.5.6 Confidentiality*

In the evaluation we also tried to assess to what extent enterprises were concerned about disclosing confidential information and if this might have reduced their willingness to participate to the survey. Technological innovation is one of the key aspects of strategic management and we expected that respondents might be concerned to disclose information about it. None of the contractors attributed the lack of collaboration from respondents to the confidential nature of the information required. Statistical offices may have an advantage over other institutions since their reputation on statistical secrecy is widely known.

Confidentiality might affect the response rate in two different ways, total non response or item non response to the most sensitive questions. It is not easy to quantify the number of enterprises which did not return the questionnaire because of confidentiality; if respondents do not want to disclose specific information it is possible that they do not return the questionnaire at all.

The Greek questionnaire asked respondents if they were willing to disclose the information provided on their own enterprise to persons other than the organiser of the survey. The results are quite interesting (see table 5.5). While 39.6% of the respondents did not want to disclose the information at all, as many as 25.6% were willing to make the results of the questionnaire publicly available. In the middle, 34.8% of the participants were willing to disclose information to researchers and academics but not to other enterprises.

We have also invited the contractors to mention the questions which, according to their experience, are more likely to affect confidentiality. It was pointed out that they correspond to questions with high item non-response. It is not easy, however, to detect if enterprises do not respond to some specific questions because they are confidential or because responding was difficult and would take too much time. Sometimes respondents may say that some questions are confidential because they are too difficult to answer, others that it is impossible to answer because they wish to keep the information secret.

---

[3] Which, in turn, might have created another problem for the comparability of the data.

*Table 5.5* *Confidentiality of the innovation survey. Results from the Greek survey.*

| Options about confidentiality | Number | Percent |
|---|---|---|
| Enterprises which did not want to disclose information to any other party than the performer of the survey | 121 | 39,6 |
| Enterprises which allowed to disclose information to researchers and academicians in the field of innovation | 106 | 34,8 |
| Enterprises which allowed to disclose information to any interested person or company | 78 | 25,6 |
| Total | 305 | 100,0 |

In some cases respondents in Belgium, Germany, Greece, Ireland and Luxembourg perceived that the questions on innovation costs were confidential. The Italian pre-test, however, shows that it was also the most time consuming question. Other questions, such as share of new products, impact of innovation and product life cycle, were also perceived as confidential by some contractors.

No clear indication emerged on the sectors more concerned about confidentiality. Only one contractor (Norway) mentioned high technology enterprises in process industries as the most concerned about confidentiality.

## 5.6 Perspectives

In the course of the interviews with the national contractors, we also asked their opinion on the perspectives of innovation surveys. In particular, we asked about the relevance, harmonisation and periodicity of this exercise.

### *5.6.1 Relevance*

All contractors believed that the development of innovation indicators is very relevant. In countries where the survey was performed by, or in close collaboration with, ministries and government agencies, it was stated that the data obtained will provide relevant information for technology and industrial policy as well as for evaluation of public intervention. The Italian contractor stated that it is crucial to enlarge the number of participating sectors into the services.

### *5.6.2 Harmonisation*

All contractors expressed the view that the survey should be harmonised. Some countries, including Italy and Norway, pointed out that their interest on the survey depends heavily on its international comparability. In other words, they are interested in performing it only if international comparability is secured. However, as we noted earlier, the actions of the national contractors were not always consistent with this wish.

Several countries urged the establishment of harmonisation on sampling, collection of data and reference periods by means of Community guidelines. A few contractors stated that the harmonised part of the questionnaire should be reduced to keep it identical in all countries. The emphasis on harmonisation expressed by the majority of the contractors is somehow contradicted by the changes they introduced into the questionnaires. From the positive side, however, we have recorded the contractors' willingness to increase the co-ordination and the harmonisation in the design, implementation and sampling of future innovation surveys.

We have also recorded that several countries are planning to use these data for panel analysis. This will put them under two opposite pressures in future surveys. On the one hand, to keep the survey as similar as possible to the previous one to allow time-series comparisons within one country. On the other hand, it will make it more compatible with the Eurostat standard to allow international comparisons. Although both the issues are relevant, we believe that it is essential first to apply a common international standard to innovation surveys and second build up panel data. Innovation surveys are still at their infancy and it is crucial to build first a robust and standardised design.

### *5.6.3 Periodicity*

All contractors insisted on the periodicity of the survey. However, only one country (Luxembourg) was prepared to do it annually. According to the contractors, the optimal periodicity is every 2 to 4 years. Some have suggested sending a shorter questionnaire with the quantitative questions every one or two years, and a longer one every 4 years. This will allow to produce longitudinal data for certain variables but also to reduce the burden on respondents and organisers.

From the perspective of the respondents there are advantages and disadvantages to have a periodical survey. On the one hand, it increases the burden of survey participation to

enterprises. On the other, the experience of Luxembourg and other countries where innovation surveys were already carried out in the past indicates that enterprises are more willing to participate if it enters in their standard routine. It has also been pointed out that enterprises will not keep records of some crucial quantitative variables such as innovation costs unless they know that they will be asked about it regularly.

## 5.7 Conclusions

Coming back, then, to the five steps of an ideal implementation process presented in section 5.2, the following provisional conclusions may be drawn:

- The conceptual basis was well developed through the work carried out in OECD in EU and in co-operation between OECD and EU.
- A harmonised questionnaire was agreed on. However, it was agreed on too late for some countries. Furthermore, the harmonised questionnaire could not be imposed on member states.
- A survey design was set up. However, it was not sufficiently developed and it was presented as recommendations and not as instructions.
- Timing was not satisfactory since some countries wanted to implement the survey before international agreement had been reached. Furthermore, in other countries, national factors delayed the implementation.
- As a consequence of these factors, a database with strictly comparable innovation data covering all EU countries and all questions of the harmonised questionnaire cannot be created.

All over this chapter we have mentioned specific cases where a greater harmonisation and standardisation in implementation could have occurred. However, we acknowledge that to carry out a survey at the European level is quite a different task than to carry out a survey in one country. It is not only much more demanding, but the results will never be as neat as those achieved in initiatives which are national in scope. This also implies that the launching of such an European survey does necessarily require a larger amount of efforts, planning and imagination to overcome unexpected problems.

The overview on the national experiences described here has also indicated quite clearly that nation-specific aspects have played a crucial role in the implementation of this survey. On the positive side, we consider that the variety of experiences carried out at the national level offers an unique learning opportunity. It is therefore essential to transmit in depth all the experiences carried out across all countries and this chapter has been an attempt to do it. The CIS experience could provide relevant lessons not only for any future innovation survey, but also to any survey with international ambitions.

## 5.8 Recommendations

This evaluation has shown that the co-ordination of the venture at the European level is essential in order to produce internationally comparable results. The analyses of the implementation of the CIS has led to the following recommendations.

- As a first step, it is essential to adopt a legal base for innovation surveys. This legal base will allow the Commission to co-ordinate the whole exercise, i.e. make sure that the timing is appropriate and that all national contractors "march together".

- In future Community Innovation Surveys we recommend Eurostat and DGXIII to release complete instructions on sampling. Equally important is to popularise among the national contractors the notion that sampling techniques may or may not allow data to be internationally comparable. It may also be useful to devote a specific workshop to this issue involving national contractors, the Commission and some independent scholars.

- The method used for the survey (voluntary/mandatory; census/sample, etc.) should also be approved at the European level. Although it does not seem feasible to apply a single method in all European countries, the Commission should be involved and discuss in advance the issue of comparing the results collected with different methodologies.

- Our field research has shown that a satisfactory implementation of the survey heavily relies on the combination of know how in statistical and economic issues. The teams which carry out the survey should include researchers with expertise in both the fields. The Commission should play an important role in the selection of the contractors. We recommend that the national statistical offices are involved in the national surveys or, at least, that the contractors have access to the frames of the statistical offices. In countries where expertise on innovation and business surveys is not yet developed, a tighter co-operation between the national contractors and the Commission is needed. Since contractors are different in terms of expertise and experience, the Commission should be flexible enough to provide guidance at different levels. For example, some national contractors might need the collaboration in statistical issues, others on the economics of technological change.

- The co-operation *among* contractors of different countries might also play an important role. In some cases it might be helpful to create working teams involving groups of a few countries.

- It might be useful to rename the survey to "Technological innovation and its obstacles in firms", and to state clearly that the survey is directed to both innovators and non-innovators. The design of the questionnaire should also improve the "route" to be followed by innovating and non-innovating enterprises.

- Pre-tests should play a greater role on the design of the survey and they should be co-ordinated at the European level. In future surveys, three different pre-test stages should be planned:

  a) Pre-test of the harmonised questionnaire at the European level and subsequent modification of the questionnaire.

  b) Distribution of the questionnaire to national contractors, with a recommendation to pre-test the questionnaire in the country if any modification is planned.

  c) National contractors report the conclusions of their pre-test to EU. The lessons learned should be disseminated across all countries and, if needed, the harmonised questionnaire should be revised before the release of the final version of the questionnaire.

- The statistical unit of the survey (legally defined enterprise, establishment, etc.) should be standardised and we recommend to use the legally defined enterprise. The Commission should provide instructions to national contractors. In countries which cannot apply the standard statistical unit (for example because an adequate frame is not available), devices to secure comparability of results should be planned in advance.

- Industrial classifications should be standardised across all participating countries. It is very important that all countries agree not only on the classification used, but also on its level of disaggregation. We recommend to record information on the NACE rev. 1, 4 digit level.

- Also information on the size of the enterprise needs to be standardised across all countries. We recommend that the number of employees of each enterprise is recorded.

- A common cut-off point should be decided on. We recommend to acquire information also on small enterprises (e.g. above 10 employees), although on a sample basis.

- The survey should be census based for all enterprises above a certain size. According to the size of the country the 'census point' may vary between 100 and 500 employees.

- In future surveys, we recommend to cover systematically also the service industries. The Commission should provide some guidelines on the service industries to be covered, also because the Oslo Manual does not provide detailed instructions on this issue. The advantages and disadvantages of a customises questionnaire for the service sector should also be explored. If needed, the customized questionnaire for the service industries should be prepared at the European level. Since CIS is the first large scale attempt to study innovation in services, we suggest to study the results from those countries which have covered services (Germany, Greece, Italy, the Netherlands and Portugal).

- Although the best methods for securing a high response rate and reliable information varies from country to country, the Commission should help to disseminate the best-practice experiences across all contractors and, when needed, provide advice and guidelines.

- In order to assess intra-enterprise variability in response, we advise the Commission to study the double responses from some enterprises in all EU countries. This small sample of questionnaires can provide valuable information on intra-enterprise variability in response.

# 6 Statistical issues

## 6.1 General aspects

### *6.1.1 Definition of sampling terms*

The population under study is defined in the Oslo Manual § 239, as follows:

> *The population of innovation surveys usually consists of enterprises in manufacturing industry. It may also be useful to include parts of the service sector, particularly enterprises that are working directly with manufacturers. It is important to obtain information on non-innovators as well as innovators, and on non-R&D performing innovators as well as R&D performing innovators.*

This population is called the "target population".

The definition is not operable in the sense that all enterprises belonging to the target population can really be identified and approached. For a complete census as well as for a sample survey always a "frame" (e.g. a file that contains at least the names and addresses of all enterprises in the target population) is needed which permits to get in contact with them.

The set of all enterprises delimited by the frame defines another population, which will be referred to as "frame population" in order to distinguish it from the target population.

The Oslo Manual § 240 the procedure to be followed:

> *Resource limitations will in most cases rule out a survey of the entire population, so a sample has to be designed. It must be representative of the industries covered and of the various types of innovators and non-innovators.*

The set of enterprises which is selected from the frame population according to the sample design is called the "gross sample". All enterprises in the gross sample are to be included in the Innovation Survey.

Generally, not all enterprises in the gross sample can be attained because the frame used for the survey is imperfect:

- the address of some enterprises is incomplete or incorrect in the frame,
- some enterprises do not exist anymore (they are collapsed or merged with another enterprise),
- some enterprises exist but do not belong to the target population (they are doing another business or they have become too small with respect to the cut-off point),
- some enterprises appear twice in the gross sample (being at least twice in the frame).

These cases are called "unreal non-response". They stand for all those enterprises in the frame population for which the information in the frame is either incorrect or not up-to-date. The gross sample adjusted for the set of unreal non-response cases is called the "net sample".

Some enterprises in the net sample do not respond or refuse explicitly to participate in the survey. These enterprises are called cases of (real) "unit non-response". The set of all responding enterprises constitute the "realized sample".

### *6.1.2 Determinants of quality*

Quality of results of the survey depends on the structural differences between the target population and the realized sample. These differences, called "errors", can be caused by three sources:

1. errors due to an imperfect frame (see subsection 6.1.3),
2. errors caused by sampling (see subsection 6.1.4) and
3. errors due to non-response (see subsection 6.1.5).

Normally, during the process of conducting a voluntary sample survey all three kinds of errors will arise and two of them in the case of a complete census.

### *6.1.3 Errors due to an imperfect frame*

Frames are imperfect because the target population, for which results are needed, deviates from the frame population to which the results refer (if the other two kinds of errors are negligible). There may be the following deviations:

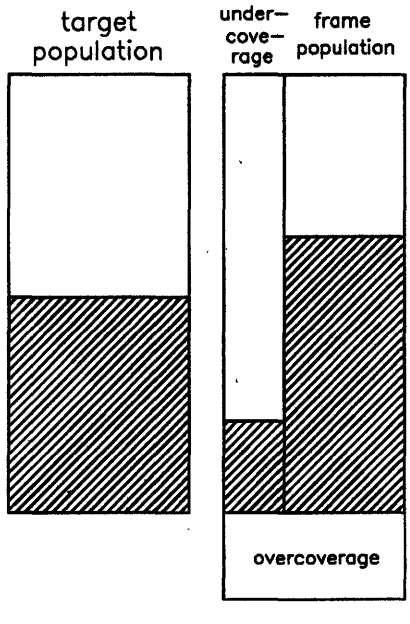
1. Deviations in coverage
   (a) Enterprises belonging to the target population are not included in the frame population (undercoverage)
   (b) enterprises belonging to the target population are several times in the frame population (excess coverage) and
   (c) enterprises contained in the frame population do not belong to the target population (overcoverage).

2. Deviations in content
   (a) The frame provides incorrect auxiliary information on enterprises (e.g. an enterprise is not engaged in manufacturing industry or the number of employees is not up to date) and
   (b) auxiliary information for enterprises is lacking in the frame.

The two groups of deviations differ fundamentally as to the type of errors caused by them. The deviations in coverage (group 1), especially undercoverage, leads to systematic errors both in a sample survey and a complete census. Deviations in content (group 2), on the other hand, only increase the random errors, if a sampling procedure is applied; generally, the results of a complete census are not at all – or only slightly – affected by deviations in content.

Graph 6.1 shows the two main deviations in coverage for a fictive target population. The areas of the rectangles labeled target population, frame population, undercoverage and overcoverage describe the number of enterprises in the corresponding sets. The shaded areas signify the number of innovative enterprises and the height of the shaded part of the rectangles the proportion of innovative enterprises in the considered set.

*Graph 6.1* *Fictive target population and deviations in coverage*




In this fictive example the proportion of innovative firms in the subset of enterprises not covered by the frame is assumed to be lower than in the frame population (21 % vs. 63 % in graph 6.1). Consequently, the survey can give neither the right proportion (assumed to be 49 %) nor the correct number of innovative enterprises. If this number in the target population would be 4900 and just two thirds of the target population were covered by the frame, the frame population would miss 700 (i.e. 14 %) of the innovative enterprises.

As mentioned above, this example is fictive, but it is not unrealistic: If the frame for the survey is deliberately restricted to those enterprises which are assumed to innovate (see table 6.1 in subsection 6.2.1), the proportion of innovating enterprises in the two parts of the target population (the undercoverage and the frame) are certainly different. Consequently, the proportion of innovating enterprises in the survey reflects a distorted picture of the target population.

Errors due to the imperfect frame occur regardless whether a complete census or a sample survey is performed and they will be the greater

- the smaller the proportion of the target population covered by the frame and
- the greater the difference between the fractions of innovative firms in the target population and in the frame population.

### *6.1.4 Errors caused by sampling*

"Sampling" is the process by which units listed in the frame, the "sampling units", are selected to form a sample, i.e. a subset of the frame population. Restricting a survey to the units in the sample implies, that a part of information gets lost, which could be gained if all enterprises in the frame population would really be included in the survey. Sample results always differ more or less from the results of the corresponding complete census. These differences are called "sampling errors".

Sampling errors can not be calculated according to this definition unless the results of the corresponding complete census were known. But the order of magnitude of sampling errors can be estimated by means of "standard errors" (see subsections 6.3.8, 6.4.1 and 6.4.2).

This procedure works if

- the sample has been selected by some sort of random mechanism and
- the probability to be included in the sample is specified for each enterprise (strictly speaking: for each set of enterprises) by the sample design.

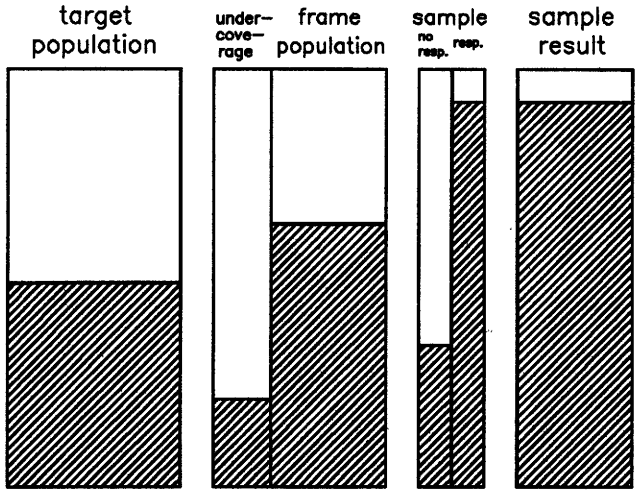
Standard errors measure the "accuracy" of the sample, i.e. the extent of random fluctuations between results which possibly could have been produced by the same sample design applied to the same frame.

### *6.1.5 Errors due to non-response*

Part of enterprises in the sample do not answer or even refuse to fill in the questionnaire. These cases of "unit non-response" bias the results if the set of non-responding enterprises and the set of responding enterprises are differing substantially with respect to their structure. The proportion of innovative enterprises generally will differ from the proportion in the set of responding firms.

Graph 6.2 on the following page enlarges the fictive example discussed in subsection 6.1.3 and shows the consequences of different proportions of innovating enterprises in the realized sample and in the set of non-responding units.

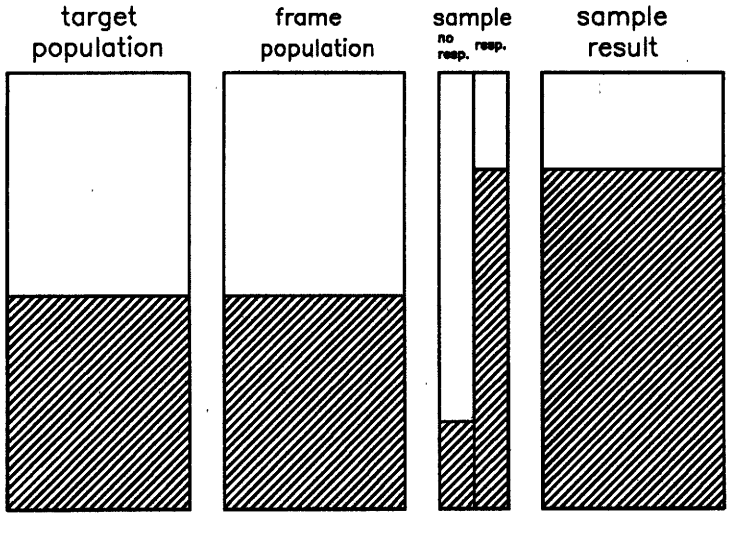
*Graph 6.2* *Fictive target population, sample and sample result in case of an imperfect frame*




The proportion of innovative enterprises is assumed to be 92 % in the set of responding firms and 34 % in the set of non-responding enterprises. Under these assumptions the estimate of the proportion of innovative enterprises is 92 % in contrast to 49 % in the target population.

If the coverage of the frame is nearly complete, as in most countries (see table 6.1), the difference between the proportion of innovating firms in the target population and the proportion in the frame population appears negligible. Nonetheless the effect of non-response can be severe:

*Graph 6.3* *Fictive target population, sample and sample result in case of a perfect frame*




For sake of comparability this fictive example assumes that the proportion of innovating enterprises is 78 % in the set of responding firms and 20 % in the set of non-responding enterprises (i.e. the difference between these proportions is the same as in graph 6.2).

Under these modified assumptions the proportion of innovating enterprises is estimated to be 78 % in contrast to 49 % in the target population.

Generally, the proportion of innovative enterprises in the non-response set is unknown. If raising factors are calculated by means of the formula

$$\frac{\text{number of enterprises in the frame population}}{\text{number of responding enterprises in the sample}}$$

and data are raised with this factor, the number of innovative enterprises would be estimated to be 7 800 (in case of the imperfect frame) and 6 120 (in case of the perfect frame) while there are really 4 900 innovative firms in the target population.

Both results gained from the sample are severely biased: Considering the case of a perfect frame, the relative bias of the estimated proportion of enterprises with innovations is about 59 % and the estimated number of innovative firms is also distorted by 59 %. Generally the bias of results due to non-response will be the greater the higher is the non-response rate.

The bias in the results cannot be estimated from the data acquired by the sample. It can only be approximately assessed by drawing and evaluating a subsample from the set of all non-responding enterprises. In order to produce reliable estimates of the bias induced by unit non-response, at least some information should be get from nearly all enterprises included in the subsample.

The fact, that some enterprises returning their questionnaire do not fill in all relevant items, is called "item non-response". This kind of non-response is concentrated on items which are not easy to be answered. The higher the rate of non-response for an item the more biased is results for this item will be.

## 6.2 Sampling techniques used

This section covers the main statistical features of CIS in the twelve member countries of the European Union and in Norway. The report rests on information given by the national contractors during the interviews of the evaluation group and later in written form. Relevant details are given in annex A.

### *6.2.1 Frame, coverage of frame and sampling units*

The frames used in the member countries for the CIS are very different both with respect to their origin and their coverage. On the other hand, in nearly all countries enterprises are used as sampling units and as reporting units (see section 5.3.4). The main facts are compiled in Table 6.1 .

*Table 6.1* *Frame, coverage of frame and sampling units*

| Country | origin of frame | coverage of frame | sampling units |
|---|---|---|---|
| Belgium | | | |
| Flemish part | official data base | not reported | establishment |
| Wallonian part | official data base | not reported | establishment |
| Bruxelles | official data base | not reported | establishment |
| Denmark | official data base | nearly complete | enterprise |
| France | official data base | nearly complete | enterprise |
| Germany | data base of Credit Reform Verein | nearly complete | enterprise |
| Greece | parts of different data bases | incomplete | enterprise |
| Ireland | official data base | nearly complete | enterprise |
| Italy | official data base | nearly complete | enterprise |
| Luxembourg | official data base | nearly complete | enterprise |
| Netherlands | data base of Chamber of Commerce | nearly complete | principal establishment |
| Norway | official data base | nearly complete | legally defined enterprise |
| Portugal | parts of different data bases | incomplete | enterprise |
| Spain | official data base | nearly complete | enterprise |
| United Kingdom | | | |
| pilot survey | official data base | nearly complete | enterprise |
| real survey | data base of mailing house | not known | autonomous enterprise |

Two countries, Greece and Portugal, have deliberately restricted their frame to enterprises which are judged to be engaged in innovation, i.e. they used a rather incomplete frame (see the definition of target population by the Oslo Manual in subsection 6.1.1). This fact will make it impossible to compare results gained by Greece and Portugal with results of other countries unless these results are curtailed to sets of innovating firms.

### *6.2.2 Information contained in the frame*

The frames used for the Innovation Survey contain in most countries not only the name and address of the reporting units, but also some auxiliary information: The number of employees (or at least the employment class) and the branch of sampling units in terms of the national classification of business activities. These classifications can be translated generally at least to two-digit NACE.

Of course, the auxiliary information is not always exactly up-to-date, but it seems to be of rather high quality in most countries. This is true especially for those countries in which the contractor has access to an official data base constructed and regularly updated for fiscal purposes. The contents of frames used in Germany and in the Netherlands have been analyzed and proved to be reliable.

The frames on which the pilot survey and the real survey in the United Kingdom have been based are very different with respect to their coverage and to their content. The Central Statistical Office (CSO) was able to use its official data base for the very small pilot study. On the other hand, the Confederation of British Industry (CBI) has engaged a mailing house to select the sample. This subcontractor has used its own address file as frame. The quality of this frame is unknown and should be checked.

### *6.2.3 Sampling method and stratification*

In five countries no sample has been selected for the Community Innovation Survey. In Ireland and Luxembourg a sample is not appropriate because the number of enterprises is too small. Greece and Portugal have surveyed all enterprises contained in their deliberately restricted frame (see subsection 6.2.1). The Statistical Office in Italy and Spain aimed at a complete census. All other countries have used some sort of sampling.

*Table 6.2* *Sampling method, stratification variables and number of strata*

| Country | sampling method | stratification variables | number of strata |
|---|---|---|---|
| Belgium | | | |
| Flemish part | systematic | branch, employment class | 20 · 7 |
| Wallonian part | systematic | branch, employment class | 20 · 7 |
| Bruxelles | systematic | branch, employment class | 20 · 7 |
| Denmark | random | branch, employment class | 24 · 4 |
| France | systematic | branch, employment class | 50 · 7 |
| Germany | modified systematic with random start | branch, employment class and regional parts | 50 · 7 · 2 |
| Netherlands | modified random | branch, employment class | 43 · 6 |
| Norway | random | employment class | 4 |
| United Kingdom | | | |
| pilot survey | random | branch, employment class | ? |
| real survey | unknown | employment class | ? |

The method of pure random sampling by strata is a safe way to secure a good sample, but it has the drawback that regional information in the frame is not used. Therefore, the subcontractor in Germany applied systematic selection with random start from a file which has been ordered according to regions; by means of this procedure a fair representation of all regions can be reached. The contractor in the Netherlands has modified the pure random method in order to select as often as possible firms included in the survey of 1989; such a panel is apt to give higher accuracy in comparing results of subsequent surveys.

These two modification should be considered for the design of further Innovation Surveys. They can be combined by means of the so-called Deming plan (Deming 1956), an easy to use procedure, which is safe and sound (see subsection 6.6.3).

### *6.2.4 Sampling fractions and sample size*

All countries listed in table 6.2 have used different sampling fractions for strata. Various methods have been applied to associate sampling fractions with strata:

- Some contractors have preferred to differentiate them only by employment classes. This form of allocating sampling fractions to strata is very easy, but it implies the drawback, that samples from branches with a small number of firms become also small and, consequently, the related results may be rather inaccurate.

- Other contractors have assorted sampling fractions taking the size of strata into account.

- The so-called "optimum allocation" according turnover has been used at least in Denmark. This method allocates a given overall sample size to a set of strata (e.g. all size classes within a branch) with the aim, to reach the highest accuracy of results for a prescribed set of strata. This method secures accurate results for turnover as well as for all characteristics highly correlated with turnover, but this is true only for results which correspond to the prescribed set of strata. The accuracy of results for single strata (e.g. a single size group in the branch) may be very poor.

These differences in the method of allocation of sample size to strata influence sampling errors of results (except those based entirely on strata with sampling fraction 100 %). In subsections 6.4.1 and 6.4.2 it is shown that the level of sampling errors depend mainly on the sample size. Therefore, the method of allocation of sample size to strata should be harmonized in order to reach a necessary condition of the comparability of results.

The size of the gross sample depends on both the size of strata and the sampling fractions applied (in annex A sample sizes by strata are listed for most countries). The general assessment of the sampling procedure will be based on overall sample sizes. This restriction makes it easy to compare the size of the gross sample with the size of the frame. The ratio of these two numbers is the average sampling fraction. This fraction should be the higher the smaller is the size of the frame.

In table 6.3 all countries are listed. For countries, where no sample has been selected, the size of gross sample equals the size of the frame.

*Table 6.3* *Size of frame and size of gross sample and average sampling fraction*

| Country | size of | | sampling |
|---|---|---|---|
| | frame | gross sample | fraction in % |
| Belgium | | | |
| Flemish part | 5953 | 1335 | 22 |
| Wallonian part | 1876 | 438 | 23 |
| Bruxelles | 787 | 176 | 22 |
| Denmark | 3071 | 1313 | 43 |
| France | 14751 | 5245 | 36 |
| Germany | 126319 | 13320 | 11 |
| Greece | 1799 | 1799 | 100 |
| Ireland | 3032 | 3032 | 100 |
| Italy | 35182 | 35182 | 100 |
| Luxembourg | 470 | 470 | 100 |
| Netherlands | 45782 | 8221 | 18 |
| Norway | 5606 | 1882 | 34 |
| Portugal | 1767 | 1767 | 100 |
| Spain | 20924 | 18002 | 86 |
| United Kingdom | | | |
| pilot survey | 40450 | 44 | 0.1 |
| real survey | 40450 | 4998 | 12 |

### *6.2.5 Subsampling for non-response*

It was to be expected from the very beginning that the response rate in Innovation Surveys would be rather low because the questionnaire is long and contains questions which are quite new and not easy to answer. Other reasons for low response rates are, that the survey in most countries must rely on voluntary participation and firms are generally fatigued to fill in questionnaires.

The contractors in Ireland, the Netherlands and Norway have reported that they had drawn subsamples from the set of non-respondents which helps to assess the order of magnitude of bias due to non-response.

In the Netherlands a subsample of 520 firms from the non-respondents has been drawn. These firms have been asked to answer two simple questions (see annex A). A series of up to seven reminders has been conducted in order to reach a response rate as high as possible for the subsample.

A similar plan was realized in Ireland. A subsample of 732 enterprises has been drawn from the set of 1466 enterprises which did not respond in the first round. They were telephoned and asked the same questions as in the Netherlands (see Annex A).

During the evaluation we have strongly recommended to supplement the former sample design by drawing and evaluating a subsample of non-respondents in order to get some information on the bias due to non-response. According to this recommendation some countries have performed or will perform non-response surveys. Details of these surveys are not yet available.

## 6.3 Features of realized samples

### *6.3.1 Reminders to overcome non-response*

In nearly all countries contractors have tried to reduce non-response by a series of reminders. Most of them have sent some reminders by mail and then attempted to get responses by phone.

### *6.3.2 Unit response*

The size of the realized sample is equal to the number of responding enterprises, i.e. the unit response (cf. subsection 6.1.5). Table 6.4 shows the size of gross sample and realized sample by country. The ratio of these two numbers is the "raw response rate". It is taken as a proxy for the correct response rate which is based on the size of the net sample. Most countries have not yet reported the size of the net sample. So raw response rates must be used for the comparison.

*Table 6.4* *Size of gross and realized sample and raw average rates of response and non-response*

| Country | size of | | average raw rate (*) of | |
|---|---|---|---|---|
| | gross sample | realized sample | response % | non-response % |
| Belgium | | | | |
| Flemish part | 1335 | . | 45 | 55 |
| Wallonian part | 438 | . | 35 | 65 |
| Bruxelles | 176 | . | 40 | 60 |
| Denmark | 1313 | 674 | 51 | 49 |
| France | 5245 | 3911 | 75 | 25 |
| Germany | 13320 | 2954 | 22 | 78 |
| Greece | 1799 | 1660 | 92 | 8 |
| Ireland | 3032 | 1003 | 33 | 67 |
| Italy | 35182 | 22493 | 64 | 36 |
| Luxembourg | 470 | 372 | 79 | 21 |
| Netherlands | 8221 | 4084 | 50 | 50 |
| Norway | 1882 | 986 | 52 | 48 |
| Portugal | 1767 | . | 70 | 30 |
| Spain | 18002 | 2372 | 13 | 87 |
| United Kingdom | | | | |
| pilot survey | 44 | 17 | 39 | 61 |
| real survey | 4998 | 182 | 4 | 96 |

(*) Estimated values if size of realized sample is unknown

The raw response rates are around 50 % in most countries, in some of them they are far smaller. Four countries have reached response rates of 70 % and more: Greece and Portugal have deliberately restricted their frames, Luxembourg has performed telephone interviews of all enterprises in its small population and France has used a questionnaire which was short compared to questionnaires in all other countries.

### *6.3.3 Gauging the impact of unit non-response*

There is a rule of thumb that results of a survey with response rate below 75 % will be distorted by errors due to non-response. It has been shown by two fictive examples (see section 6.1.5) that non-response of 50 % can cause severely biased results. The real impact of unit non-response can be gauged if not only a subsample from the non-respondents has been drawn but also a very high response rate has been reached in this subsample.

According to subsection 6.2.5 three countries have planned and also realized such a subsample; two of them have reported results of their subsample. In the Netherlands 79 % of the firms in the subsample have answered and in Ireland the even higher response rate of 87 % has been reached. The results gained by subsampling are described in annex A.

In subsections 6.4.3 and 6.4.4 some results reported by contractors are evaluated. This evaluation shows that severe problems are induced by non-response.

### *6.3.4 Dealing with item non-response*

Generally, item non-response (cf. subsection 6.1.5) is concentrated on questions which are difficult to answer. Contractors in nearly all countries have reported that especially the following questions have been found difficult and are suffering from high item non-response:

innovation costs
sales according to life cycle
sales from new or improved products

Most countries have reported serious efforts to reduce item non-response by means of phone interviews or by mail. Anyway, the rates of item non-response may be different between the countries. These rates have not been reported but for the Netherlands (see annex A).

Because item non-response is an additional source of bias, it seems advisable to ask contractors in all countries to find out the rates of item non-response for all main questions during the procedure of data processing.

### *6.3.5 Imputation of missing data*

Item non-response causes missing data which hamper the evaluation of data obtained by the survey. The best way to overcome this obstacle and, possibly, also to reduce bias due to item non-response is imputation. These methods try to replace missing data by proximate values which are estimated on the basis of complete data provided by similar units (e.g. an enterprise of the same branch and size class, which has nearly the same values in some important characteristics).

Eurostat has envisaged to use imputation in its data processing.

### *6.3.6 Plausibility and compatibility of data*

Checks of both plausibility and compatibility have been applied in all countries. These checks are done during the course of data processing. Methods used and characteristics checked are defined by the programs used for processing and have not been detailed during the interviews.

### *6.3.7 Raising factors and estimation of aggregates*

The process of diminishing the frame population to the sample has to be reversed by raising the sample data to results pertaining to the population. Stratified sampling and the use of different sampling fractions imply that also the reversion of the sampling procedure must be done separately for each stratum.

If the rate of unit non-response is low (e.g. markedly less than 25 %) the following formula is applicable for calculating raising factors for free expansion:

$$\text{raising factor} = \frac{\text{total number of units in stratum}}{\text{number of units in realized sample arising from this stratum}}$$

This formula implies the assumption that the set of non-responding units behaves exactly as the set of responding units. If non-response is low this assumption may be justified, at least approximately.

This will hardly be true if the non-response rate is higher than 25 %. The higher this rate the more unrealistic is this assumption, the larger are therefore the errors induced by these raising factors (see the discussion of bias in subsection 6.3.3).

The results produced by this raising method represent at best values of the frame population (see section 6.1.1). If the target population is nearly equal to the frame population, the results may be applicable also for the target population. If, however, the frame used implies a severe discrepancy between target population and frame population, the results produced by the survey cannot be regarded to be valid for the target population.

Some contractors seemingly did not understand the necessity to allocate weights to data coming from strata with different sampling fractions, other contractors intend to use free estimation of aggregates, i.e. multiplication of data by raising factors and adding these products. Eurostat prepares weighting for the estimation of aggregates.

### *6.3.8 Assessing the accuracy of results*

Use of sampling methods always has the consequence that results gained from the sample differ from the comparable results of a total census. The difference between these two kinds of results is called "sampling error". The order of magnitude of sampling errors can be estimated by means of standard errors, if random sampling has been done.

The calculation of standard error, at least for the most important results, has not been reported in any of the interviews. Therefore, the contractors cannot assess the accuracy of results, i.e. the importance of sampling fluctuations. Eurostat intends to calculate standard errors of main results.

## 6.4 Reliability of aggregated data

The Community Innovation Survey (CIS) is a new statistical project and till now no information is available, which could help to assess the reliability of data and results by aggregating data. Contractors have not reported any attempts to gauge sampling errors (see subsection 6.3.8). So it may be, that data and results will be used without regard to their reliability.

In this section some quantitative statements are presented in order to give at least a feeling for the magnitude of errors due to sampling and non-response. It should be clear that these statements cannot replace investigations regarding the specific conditions of different countries.

The following investigation treats sampling errors (in subsections 6.4.1 and 6.4.2) and errors due to non-response (in subsections 6.4.3 and 6.4.4).

### *6.4.1 Sampling errors of proportions*

Sampling errors can be judged by means of standard errors, provided that the sample has been drawn at random. Generally, in most cases (approximately 95 %) the absolute value of a sampling error is less than the twofold standard error.

Standard errors depend on the sampling device, the variability of the investigated characteristics and the sample sizes within the strata. In order to simplify matters we restrict this overview to random sampling without stratification (i.e. simple random sampling). Stratification is apt to reduce sampling errors. Consequently, standard errors for stratified sampling are smaller than those for simple random sampling. The extent of reduction depends on different facts, but in general the stratification effect is less than a quarter of the standard error.

At first we consider standard errors for proportions of enterprises in groups defined by some characteristic. These standard errors are lower bounds of the standard errors for estimated totals of quantitative characteristics in the same groups.

For simple random sampling the standard error $s_p$ of the proportion $p$ in a group of $n$ units can be approximated by the formula

$$s_p \approx \sqrt{(1-f) \cdot \frac{p \cdot (1-p)}{n}} \quad ,$$

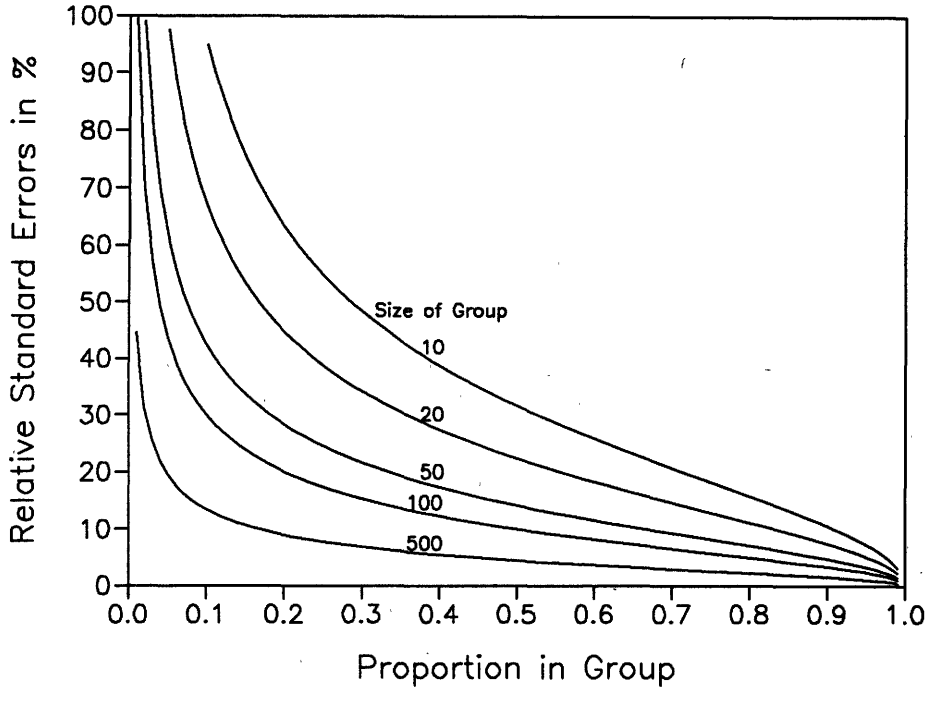
if the size $n$ of the group is at least 10 and if the proportion $p$ is greater than 5 % or less than 95 %. The symbol $f := n/N$ denotes the sampling fraction, i.e. the ratio of sample size $n$ and size $N$ of the frame population.

In most circumstances it is easier to deal with relative standard errors. They can be approximated by the formula

$$\frac{s_p}{p} \approx \sqrt{(1-f) \cdot \frac{1-p}{p \cdot n}}$$

In graph 6.4 relative standard errors for different sample sizes are displayed; the sampling fractions are assumed to be less than 10 % so $\sqrt{1-f} \geq 0.95 \approx 1$. If sampling fractions are 20 % (or 30 %), the values of relative standard errors shown in graph 6.4 are to be multiplied by 0.89, respectively by 0.84.

*Graph 6.4* *Relative standard errors of proportions*
*– approximation for simple random sampling –*




For example, the proportion $p = 20\%$ in a group of size $n = 20$ has a relative standard error of 45 %, so that the absolute standard error is $0.45 \cdot 20\ \% = 9\ \%$. That means that the proportion $p = 20\%$ may fluctuate due to sampling errors in the interval

$$\text{from} \quad 20\ \% - 2 \cdot 9\ \% = 2\ \% \quad \text{to} \quad 20\ \% + 2 \cdot 9\ \% = 38\ \% \ .$$

If the same proportion would be based on a sample of $n = 500$ the random fluctuations of $p = 20\%$ would lie in the range from 16.5 % to 23.5 % in case of sampling fractions less than 10 %. If the sampling fraction is 20 % the random errors fluctuate in the range from 17 % to 23 %.

The size of sampling errors should always be taken into account when results of a sample survey are reviewed. In general, the higher the proportion and the larger the size of the group the smaller is the relative standard error of the proportion.

Sampling errors impede the interpretation of results, but they can cause severe failures, if results for two groups are compared and interpreted to be important. The comparison of the proportion $p_1$ for group 1 with the proportion $p_2$ for group 2 seems to be extremely simple: The difference $p_1 - p_2$ of these proportions in the sample is considered. If the difference is greater than zero, i.e. $p_1$ is greater than $p_2$, inexperienced users are tempted to conclude that this is also true for the corresponding groups in the population. They do not pay attention to the fact that sampling errors of differences of two proportions are even larger than sampling errors of a single proportion.

The risk of incorrect conclusions can be avoided in the following way: The obtained difference of results must be compared with its standard error calculated under the assumption that their counterparts in the population do not differ. If the difference does not exceed

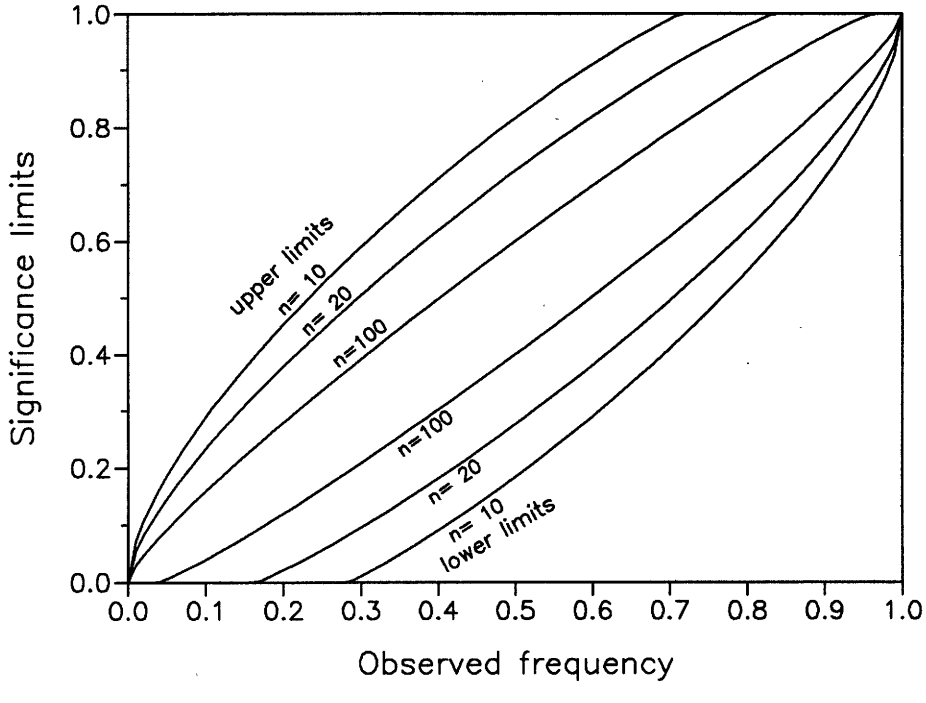
twice its standard error; then it must not be argued that there is a real difference in the population.

A difference of proportions $p_1$ and $p_2$ based on groups with size $n_1$ respectively $n_2$ can be taken to be significant if the following condition is fulfilled:

$$|p_1 - p_2| > 2 \cdot \sqrt{\frac{p_1 n_1 + p_2 n_2}{n_1 + n_2} \cdot \left(1 - \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2}\right) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

This condition depends both on the sizes $n_1$ and $n_2$ of the two groups and on the two proportions $p_1$ and $p_2$. In order to give some clues of the importance of random fluctuations of differences, a simplified case is considered: It is assumed that the two groups are of the same size (i.e. $n_1 = n_2$) and the proportion $p_1$ in the first group is given. Then the upper and lower limits of the proportion $p_2$ in the second group can be calculated which assure that the difference $p_1 - p_2$ is significant.

*Graph 6.5* *Limits of proportion in group 2 given the proportion in group 1 and equal sizes n*




Graph 6.5 shows that if the size $n$ of both groups is 100 and the proportion in group 1 is $p_1 = 20\%$ the proportion in group 2 has to be smaller than 11 % or larger than 29 %, otherwise the difference must not be stated relevant.

### *6.4.2 Sampling errors of total values*

The discussion was restricted so far to standard errors of proportions although innovation surveys are meant to give results for a lot of *quantitative characteristics*, e.g. innovation

expenditures and R&D expenditures in predefined groups of enterprises. Relative standard errors $v_{\hat{x}}$ for estimates of *total values* $\hat{x}$ of characteristic $X$ in a domain of study can be approximated by the formula

$$v_{\hat{x}} \approx \sqrt{(1-f) \cdot \frac{v_g^2 + (1-p)}{p \cdot n}} \quad ,$$

where $v_g$ is the coefficient of variation of characteristic $X$ in the group g of enterprises in the domain, for which the characteristic is unequal zero. A comparison of this formula with the formula for the relative standard error of proportions (see subsection 6.4.1) shows, that standard errors for quantitative characteristics are always larger than standard errors of numbers in the same group.

The inflation of standard errors depends on the coefficient of variation $v_g$ of the characteristic in group g. The German subcontractor has kindly submitted on demand material for some branches, size groups and three characteristics. The calculated coefficients of variation are superimposed by sampling errors, so the results have been graduated (see table 6.6 on the following page).

The relative standard errors exhibited in graph 6.6. are calculated for a medium sized group on the basis of graduated values.

*Graph 6.6* *Relative standard errors of total values*
*– approximation for simple random sampling –*

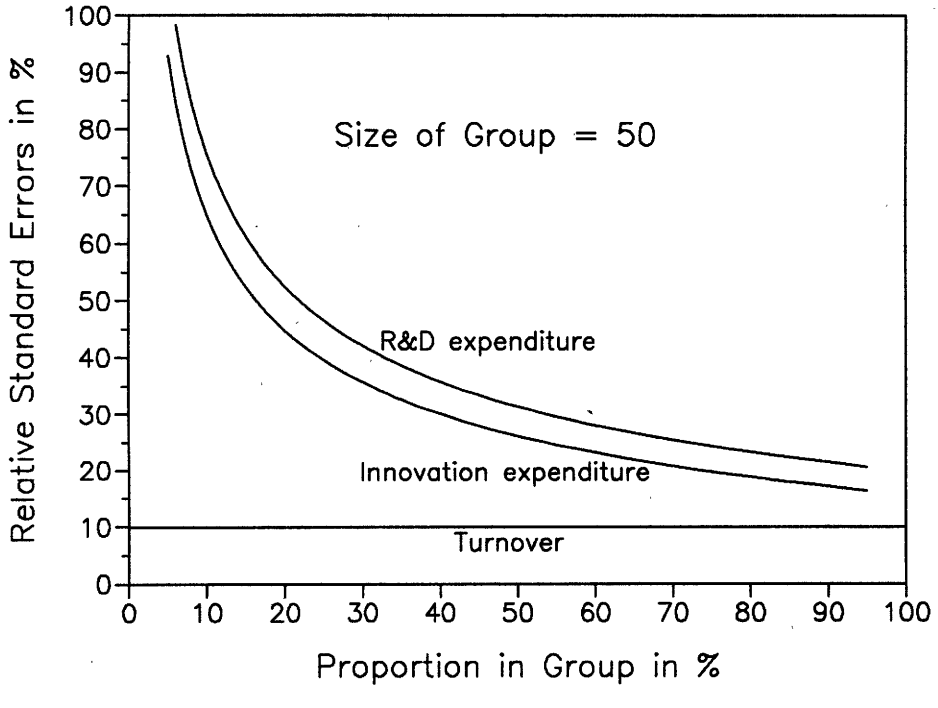


*Table 6.5 Graduated coefficients of variation*

| characteristic | coefficient of variation |
|---|---|
| turnover | 0.7 |
| innovation expenditure | 1.1 |
| R&D expenditure | 1.4 |

Graph 6.6 shows that relative standard errors of the two expenditures are in all cases considerably larger than the relative errors of turnover, a characteristic which applies for all enterprises. Sampling errors of total values for groups, in which only a small part of enterprises is innovating, are much higher than sampling errors of turnover.

### *6.4.3 Errors of proportions due to non-response*

The extent of bias induced by non-response cannot be assessed but by an evaluation of results which are gained by subsamples drawn from the set of non-responding firms. At first results from the Netherlands (see annex A) are used to provide some information about bias due to non-response. This material is broken down by six size classes and by two groups (manufacturing and services) for the two questions asked in the subsample (R&D department and Innovation).

According to table 6.6 the proportion of manufacturing firms reporting to have a R&D department is 28.3 % in the sample and 35.2 % in the subsample of non-respondents. Obviously the real proportion of firms having a R&D department in manufacturing is somewhat underestimated due to non-response. To get a more precise proportion of R&D firms the two results must be combined. The calculation of the combined proportion has to take into account the following information:

| | in sample | in subsample |
|---|---|---|
| number of responding firms | $n_r$ | $n'_r$ |
| number of non-responding firms | $n_{nr}$ | $n'_{nr}$ |
| response rate | $q_r$ | $q'_r$ |
| proportion of firms with R&D dept. (or proportion of innovating firms) | $p$ | $p'$ |

The proportion of firms having a R&D department (respectively the proportion of innovating firms) in the frame population can be estimated according to formula

$$p_c = \frac{n_r \cdot p + n_{nr} \cdot [\lambda + (1-\lambda) \cdot q'_r] \cdot p'}{n_r + n_{nr}} \quad .$$

$\lambda$ designates the rate of *non-responding* enterprises in the subsample which are supposed to have the same behavior (proportion of R&D department) as the *responding* firms in the subsample .

If the response-rate $q'_r$ in the subsample is less than 1, the estimate depends on the assumptions made for $\lambda$. Assuming that all non-responding enterprises in the subsample do not have an R&D department, the parameter $\lambda$ is equal zero. Under the assumption, that the proportions of firms with R&D are equal

in the set of responding firms in the subsample and

in the set of non-responding firms in the subsample,

the parameter $\lambda$ is equal 1 and the proportion in the frame population is to be estimated by

$$p_c = \frac{n_r \cdot p + n_{nr} \cdot p'}{n_r + n_{nr}} \quad .$$

The two values of $\lambda$, which reflect different assumptions on the structure of non-responding firms in the subsample, may be used to show how much the estimated proportion of firms having a R&D department (are innovating) in the population depends on the guessed parameter $\lambda$. For $\lambda = 0$ the example mentioned above gives

$$p_c = \frac{2093 \cdot 28.3 + 2028 \cdot \frac{206}{260} \cdot 35.2}{2093 + 2028} = 28.1$$

and for $\lambda = 1$

$$p_c = \frac{2093 \cdot 28.3 + 2028 \cdot 35.2}{2093 + 2028} = 31.7 \quad .$$

So the proportion in the population can be expected to lie in the interval from 28.1 % to 31.7 %. The result 28.3 % obtained from the sample is greater by 0.2 % respectively smaller by 3.4 % than the boundaries of the interval. These differences are point estimates of the absolute bias under different assumptions. The real value of bias cannot be ascertained, because it depends on the unknown properties of non-responding firms, which is reflected by the parameter $\lambda$.

The assumption $\lambda = 0$ and the corresponding result is highly unrealistic. It seems more realistic to assume, that non-respondents in the subsample have nearly the same properties as respondents in the subsample. Under this assumption also the problem is avoided, that the response rate must be known for each group, if the parameter $\lambda$ is guessed to be less than 1. Therefore, the results are evaluated according to the formula for $\lambda = 1$.

According to the results in table 6.6 (on the following page) non-response of firms in the sample has the consequence that nearly all proportions are underestimated. The negative bias is rather large for most results, especially in the domain of services. On the other hand it seems that proportions for the smallest manufacturing firms are overestimated by the sample.

*Table 6.6 Non-response analysis of results from the subsample in the Netherlands*

| Number of employees | Number of firms | | | Proportion of firms | | | Bias of proportion | | Std.error of abs. Bias |
|---|---|---|---|---|---|---|---|---|---|
| | in sample | | in subsample | in sample | in subsample | combined | | | |
| | resp. | n-resp. | | | | | abs. | rel. | |
| | | | | % | % | % | % | % | % |
| *R&D department reported in Manufacturing* | | | | | | | | | |
| 10 - 19 | 299 | 313 | 22 | 6.5 | 4.5 | 5,5 | 1,0 | 19 | 0,6 |
| 20 - 49 | 378 | 384 | 33 | 12.0 | 15.2 | 13,6 | - 1,6 | -12 | 0,9 |
| 50 - 99 | 720 | 740 | 78 | 28.3 | 32.1 | 30,2 | - 1,9* | - 6 | 0,8 |
| 100 - 199 | 387 | 367 | 48 | 40.2 | 54.2 | 47,0 | - 6,8* | -14 | 1,3 |
| 200 - 500 | 228 | 173 | 19 | 52.2 | 68.4 | 59,2 | - 7,0* | -12 | 1,7 |
| 500 - | 81 | 51 | 6 | 54.2 | 50.0 | 52,6 | 1,6 | 3 | 3,1 |
| Total | 2093 | 2028 | 206 | 28.3 | 35.2 | 31,7 | - 3,4* | -12 | 0,5 |
| *R&D department reported in Services* | | | | | | | | | |
| 10 - 19 | 763 | 717 | 49 | 4.8 | 2.0 | 3,4 | 1,4* | 39 | 0,3 |
| 20 - 49 | 422 | 324 | 43 | 3.6 | 11.6 | 7,1 | - 3,5* | -49 | 0,6 |
| 50 - 99 | 281 | 246 | 27 | 8.2 | 3.7 | 6,1 | 2,1* | 34 | 0,7 |
| 100 - 199 | 171 | 163 | 19 | 8.2 | 15.8 | 11,9 | - 3,7* | -31 | 1,2 |
| 200 - 500 | 269 | 172 | 19 | 19.4 | 31.6 | 24,2 | - 4,8* | -20 | 1,4 |
| 500 - | 85 | 50 | 2 | 24.7 | 100.0 | 52,6 | - 27,9 | -53 | • |
| Total | 1991 | 1672 | 159 | 8.7 | 14.4 | 11,3 | - 2,6* | -23 | 0,4 |
| *Innovation reported in Manufacturing* | | | | | | | | | |
| 10 - 19 | 299 | 313 | 22 | 29.0 | 22.7 | 26,3 | 2,7* | 10 | 1,2 |
| 20 - 49 | 378 | 384 | 33 | 42.3 | 36.4 | 39,6 | 2,8* | 7 | 1,2 |
| 50 - 99 | 720 | 740 | 78 | 55.1 | 70.5 | 63,1 | - 8,0* | -13 | 0,9 |
| 100 - 199 | 387 | 367 | 48 | 68.3 | 70.8 | 69,7 | - 1,4 | - 2 | 1,2 |
| 200 - 500 | 228 | 173 | 19 | 73.0 | 79.0 | 75,7 | - 2,7 | - 4 | 1,5 |
| 500 - | 81 | 51 | 6 | 74.4 | 100.0 | 85,3 | - 10,9 | -13 | • |
| Total | 2093 | 2028 | 206 | 54.4 | 62.4 | 58,4 | - 4,0* | -7 | 0,5 |
| *Innovations reported in Services* | | | | | | | | | |
| 10 - 19 | 763 | 717 | 49 | 22.9 | 34.7 | 27,5 | - 4,6* | -17 | 1,0 |
| 20 - 49 | 422 | 324 | 43 | 25.6 | 58.1 | 42,0 | - 16,4* | -39 | 1,4 |
| 50 - 99 | 281 | 246 | 27 | 3.1 | 55.6 | 28,8 | - 25,7* | -89 | 1,2 |
| 100 - 199 | 171 | 163 | 19 | 43.3 | 57.9 | 51,0 | - 7,7* | -15 | 2,3 |
| 200 - 500 | 269 | 172 | 19 | 51.1 | 68.4 | 58,3 | - 7,2* | -12 | 2,0 |
| 500 - | 85 | 50 | 2 | 56.5 | 100.0 | 64,8 | - 8,3 | -13 | • |
| Total | 1991 | 1672 | 159 | 32.5 | 55.2 | 42,6 | -10,1* | -24 | 0,7 |

• approximation is too bad or impossible

The estimated bias of proportions is superimposed by sampling errors. The level of these sampling errors can be judged by means of standard errors if random sampling has been used. They are approximated in a rather coarse manner because the stratification effects (which are apt to diminish the errors) cannot be considered here. The approximated standard error $s_b$ of the estimated absolute bias $b$ of the proportion estimated under the assumption $\lambda = 1$ is

$$s_b \approx \sqrt{\left(\frac{n_r}{n_r + n_{nr}}\right)^2 \cdot \frac{p(1-p)}{n_r + n_{nr}} + \left(\frac{n_{nr}}{n_r + n_{nr}}\right)^2 \cdot \frac{p'(1-p')}{n'_r + n'_{nr}}} \quad .$$

The approximations calculated by this formula are listed in the last column of table 6.6. Some of them are not shown because the approximation is too bad (the number of firms in the subsample is too small or the proportion in the subsample is 100 %).

If the ratio

$$\frac{\text{bias}}{\text{standard error of bias}}$$

is greater than 2, the probability is less than 2.5 % that the estimated bias differs from zero not at random. The appropriate results are marked in table 6.6 by asterisks.

The contractor in Ireland has also drawn a subsample from the set of non-responding firms (see the detailed results in annex A). Table 6.7 (on the following page) shows the non-response analysis based on these results.

The main results from the Netherlands (in table 6.6) and from Ireland (in table 6.7) are to be collated.

*Table 6.8 Main results of the non-response analysis in the Netherlands and Ireland*

| Manufacturing firms | Proportion | | | | | | bias | |
|---|---|---|---|---|---|---|---|---|
| | in survey | | | revised | | | | |
| | NL % | | IRL % | NL % | | IRL % | NL % | IRL % |
| with R&D department | 28.3 | < | 61.0 | 31.7 | > | 29.5 | - 3.4 | 31.6 |
| reporting innovations | 54.4 | < | 71.2 | 58.4 | > | 33.0 | - 4.0 | 38.2 |

These findings may be summarized as follows:

- The proportions shown by the survey are smaller for the Netherlands than for Ireland, but the opposite is true if the revised proportions are considered.
- The direction of bias due to non-response in the Netherlands is opposite to the direction found in Ireland.
- In absolute terms the bias in Ireland is even much greater than the bias in the Netherlands (and significantly different from zero in nearly all groups).

*Table 6.7 Non-response analysis of results from the subsample in Ireland*

| Number of employees | Number of firms | | | Proportion of firms | | | Bias of proportion | | Std.error of abs. Bias |
|---|---|---|---|---|---|---|---|---|---|
| | in sample | | in subsample | in sample | in subsample | combined | | | |
| | resp. | n-resp. | | % | % | % | abs. % | rel. % | % |
| *R&D department reported in NACE group 2* | | | | | | | | | |
| 10 – 49 | 85 | 127 | 38 | 60.0 | 15.8 | 33.5 | 26.5* | 44 | 3.3 |
| 50 – 99 | 27 | 58 | 10 | 74.1 | 30.0 | 44.0 | 30.1* | 41 | 7.2 |
| 100 – | 50 | 21 | 7 | 72.0 | 57.1 | 67.6 | 4.4 | 6 | • |
| *R&D department reported in NACE group 3* | | | | | | | | | |
| 10 – 49 | 218 | 360 | 108 | 57.3 | 7.4 | 26.2 | 31.1* | 54 | 1.7 |
| 50 – 99 | 60 | 141 | 15 | 76.7 | 13.3 | 32.2 | 44.4* | 58 | 3.6 |
| 100 – | 75 | 101 | 17 | 70.7 | 52.9 | 60.5 | 10.2* | 14 | 4.2 |
| *R&D department reported in NACE group 4* | | | | | | | | | |
| 10 – 49 | 270 | 791 | 328 | 51.5 | 10.1 | 20.6 | 30.9* | 60 | 1.9 |
| 50 – 99 | 99 | 273 | 46 | 60.6 | 10.9 | 24.1 | 36.5* | 60 | 3.0 |
| 100 – | 119 | 157 | 52 | 68.9 | 30.8 | 47.2 | 21.7* | 32 | 3.1 |
| *R&D department reported in manufacture* | | | | | | | | | |
| 10 – 49 | 573 | 1278 | 474 | 55.0 | 9.9 | 23.9 | 31.1* | 57 | 1.2 |
| 50 – 99 | 186 | 472 | 71 | 67.7 | 14.1 | 29.3 | 38.5* | 57 | 2.3 |
| 100 – | 244 | 279 | 76 | 70.1 | 38.2 | 53.1 | 17.0* | 24 | 2.2 |
| Total | 1003 | 2029 | 621 | 61.0 | 13.8 | 29.5 | 31.6* | 52 | 0.1 |
| *Innovation reported in NACE group 2* | | | | | | | | | |
| 10 – 49 | 85 | 127 | 38 | 72.9 | 15.8 | 38.7 | 34.2* | 47 | 3.0 |
| 50 – 99 | 27 | 58 | 10 | 81.5 | 30.0 | 46.4 | 35.1* | 43 | 6.8 |
| 100 – | 50 | 21 | 7 | 78.0 | 57.1 | 71.8 | 6.2 | 8 | • |
| *Innovation reported in NACE group 3* | | | | | | | | | |
| 10 – 49 | 218 | 360 | 108 | 65.1 | 9.3 | 30.3 | 34.8* | 53 | 1.7 |
| 50 – 99 | 60 | 141 | 15 | 86.7 | 13.3 | 35.2 | 51.4* | 59 | 3.4 |
| 100 – | 75 | 101 | 17 | 86.7 | 52.9 | 67.3 | 19.4* | 22 | 3.7 |
| *Innovation reported in NACE group 4* | | | | | | | | | |
| 10 – 49 | 270 | 791 | 328 | 60.7 | 9.1 | 22.3 | 38.5* | 63 | 1.7 |
| 50 – 99 | 99 | 273 | 46 | 69.7 | 10.9 | 26.5 | 43.2* | 62 | 2.8 |
| 100 – | 119 | 157 | 52 | 83.2 | 36.5 | 56.7 | 26.5* | 32 | 2.9 |
| *Innovation reported in manufacture* | | | | | | | | | |
| 10 – 49 | 573 | 1278 | 474 | 64.2 | 9.7 | 26.6 | 37.6* | 59 | 1.1 |
| 50 – 99 | 186 | 472 | 71 | 76.9 | 14.1 | 31.8 | 45.0* | 59 | 2.1 |
| 100 – | 244 | 279 | 76 | 83.2 | 42.1 | 61.3 | 21.9* | 26 | 2.0 |
| Total | 1003 | 2029 | 714 | 71.2 | 14.2 | 33.0 | 38.2* | 54 | 0,9 |

• approximation is too bad or impossible

The proportions considered in table 6.8 are results of CIS in two countries which have succeeded to reach a rather good response rate. In spite of this favorable basis not even the corner-stones of their results are comparable: Comparisons based on the results as given by the survey would be seriously misleading.

Tables 6.6 and 6.7 show that the bias due to non-response may be greater for more detailed results. For example, according to table 6.6 the proportion of service firms reporting innovation has an overall bias of – 10,1 % (see the last line), whereas the bias of this proportion for firms with 50 – 99 employees is – 25,7 %. It would be extremely hazardous to deduce substantial conclusions from these unreliable figures.

In some other countries the response rates are rather low (see table 6.4) and their results will suffer even more from errors caused by non-response. This can proved as follows: The absolute bias of the proportion $p$ is defined by

$$b_p := p - p_c$$

where $p_c$ is the proportion corrected for non-response defined above. With $q_r$, the response rate in the sample, the absolute bias of the proportion $p$ is

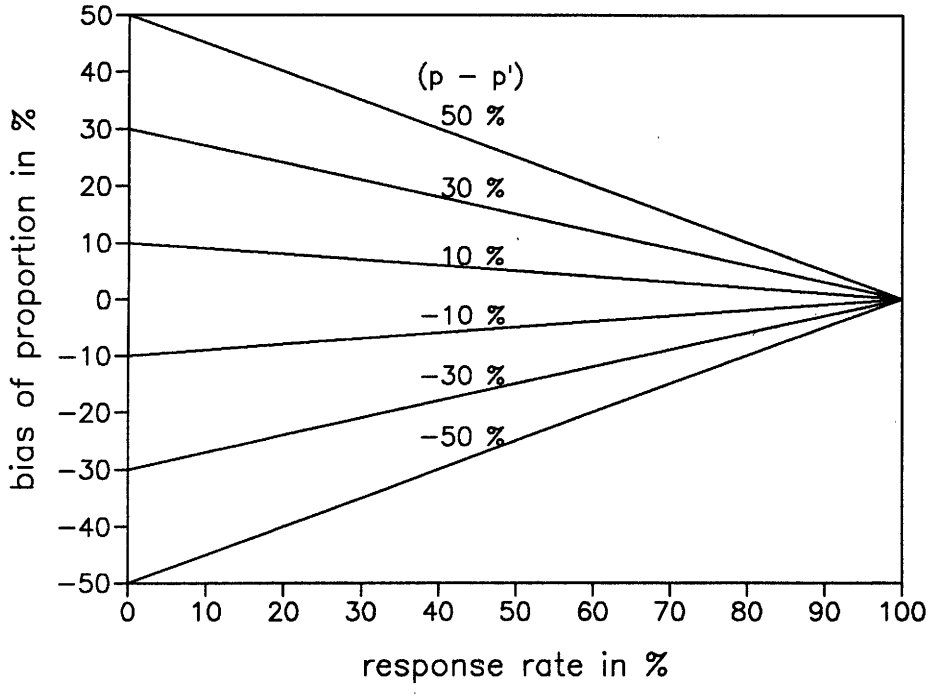
$$b_p = (1 - q_r) \cdot (p - p')$$

and the relative bias of the proportion $p$

$$b_p^{rel} = (1 - q_r) \cdot \left(1 - \frac{p'}{p}\right) \quad .$$

The implications of the formula for the absolute bias $b_p$ can be seen in graph 6.5 on the following page.

*Graph 6.5* *Dependence of the absolute bias of a proportion from the response rate $q_r$ and the difference $(p - p')$*




This graph shows that

1. the bias will be the more important the smaller the response rate $q_r$ is,
2. the bias will be the larger the larger the difference $(p - p')$ is between
   the proportion $p$ in the set of responding enterprises and
   the proportion $p'$ in the set of non-responding enterprises,
3. the direction of the bias depends on the sign of the difference $(p - p')$ .

According to tables 6.6 and 6.7

- the differences $(p - p')$ have opposite sign in the Netherlands and in Ireland,
- one half of the absolute values $|p - p'|$ of differences are larger than 10 % and more than a quarter of these values are greater than 30 %.

Lacking information for other countries it is not yet possible to say in which range the differences $(p - p')$ of proportions for responding and non-responding enterprises are lying. Results gained by the first Innovation Survey in countries, for which an analysis of non-response is not yet available, must not interpreted as if they were entirely free of any bias. Interpretations, which are implicitly based on this highly unrealistic assumption, are not only risky, but are not conscientious. The description of results, even more their use for inference, should rest on the safer assumption, that in countries with a response rate around 50 % (like Ireland and the Netherlands) proportions in half the cases are biased by 5 % and more. Corresponding proportions for two countries differing by less than 5 % must not be stated to be relevant, because such a difference may not reflect the reality but just an artifact caused by non-response.

### *6.4.4 Errors of total values due to non-response*

Because innovation surveys are focussed on quantitative characteristics, like innovation expenditures, it is extremely interesting to know something about the effect of non-response in the case of quantitative characteristics.

The mean value $\bar{x}$ of a characteristic $X$ in a domain of study can be estimated on the basis of the following information:

| | in sample | in subsample |
|---|---|---|
| response rate | $q_r$ | $q'_r$ |
| proportion of innovating firms | $p$ | $p'$ |
| mean value of innovation expenditures in the set of innovating firms | $\bar{x}_g$ | $\bar{x}'_g$ |

If $q'_r$ is less than 1, it is necessary to guess the rate $\lambda$ of non-responding enterprises in the subsample which have in the average the same innovation expenditures as the responding units in the subsample. The estimate of $\bar{x}$ is

$$\bar{x}_c = \frac{n_r \cdot p\,\bar{x}_g + n_{nr} \cdot [\lambda + (1-\lambda) \cdot q'_r] \cdot p'\,\bar{x}'_g}{n_r + n_{nr}}$$

and its relative bias

$$b^{rel}_{\bar{x}} = (1 - q_r) \cdot [\lambda + (1-\lambda) \cdot q'_r] \cdot \left(1 - \frac{p'}{p} \cdot \frac{\bar{x}'_g}{\bar{x}_g}\right) \quad .$$

The relative bias of a total value $\hat{x}$ for the characteristic $X$ in a domain of study is equal to the relative bias of the mean $\bar{x}$ for the same characteristic and domain: $b^{rel}_{\hat{x}} = b^{rel}_{\bar{x}}$ .

Comparing the relative bias of the total value $\hat{x}$ under the assumption $\lambda = 1$

$$b^{rel}_{\hat{x}} = (1 - q_r) \cdot \left(1 - \frac{p'}{p} \cdot \frac{\bar{x}'_g}{\bar{x}_g}\right) \quad .$$

with the relative bias of the proportion under the same assumption

$$b^{rel}_{p} = (1 - q_r) \cdot \left(1 - \frac{p'}{p}\right) \quad ,$$

one can see, that both are equal, if the mean $\bar{x}'_g$ in the subsample is equal to the mean value $\bar{x}_g$ in the sample. On the other hand, if the mean value $\bar{x}'_g$ in the subsample is smaller than the mean $\bar{x}_g$ for responding enterprises in the sample, the relative bias of a total value is larger than the relative bias of the proportion. Consequently, the relative bias of a proportion is just a lower bound of the relative bias of total values.

The level of the relative bias of proportions, shown in the last columns of tables 6.6 and 6.7, is really alarming: In most cases it is greater than 20 %. Consequently, quantitative results should be used and interpreted with extreme care.

## 6.5 Summary

### *6.5.1 Sampling techniques used*

Greece and Portugal deliberately restricted their frame to probable innovators, i.e. to a part of the target population to be covered by innovation surveys. Thus their contribution to the first CIS is weak and should be used only in analysis restricted to innovating enterprises.

Sampling methods used are also very different: In a few countries an adequate sample design has been established, whereas other countries have used a less satisfactory design, so that the accuracy of their results may be impaired.

Most countries did not make provision in the first instance for an investigation of the distorting effects caused by non-response. Some countries are performing or will perform such investigations.

No country using sampling methods has reported the calculation of standard errors, which are indispensable for sample surveys, because the inevitable sampling fluctuations of results must be duly assessed for their correct interpretation. These calculations will be done by Eurostat.

### *6.5.2 Comparability of results between countries*

Response rates are around 50 % in most countries, smaller in some countries and tiny in another one. The response rates in France, Greece, Luxembourg and Italy are satisfactory.

An evaluation of data provided by Ireland and the Netherlands, countries with response rate of about 50 %, shows that some key results of the first Innovation Survey are severely biased due to non-response. These results are not comparable in the weak sense (see definition in section 4.3).

Errors due to non-response in other countries are not known because either they have not at all investigated them or the investigations are not yet finished. The analysis of results in these countries should rest on the following assumptions, which are based on the experiences in Ireland and the Netherlands: The errors due to non-response

- in countries with response-rate around 50 % are nearly equal to those in Ireland and the Netherlands,
- in France, Luxembourg and Italy may be half as large and in Germany likely twice as large as the level in Ireland and the Netherlands,
- in Greece and Portugal cannot be gauged,
- in Great Britain cannot be gauged and are likely very large.

According to the definition of comparability of aggregated data (see section 4.3) not only results from Ireland and the Netherlands, but also results of the other countries cannot

be regarded as comparable in the weak sense. All the more they are not comparable in the strict sense (as defined in section 4.3), because standard errors are not calculated.

The first CIS has missed its aim to produce results which are comparable between all – or at least most – participating countries. If aggregated data of CIS are nonetheless used for some purely descriptive analyses, their quality should be kept in mind and results interpreted with utmost care. For policy-related advice these results may be used to assess, at a general level, questions like the non-R&D costs of innovation, the sources of and barriers to innovation, etc. Because comparability between countries is low the data must not be used for detailed analysis of differences between countries that result in policy advice on the (re)distribution of EU resources to countries or regions, or initiatives aiming at harmonizing structural or institutional factors across countries.

On the other hand, the first CIS has provided valuable experiences, how comparable results can be reached in further surveys (see recommendations in subsection 6.6.3).

### *6.5.3 Usability of results within countries*

The difficulties of the first CIS do not imply necessarily that data may not be used for analytical purposes within countries. As stated in chapter 1, empirical studies so far have been based on much less comprehensive data as those obtained by the first CIS. It remains a matter of judgment whether these data are usable to analyze structural relations between innovation, R&D and other characteristics covered by the questionnaire. However, it must be born in mind, that investigations on this basis may give results which are severely misleading. Such like results might be used as a pure description, but it must be warned to use these results as basis for detailed policy advice (cf. above).

It may be that sampling errors are smaller, if all non-innovating enterprises are dropped, i.e. explorations are restricted to the set of innovating enterprises (the same applies for enterprises engaged in R&D). For example, an analysis of inter-industry differences in the sources of innovation may give usable results, if the inevitable sampling errors are duly taken into consideration.

## 6.6 Recommendations

The experiences gained by the first Innovation Survey must be exploited systematically in order to improve statistical methods in further Innovation Surveys. The following recommendations should be considered.

### *6.6.1 General recommendation*

The Oslo manual does not treat several statistical issues which have to play a fundamental rôle in the realization of innovation surveys. It should be tried either to get enlarged the present manual or to fill in the gaps by producing a supplementary document, which describes all relevant procedures to be followed in future Community Innovation Surveys (CIS). Some important aspects are listed in sub-section 6.6.3.

### *6.6.2 Recommendations for the first CIS*

- Although standard errors are needed – at least for key results – to recognize the impact of their sampling errors, no country reported to have calculated standard errors. We recommend to ascertain means, variances and covariances within strata for some important characteristics. These statistics are needed
  - to calculate standard errors of selected results and
  - as a basis of an improved and harmonized sample design for the next CIS.
- Based on these measures a calculation of standard errors should be done for proportions, total values and means of selected charácteristics and also for relations of pairs of characteristics (e.g. innovation expenditure / turnover). We suggest to consider the accuracy of results both for all enterprises within a domain of study and for the subset of innovating enterprises. It should be checked, if results restricted to these enterprises are more accurate.
- Data gained by subsampling non-respondents have been reported up to now only in an aggregated form by two countries. This information is extremely valuable, but not sufficient to reach a general assessment, whether they can be used in an operational way, i.e. in order to adjust results for non-response errors. For this purpose detailed data are needed which cover supplementary information from the frame and the subsample. These data should be procured and analyzed in the hope that a reliable decision can be reached.
- The conjecture, that the method used for gathering information for non-respondents (by telephone or by written reminder) may affect the quality of responses, should be investigated. We recommend to give some money to countries for splitting their subsample of non-respondents in two parts of equal structure and trying each method in one part.

### *6.6.3 Recommendations for future CIS*

The main statistical aspects, which should be investigated in order to develop improved methods for future CIS, are sketched in the following list:

- The target population should be clearly defined at the very beginning of the planning stage. Because the Oslo manual states that

  *"it may be useful to include parts of the service sector",*

  a decision must be reached whether the service sector is to be covered by future CIS and which branches should be included. If this sector of growing relevance becomes part of the target population, the advantages and disadvantages to develop a special questionnaire should be explored.

- The survey should be split up into a main survey and a supplementary survey which is addressed only to a subset of enterprises in the main survey. The questionnaire for the main survey should be restricted to those questions which are most important and rather easy to answer; all other questions should be reserved for the supplementary survey. The general experience is: The shorter the questionnaire the higher is the response rate and the more reliable are the responses. Observing this experience seems to be the safest way to reach high response rates and to get rid of the main obstacle for comparability of results.

- All contractors should be urged to secure a satisfactory frame for the target population as defined by the Oslo manual. A deliberate restriction of the frame to likely innovating enterprises is not acceptable, because it hampers severely the comparability of results between countries.

- Domains of study should be defined, i.e. sets of enterprises, for which results with high accuracy are needed. This should be done before the sampling design is started: A precise specification of domains of study plays a fundamental rôle in the design and can prevent that some highly interesting results turn out to be not as accurate as they should be. The number of these domains should be limited according to the sample size and their boundaries set in such a way that results are really of main interest. It seems not advisable to equalize these domains with some prescribed level of NACE.

- The stratification for the sampling procedure should be harmonized and planned with respect to its feasibility and its efficiency. The sampling design rests on some measures which can be gained as a by-product if standard errors of results of the first CIS are calculated (see subsection 6.6.2).

- The allocation of sample size to strata should be harmonized according to the following principle: The more important results for a domain of study are, the higher their accuracy should be.

This can be reached if

- those parts of the population are specified for which results with high accuracy are to be realized (e.g. small and medium/large enterprises in domains which are not necessarily equal to a prefixed level of NACE),
- the desired relative accuracy of these parts is defined and
- allocation variables are prescribed on the basis of an analysis of data gained by the first CIS.

- We suggest to consider not only the variance of the most important characteristics, but also the propensity of non-response in the strata.
- Drawing the sample from each stratum should not be done by simple random sampling but by means of the so-called Deming-plan (W. E. Deming 1956), because this method is apt

  - to secure a random selection of units,
  - to utilize regional information (if available in the frame) and
  - to prefer the selection of ear-marked units (e.g. for a panel of enterprises).

- Common procedures for sending reminders and criteria for closing the field work should be developed.
- Methods for treating non-response should be investigated thoroughly in order to find out the best method for treating non-response in CIS (see the three–volume book of Madow and Olkin, 1983).
- On this basis a complete design for subsampling the set of non-respondents in the survey should be prepared and executed promptly, when the main survey has been declared finished. An investigation of non-response effects on the base of a well designed subsample is badly needed in all countries.
- Methods used to edit, to check and to weight data should be harmonized.
- It may also be advisable to define rules for field work. They should especially include rules which may prove helpful to assess non-response effects.

# Part III

# Conclusions and recommendations

# 7. Summary, conclusions and recommendations

The endeavour to develop innovation surveys has been characterised as a venture creating 'blue sky statistics'[1]: Starting from scratch, it aims at developing a new internationally harmonized source of statistics. The CIS project is an important step in this process since it is the first project based on a common conceptual background (the Oslo manual) and the first project to use an internationally agreed questionnaire[2]. Thus the CIS is the first major project aiming at developing internationally comparable innovation statistics.

Therefore it should be evaluated as a project constituting a step forward in an ongoing process and not as the ultimate design of innovation surveys. This entails that the evaluation should have a positive point of departure: This large-scale type of exercise has never been tried before; Where did the survey succeed? What can the data be used for? And what can be improved in future surveys?

These three questions are, in key words, what this concluding chapter will concentrate on. We will start out with a discussion of the policy background and the conceptual framework. Then, in sections 7.2 and 7.3 we will summarise the findings on implementation and statistical issues (covering both the EU level and the national level). Section 7.4 will present the main conclusions on the first round of the CIS, and in section 7.5 and 7.6 we will attempt to develop some 'best practice recommendations' for future internationally based innovation surveys. Finally, section 7.7 sketches some issues for future research.

## 7.1 Context

The history of the CIS can, very briefly, be summarised as follows. The CIS initiative started in 1991 and involved different phases. In a first phase of the project a model questionnaire for

---

[1] An expression introduced by Allison Young from STI/OECD.

[2] Previously only the Nordic countries have made a co-ordinated effort to make comparable data. However, in this project no harmonized questionnaire was agreed on (cf. chapter 2).

innovation surveys in Member States was developed by Eurostat, based on the Oslo manual and input from a Eurostat working party. In a second phase in 1991/92 this questionnaire was pre-tested in five countries (the so-called Kleinknecht project - see Kleinknecht, 1993). On the basis of this pre-test and a new Eurostat working party a revised questionnaire was developed by Eurostat. In a third phase this questionnaire was further developed by an expert group during the spring of 1992, and in April 1992 a draft harmonised questionnaire was presented at the OECD NESTI (National Experts on Science and Technology Indicators) meeting. The final harmonized questionnaire was agreed in June 1992 by Officials from EC and OECD and leading scholars in innovation surveys. In 1992/93 the national surveys were implemented and now the international data base is being constructed in Eurostat.

The CIS should, of course, be evaluated on the basis of this historical background and the specific context in which it came into being. The two succeeding sub-sections, following the structure of the report, assess the rationale behind the CIS and the conceptual context on which the CIS project was developed.

### *7.1.1 Rationale*

In recent years the political awareness of the crucial role played by technology in economic development has been rapidly increasing. Whereas, only 15 to 20 years ago, technology related policy instruments were relatively unknown and little utilised tools in economic policy, they have increased in importance and today they make up some of the core instruments in industrial policies in many Western countries.

These types of policies, initiated both from the national level and from the EU level, include programmes to initiate and promote R&D co-operation between firms and between firms and research institutions as well as programmes to initiate and promote R&D activities within single firms. Also included are programmes aimed at dissemination of public research results, initiatives to improve the technological infrastructure and programmes aiming at accelerating technological development in key industries or in key technologies, etc.

In these technology related industrial programmes there has been a strong tendency to focus on R&D and R&D aspects of technological development. This reflects, first and foremost, that R&D is the theoretically most examined aspect of technological development and the aspect for which statistics is most well-developed. This does *not*, however, entail that R&D is also the (single) most important contributor or even the determinant of technological development. Rather, new theory on technological development emphasises, *innovation and diffusion* as richer and broader concepts are what we should focus attention on (cf. section 7.1.2).

Thus, using solely existing data as a basis for the design of policies aiming at stimulating technological development entails a severe risk of employing less efficient instruments to less suitable target groups - or even a risk of employing wrong policies, resulting in retardment of technological change.

However, once we move from the 'safe' grounds of *R&D-based* technology policy to the broader *innovation - and diffusion based* technology policy we experience a serious lack of empirical knowledge that can guide our efforts. Data on innovation and diffusion consist mostly on indirect indicators, often developed for other purposes. This goes for, for example, patent statistics, balance of payments for technology and trade in high-tech products that are collected for legal or accountancy reasons. Other indicators, developed with the specific aim of collecting information on innovation, inform about different aspects (also mainly indirectly) of the process of technological development. This applies, for example, to bibliometric indicators and patent citations.

Even though much information has been retrieved from this data through refinement, extension and re-classification, much still has to be explored if technology policy should be build on a sound empirical basis. What we know now can be considered to picture only the 'tip of the innovative iceberg' but there is a hidden part which we must explore.

At the moment, direct firm-based surveys of innovation is the best possible method to shed light on this hidden part of this 'innovative iceberg' since this type of survey can supply information on e.g., innovative strategies, sources of innovation, barriers to innovation, innovative efforts and innovative results.

Furthermore, as is the case with most statistics, the value of this type of information is multiplied once it is comparable across countries or over time. For example, information on innovative efforts and innovative results are relevant in themselves but once it is possible to access whether national results are 'good or bad' or 'under or over average' compared to other countries or over time, the value of this data is immensely increased - then it is possible to access also the background information on this basis and this greatly increases the value of this data.

Thus, this new set of international data can, if quality and comparability prove to be satisfactory, provide vital information for the setting up of technology policy both at the national level and at the EU level.

### *7.1.2 Conceptual context*

Using solely information on R&D performance as the basis for technology policy has been heavily criticised in new innovation theory since it, implicitly, assumes 1) that R&D as an input in the innovation process is the main determinant of innovative activities and 2) a known (linear) relationship between R&D as input in the innovation process on the one hand and innovations as output on the other hand (both based on the so-called linear model of innovation).

'New' economic theory[3] instead emphasises that *innovation* is the key concept in economic development since it is the engine of a variety of phenomena, including growth, productivity, competitiveness and employment. However, innovation is a complex process, requiring a whole set of inputs and involving feed-back loops and trial and error processes and it is thus much more difficult to comprehend than R&D.

The most widespread and commonly accepted model to describe the complex innovation process is the Chain-Linked model presented by Kline and Rosenberg (1986). In short, this model insists on the interaction of market signals and technological opportunities for initiation of innovations and for the development from prototypes to final designs. Furthermore, it emphasises that feed-back links between the processes of development (and thus between the departments of the firm) are vital. Finally, the firm's knowledge base and research capabilities are not necessarily seen as initiators but are seen as vital to the whole process of innovation, functioning as input to all stages of the process.

It is this conceptual frame which forms the basis for the Oslo manual (OECD 1992), which, in turn, is the conceptual basis for the CIS. Two aspects of this framework have been discussed in Chapter 3. First, is the Chain-linked model appropriate or sufficient as the conceptual basis for innovation surveys, or should it be further developed? And second, how accurate is the connection between this conceptual background (the Oslo manual) and the CIS questionnaire?

It is concluded that the Chain-Linked model is the most adequate model developed so far to picture the innovation process and thus the best possible point of departure for innovation surveys. Nevertheless, the Oslo manual should be developed to discuss in more detail the limitations of the model and the possible expansions. A series of relevant subjects are discussed in Chapter 3.

Further, it is concluded that the harmonized questionnaire is reasonably well adapted to the Oslo manual, particularly in some key topics. However, also here a series of suggestions for developments are presented for future innovation surveys. For example, the CIS is based on the subject approach (focusing on the behaviour of firms) and this should be made more explicit in the questionnaire. The questions on the intensity of innovation should be restated in order to avoid that less firms than what is actually the case declare themselves innovative.

Finally it should be emphasised that the Oslo manual should not be (and it has not been in the CIS project) considered to be a 'bible' for innovation surveys, providing all the correct questions and answers. Rather, the Oslo manual is to be seen (as have been discussed at several OECD and EU workshops) as a first step in the development of a conceptual basis for innovation studies. Therefore, it was set as a priority in the evaluation, both for coming revisions of the Oslo manual and for the design of future innovation surveys, that the Oslo

---

[3] 'New' is put in inverted commas since the theories putting innovation in centre of the analysis of economic development to a large extent builds on Schumpeter's works from 1912 and onward.

manual should be critically assessed. Therefore, the discussions in Chapter 3 are equally relevant for the revisions of the Oslo manual, and it may be viewed as an opening of this process.

## 7.2 Implementation issues

How the CIS were implemented at the national level to some extent depended on the guidance that the national contractors were offered by the Commission. Thus, in section 7.2.1, we summarise the role played by the Commission in the implementation of the CIS (examined in section 5.2), and thereafter (section 7.2.2) we study implementation at the national level (analysed in section 5.3-5.5).

### *7.2.1 Implementation at the EC level*

The process of initiating and developing the CIS was not an ideal one in which a series of different phases (creation of a conceptual framework, creation of a common questionnaire, development of guidelines for implementation and sampling, implementation across member states, creation of a common database) followed each other consecutively and at a moderate pace.

Rather, due to several factors, the project was delayed. First, because of the pilot-character of the project, several difficult issues had to be investigated in depth, e.g. the possibility of measuring innovation costs. Second, because of the international character of the project the Commission had to negotiate with national contractors that, a priori, had very different experiences and expectations. Third, because of the aim of making the CIS survey comparable also with surveys in non-EU countries there were negotiations with non-EU OECD countries. These factors meant that the project was delayed. Some member countries did not await this development but wanted to implement nationally developed surveys (however, the surveys were based on the Oslo manual and the draft harmonized questionnaire). These processes entailed that the project became difficult to control and co-ordinate and that the process became partly self-driven, timing being unsatisfactory.

Moreover, because no legal basis exists for collection of innovation data, the Commission was unable to impose any demands on member states and were limited to compile a list of implementation and sampling issues which was presented to member states as *recommendations*. Thus, the Commission was highly dependent on member states' co-operation in the first CIS project.

However, it is the opinion of the evaluation team that the recommendations supplied by the Commission were too general and did not cover all relevant aspects. Even though still voluntary in nature, more detailed recommendations on sampling and implementation could probably have served as guide-lines for some of the more inexperienced national contractors.

### *7.2.2 Implementation at the national level*

In the following the findings on key aspects of implementation in member states and Norway will be summarised. This will be part of the basis for the recommendations presented in sections 7.5 and 7.6.

*Contractors*

In half of the countries the national contractor was the national statistical office. In the majority of the other cases the surveys were organized at the Ministerial level or by other state or semi-state agencies. When surveys were implemented by national statistical offices they were able to make use of the official data register of enterprises as sample frame. In other cases contractors were forced to work with frames that may be less satisfactory, but only in three countries (Greece, Portugal and United Kingdom) the sample frames can be considered to be incomplete (in Greece and Portugal because they deliberately restricted the frame to cover only probable innovators and in the United Kingdom because the contractor did not have access to a high-quality frame).

In many countries contractors collaborated with (other) research institutes or benefited from advice from external advisers (in some cases in ad hoc working parties). It has proved to be the case that previous involvement in technology-related business surveys and/or intensive interaction with experts have been an advantage for national contractors in carrying out the survey.

*Nature of the surveys*

Except in Ireland, Italy and Luxembourg (with Spain representing a special case) the surveys were sample based. In most countries all enterprises above a certain size were included. However, in the sampling procedure a variety of different methods has been used, notably in Greece and Portugal where the sample frames do not follow the recommendations in the Oslo manual.

The participation in the surveys has been voluntary in all countries except Italy and Portugal. Only in these two countries exist legal frameworks that allow national contractors (i.e. national statistical offices) to make mandatory surveys.

In all countries except Greece and Luxembourg (and partly Ireland) the survey was made as a postal survey.

*Coverage of the surveys*

With the exception of the Netherlands (which used principal establishment) and the United Kingdom (which used autonomous enterprise) all countries used the enterprise as a legal unity as the survey unit.

Five countries have included (part of) the service sector.

In several countries a cut-off point of 20 employees has been applied. In some small countries (Belgium, Ireland and the Netherlands) a cut-off point of 10 employees has been used - in Norway a cut-off point of 5 employees was applied. Germany represents a special case since all enterprises with more than 4 employees were in the frame. Greece and Portugal did not apply any cut-off point since their target population included all innovative enterprises.

*Classifications available*

Since classification systems and levels should be the same over countries to secure harmonized sampling methods and to make possible cross-country analysis, this issue was also considered in the evaluation.

All EU countries have used the NACE classification for branches, however, not on the same level of aggregation and furthermore, some countries have used nationally adapted NACE versions (or NACE 70). For sampling a two digit NACE level is sufficient but for analysis it is preferable to have assess to 4 digit NACE to facilitate disaggregated branch analysis (e.g. analysis of pharmaceuticals or computer equipment). Most countries have used at least 4 digit NACE but a few countries (Greece and Spain) have used only 2 digit NACE. Norway has used ISIC classification which, with some effort, can be translated into NACE.

As a measure for firm size all countries have used number of employees. In Belgium and the Netherlands they use 9 and 6 employment classes but in all other countries employment was measured as a continuum.

Some kind of regional disaggregation has been used in most countries. Several small countries expressed that the regional disaggregation should not be used for small countries since headquarters (and R&D departments) often are located in or near the economic centres (= the metropolis) whereas production units may be located in the low-cost areas of the country.

*Pre-tests*

The pre-test has proven to be a very valuable step to improve the design of the questionnaire and to sharpen the whole survey. A first pre-test in five European countries (Great Britain, Norway, Denmark, Germany and the Netherlands) was sponsored by Eurostat (for a synthesis of the results, see Kleinknecht, 1993). However, the conclusions of this pre-test are not always consistent with what emerged during the survey, possibly also because the questionnaire was further modified after the pre-test. The most notable case is represented by the question on the life cycle of the product, which appeared to be well understood by respondents during this pre-test while it was one of the questions with the highest item non-response in the real survey.

Some countries which did not modify the harmonised questionnaire (Belgium and Luxembourg) have not done any additional pre-test. Other countries which were involved in the

European pre-test, such as Denmark and Norway, did not either engage in additional pre-tests. Very detailed pre-tests were carried out in France, Germany, Ireland and Italy (see column 4 of table 5.2). In some cases these pre-tests dramatically changed the design of the questionnaire. The most remarkable case is France, were the organiser of the survey (SESSI) drastically simplified the harmonised questionnaire and dropped all controversial questions.

*Design of the questionnaire*

According to contractors, the key concepts (notably innovation and R&D) were clearly defined in the questionnaire. However, a problem raised by a few contractors was that some respondents were confused about the relationship between R&D and innovation. A few contractors mentioned the possibility of entering more (and more relevant) examples in the definitions.

Apparently there was no serious systematic bias in problems of comprehension. However, some countries (Belgium, Italy, the Netherlands, Norway and Spain) mentioned that some problems of comprehension occurred in traditional sectors where concepts like technology, innovation, design and even R&D sometimes caused confusion. Thus, the questionnaire seems to be more well adapted to medium- and high-tech sectors.

In some countries respondents had been confused by the fact that part of the questionnaire refer to the period 1990-1992 whereas other parts refer to 1992 only.

The question on innovation costs (Q. 13) was the question which caused most (both conceptual and statistical) difficulties (see also section on item non-response). National contractors had different views on how the question should be interpreted (as discussed in detail in section 5.4.2) and this sometimes led to changes in the formulation of the question and in one case (France) the contractor even decided to drop the question. This should be borne in mind when comparing this information across countries. Further, the question on life cycle of products (Q. 14) caused problems for many respondents.

A general complaint from national contractors was on the length of the questionnaire. However, this problem was intensified because most countries included additional questions (as described in the next section). A few countries (e.g. Germany and Norway) collected information on the time needed to fill in the questionnaire and they reported an average response time of app. 1/2 hour for non-innovators and app. 2 hours for innovators - varying of course depending on the size of the responding firm. In Greece and Luxembourg where direct interviews were carried out, the interviewers recorded the time needed to complete the questionnaire. The minimum time required was 30 minutes and the maximum 1 1/2 hour.

*National adaptations*

A detailed analysis of the comparability of national questionnaires has been performed by Eurostat (Eurostat 1994) so the evaluation concentrated on the type of differences introduced and why they had been introduced.

Even though all contractors felt that the harmonized questionnaire was too long, all countries except France either adopted the questionnaire unchanged (Belgium and Luxembourg) or added one or more questions.

In some cases the reason for differences in questionnaires was the timing. In Greece and Spain the survey was planned before the harmonized questionnaire was adopted. In most cases, the reason for national modifications was the desire to include some kind of background information. This was particularly the case in countries where the contractors were not the national statistical offices (which has a lot of background information on the enterprises from other surveys). Italy and the Netherlands included a few questions to secure comparability with former innovation surveys.

Germany customized the questionnaire to three response groups: 1. small manufacturing firms (< 50 employees), 2. medium sized and large manufacturing firms, 3. service firms. The Netherlands and Italy customized the questionnaire to two response-groups: 1. manufacturing firms, 2. service firms. In both countries this was done to make the questionnaire more understandable and to reduce the response burden, and thus increase the response rate and secure more reliable responses.

Only one important term, namely 'incremental' caused problems in translation of the questionnaire. This was the case in several countries.

*The participation*

The response rate varied very much across the participating countries. In countries where the survey was mandatory it was high (Portugal app. 70% and Italy 64%). In Luxembourg and France the response rate was 79% and 74%, respectively. In other countries it varied from about 50% for several countries to 23% in Germany - United Kingdom being a notable exception with a response rate of only 4%. In most countries contractors judged the response rate as 'standard' compared to other business surveys.

A series of devices (described in detail in section 5.5.1) were employed to induce respondents to participate in the survey. For example high-level supporting letters, a long series of reminders, telephone calls, customization of questionnaires, collaboration with industry associations, promise of results from the survey. Many of these devices are country-specific and cannot be transferred to other countries.

Who (i.e. which type of manager: general manager, R&D manager, marketing manager, etc.) should be selected as the 'best' respondent is not entirely clear. Therefore one of the most

important actions which can be taken to secure high response rate and reliable answers is to pre-identify the relevant respondent in each participating enterprise. This pre-identified respondent may not be the person to fill in the questionnaire, but it should be the person empowered to ask other persons to fill it in. Experience in several countries shows that often the questionnaire is circulated in several departments (most often 2-3 departments) in the enterprises to secure reliable information is supplied (this is, of course, especially the case in large enterprises).

Not only the quantity of results but also the quality of results is important, and across member states a variety of actions were taken to secure the quality of responses. Most countries have made checks on plausibility and compatibility on all incoming questionnaires, following up discrepancies by telephone immediately. Furthermore, most countries have phoned firms and asked for any kind of missing information. One exception from this is Denmark where this was done only in very few cases.

*Problems of confidentiality*

On the basis of their contact with respondents no contractors regarded concern about confidentiality as a problem for the realisation of the CIS. In the cases of statistical offices or government agencies contractors often attributed this to the acts of statistical secrecy.

In general national contractors did not believe that special questions were seriously affected by confidentiality problems, neither did they find that special size groups, branches, etc. were more concerned about confidentiality than others.

However, since concern about confidentiality may not only cause item non-response but also cause total non-response, this may contribute to the (sometimes high) non-response. Whether this could be the case should be investigated closer in a separate project.

*Reliability of the responses*

The statistical reliability of the data is considered in the next section. The type of reliability discussed here is the reliability in the individual responses. The conclusions presented here is based on contractors' follow up on item non-response and their plausibility checks.

The majority of the contractors judge the quality of the responses as either standard or good - the quality of responses was often compared to the quality of responses to the R&D survey and in this comparison the innovation data was judged to be moderately less reliable. In the Netherlands the quantitative questions included the possibility to tick whether the information supplied was 'fairly accurate' or 'a rough estimate' or whether it was 'not possible to answer'. The conclusions from this analysis have not been published and setting up a small project to assess these results may be advisable.

In one country (Denmark) we received copies of a 'set of double responses' - i.e. a case where the original questionnaire was returned at the same time as a reminder with a new questionnaire

was sent out, and the second questionnaire also was completed. This example shows considerable response variability (a well-known problem in questionnaire or interview based surveys), and this should be taken into account when assessing the reliability of single responses. An analysis of all 'double responses', i.e. response variability, at the European level might be set up to evaluate the problem thoroughly.

It is our feeling that the quality of the data was slightly better in countries with a 'tradition' for innovation surveys because of the 'learning effect' described earlier and in section 7.6.3).

## 7.3 Statistical issues

*Sampling frame and coverage of frame*

The frames used across the member states were different in their origin and coverage. In countries where the survey was carried out by (or in close co-operation with) the national statistical office the sampling frames were the official directories of business firms and the coverage was nearly complete. In other countries a variety of other sources have been used but frames and coverage are generally less precise and satisfactory. However, the frames are acceptable in all countries, possibly with the exception of the United Kingdom. In this case the quality of the sampling frame is unknown and should be checked.

Greece and Portugal have used sample frames different from what has been used in all other countries (both countries have concentrated on probable innovative firms only), and this prevents cross-country comparisons on the level of the whole industry with these two countries. However, studies concentrating on innovative firms only may use also the data from Greece and Portugal.

*Information contained in the frame*

Since the information contained in the sample frame, especially information on size and branch, should be used for stratification of the sample it is important that this information is in the frame and reasonably up to date.

Although the information was not always up to date, the necessary information was contained in all sampling frames and seemed to be of a high quality. This was especially the situation in cases where the contractor has access to an official register of enterprises.

*Sampling methods and stratification*

Five countries decided to make census surveys: Ireland and Luxembourg made census surveys because of their limited size. The tradition in Italy is to make census surveys and the Italian contractor decided to make a census survey also in this case. Greece and Portugal made census surveys inside their deliberately restricted frames.

In the remaining countries some kind of sample survey was implemented. In Belgium, Denmark, Norway, Spain and United Kingdom sampling was done as pure random sampling, whereas the contractor in Germany stratified according to eastern and western parts and applied systematic selection to secure a fair representation of regions. The contractor in the Netherlands modified the random selection in order to include as often as possible firms included in previous innovation surveys.

In all countries except three employing sample surveys, branches and employment classes were used for stratification. The three exceptions were Germany where also a regional aspect was used (as described above) and Norway and the United Kingdom where only employment classes were used.

*Sampling fractions and sample sizes*

The countries implementing sample surveys have used various sampling fractions for different strata, as listed in section 6.2.4.

Average sampling fractions vary between 11% and 12% for Germany and the United Kingdom and 43% in Denmark (the sampling fraction being 100% for countries with census surveys). The average sampling fraction is (and should be, according to textbooks on sampling) higher for countries with small frames than for countries with large frames.

*Reminder procedure*

All countries have used reminders to increase the response rate. There were, however, big differences in the intensity of this follow up. In the United Kingdom only one written reminder was sent out, in most countries (for example Denmark and the Netherlands) two to three written reminders were sent out. A few countries were more thorough and combined a series (3 to 4) written reminders with telephone reminders (the case in Ireland and in Germany for large enterprises).

*Response rates and impact of unit non-response*

In a new venture as the CIS, employing a big and complicated questionnaire, it must be expected that non-response can be high - in effect there is a trade-off between the length and the complexity of the questionnaire (and thus the amount of relevant information) and the response rate. In most countries (the exception being France) it was decided to use a long questionnaire with complex questions and thus low response rates may be expected. This has, however, been the case only in a minority of the countries. However, as stressed in the last part of this subsection distorting effects of non-response may nevertheless be serious.

Table 6.4 reports the realized raw response rates in the all countries. In 5 countries the response rate lay in the interval between 64% and 79% and must be judged to be (almost) satisfactory

(cf. below). 4 countries had around 50% response rate and the response rate in the three Belgian regions were between 35% and 45%.

In Germany the response rate was just 22% and in Spain where the survey was cancelled shortly after mailing the questionnaires a response rate of 12% was reached. The response rate was just 4% in United Kingdom and this is clearly unsatisfactory.

According to a 'rule of thumb' results in sample surveys with response rates below 75% will be distorted by errors due to non-response. Evaluating two non-response surveys (Ireland and the Netherlands) it is shown (section 6.4.3 and table 6.8) that a non-response of 50% can severely bias the key-results of the first CIS.

Thus results of non-response analysis that has been or shall be implemented should be analysed in detail. Absence of such results for some countries must not be interpreted as if the results were free of biases caused by non-response. Furthermore, data on non-response may be useful in the weighting procedure to increase the reliability of the data (cf. later in section 7.5).

As long as no results have been reported for other countries, the analysis of the Irish and Netherlands data suggests the following assumptions: Errors due to non-response

- in countries with response rates around 50% ((Belgium) Denmark and Norway) may be nearly equal to those reported in table 6.6 and 6.7 for the Netherlands and Ireland
- in France, Italy and Luxembourg may be half as large and in Germany twice as large as the level in the Netherlands and Ireland
- in Greece and Portugal cannot be gauged
- in Spain and the United Kingdom are likely to be very large.

The consequences of this for comparability is discussed in section 7.4.

*Item non-response*

Also, one might expect that questions that are especially complex or are introducing new concepts to respondents may suffer from high item non-response. This has proven to be the case in the CIS for especially two questions.

In all countries a high rate of item non-response were reported on the question on innovation costs. Several explanations for this were provided by the national contractors: 1. The question asked for types of costs which normally cannot be retrieved through the normal accounting systems. 2. The question asked quantitative information and response to quantitative information is normally more difficult to retrieve. 3. The terms used in the question may have been difficult to understand since many respondents were not familiar with them. 4. The question was fairly detailed and looked complicated.

Another question with high non-response in all countries was the question on sales according to phases in the life cycle. Almost unanimously national contractors said that respondents had

reported that this question had been difficult to understand and relate to their turnover which was made up by a whole range of products in different stages of their life cycle.

Also the question on sales from new or improved products had a relatively high item non-response. Supposedly because this way of distributing the turnover is unfamiliar to respondents.

Most countries reported to have made serious attempts to reduce item non-response by telephone follow-up.

All national contractors shared the view that there is a learning effect which can be utilised in coming surveys so item non-response gradually will fall and the response rate will rise.

*Plausibility and compatibility of data*

All countries have made plausibility and compatibility checks of the data.

*Assessing the accuracy of results*

To assess the accuracy of the results of the first CIS we need to know the sampling errors. However, till now sampling errors have not been assessed for any country - this will be done by Eurostat, though.

Since we have no precise knowledge of the accuracy of the results of the first CIS, section 6.4.1 and 6.4.2 using a standard example, show that the sampling errors can be considerable and should be calculated and taken into account when the data is used for analysis.

## 7.4 Conclusions

It is the opinion of the evaluation team that the realisation of the first CIS has been successful in its first aim. From this pilot action the Commission, the national contractors and scholars in innovation measurement have learned a lot which can secure a high quality of a possible next CIS (or any other national or international innovation survey).

Regarding the second aim of the CIS - collection of comparable innovation data across EU member states - the level of success is more moderate. Data is not comparable across all countries and all variables (cf. below). However, it is the opinion of the evaluators that this aim - given the conditions under which the project came into being and was developed (especially in view of the novelty involved and the international character of the project) - could not be reached. Under these circumstances it is an achievement that a new data source which may be used for some types of analysis of innovation in Europe (cf. section 7.4.3 and 7.5) has been created.

If we make a 'relative' assessment of the degree of success in the CIS - i.e. an assessment of the CIS compared to other types of data on technological development we may conclude that the CIS has been more successful. In one venture the CIS project has gathered and - with this report - disseminated much information on the 'field-methodology' of innovation surveys compared to what has been collected and disseminated for other types of surveys of technological development. Furthermore, there are also problems with the international comparability of these types of data even though, in several cases, they have been collected for many years. Thus, in a 'relative light' the CIS has come far in its first year.

### *7.4.1 Overall quality of the realized CIS data*

We concluded above that it should not be expected that the CIS would reach full international comparability of data in its first attempt and in the following we shall, to deduce as many lessons as possible for future innovation surveys, assess critically and in detail the quality of the realized CIS data and the problems in the realisation of the first CIS.

According to the definition presented in section 4.3 the following conditions should be fulfilled for aggregate data to be comparable in the strict sense:

- Questionnaires should be comparable at least in the weak sense (i.e. questions included in the questionnaires should correspond to each other exactly).
- Data records should be comparable (i.e. the questionnaires should be at least comparable in the weak sense and the measurements should be reliable and valid).
- For each data record the correct raising factor should be calculated and the raising factors should be taken into account during the aggregation of data records.
- Some margins of errors should be estimated and used for assessment of results.

On the basis of these definitions the results of the first CIS cannot be regarded to be statistically comparable between all countries in the strict sense. This does not imply that the data cannot be used for analysis, but it implies that the analytical possibilities are restricted (cf. section 7.4.3).

It should be emphasised that since, up till now, neither Eurostat nor member countries have calculated margins of error and only few countries till now have reported information on non-response, we have no real knowledge of the quality of the realised data across countries. Thus we do not know whether data from some countries are comparable. These problems will, however, be properly assessed by Eurostat in building up the EU database.

What we do know about the quality of the data at this stage is

- that data from Greece and Portugal are incomparable with other countries at the level of the total industry because they have deliberately restricted their sample to probable innovators (cf. later)
- that data from the United Kingdom is not comparable because of the unsatisfactory realisation of the survey
- that data form Ireland and the Netherlands are biased due to non-response.

What we do not know is

- whether data from Ireland and the Netherlands can be made comparable with other countries if results from their non-response analysis are utilised in the weighting procedure and margins of errors are calculated
- whether data from the remaining countries are comparable since we have no information on either non-response or margins of errors. These measures should be assessed to decide whether data for theses countries are comparable and information on non-response should be utilised in the weighting procedure.

  In particular, data from Spain should be thoroughly assessed to decide whether the cancelling of the survey has caused severe biases. In Germany the magnitude of errors caused by the low response rate should be carefully assessed.

However, having no real knowledge on the precision of results must not mean that we ignore that problems may be severe. Thus, until these margins are calculated results should be interpreted with utmost care.

### *7.4.2 Problems with the realisation of the first CIS*

In our view the main factors to account for the lack of comparability (cf. the definitions presented in chapter 4) are:

- Some contractors modified some questions and thus questionnaires are not comparable in the weak sense between all countries (cf. Eurostat 1994).
- The survey frame was not comparable for two or three countries (Greece and Portugal (and the United Kingdom)).
- The sampling methods were not sufficiently harmonised.
- High levels of total and/or item non-response occurred in some countries.
- Raising factors have not been calculated for all countries and used in the aggregation of data records (this will be done, however).
- Margins of errors have not been calculated for any country (this will also be done).

The main factors, in turn, to account for these deficiencies are:

- Lack of co-ordination.
- Lack of instructions on sampling and implementation.
- Lack of awareness of the importance of international comparability.

Again, however, it is possible to go back one more step and assess the reasons for these problems, and this can, in our opinion, be summarised into six points:

1. The international character of the project. The Commission had to negotiate with national contractors that, a priori, had different experiences and expectations. Furthermore, since a key-priority for the CIS was to make the data comparable also with other OECD countries negotiations took place with the OECD. These processes delayed the project and some member states did not await the results of this endeavour but implemented the survey before a final harmonised questionnaire was agreed on. Therefore both timing and harmonisation was retarded.

2. Lack of co-ordination power. Since no legal basis exists on collection of innovation data the Commission was compelled to make only recommendations on sampling and implementation to national contractors. They could make no demands on the services to be rendered by national contractors and they could not pick the best possible national contractors (cf. the discussion of 'best practice' national contractors in section 7.6.2). This seriously hampered harmonisation and in some cases influenced the quality of data.

3. Lack of advice. The set of recommendations worked out by the Commission was not sufficiently detailed. Even though still voluntary in nature the recommendations could have been more itemised, providing detailed advice to some of the more inexperienced national contractors. This may have reduced the quality of data for some countries.

4. Lack of will. Even though all national contractors agreed on the importance of creating an internationally comparable data base of innovation statistics, some of the national contractors did not seem to have the will to comply with this aim of the project. Therefore, they introduced various national-specific changes and this hampers the comparability of the data.

5. Lack of expertise. It seems that in a few countries national contractors did not have the full economical or statistical expertise to carry out the innovation survey in a satisfactory way. This hampers the quality of the data for these countries.

6. Lack of comparison of experiences. Too little was done to facilitate an interactive learning process where national contractors could learn from each other (best practices, errors, difficulties, etc.).

### *7.4.3 Analytical possibilities with the CIS data*

The problems on comparability presented above do not imply that the CIS data may not be used for analysis if margins of errors are taken duly into account. Three different uses of the CIS data may be envisaged:

1. Descriptive analysis of differences between countries.
2. Analysis of innovation in selected industries across countries.
3. Analysis of innovative structures within countries.

Within each of these uses a variety of projects may be performed. However, it is not a task for this evaluation to go into detail about possible analysis of the CIS data since this must depend on the needs of the Commission (to some degree future analysis have been described in EIMS/ Eurostat, 1994).

Since sampling errors may be smaller if all non-innovating enterprises are left out of the analysis, i.e. if explorations are restricted to the set of innovative enterprises, such analysis may provide more manifest results. In these cases, also the data from Greece and Portugal may be used.

Before engaging in analysis the Commission should assess the reliability of single variables across countries as a sort of users guide to the data base considering which variables can be compared across which countries (phase II of the evaluation of the CIS).

The CIS data should not be used for assessment of EU totals. For example assessment of total innovation costs in EU or the share of turnover used for innovation across EU cannot be made because of deficiencies of the data for some countries. For policy-related advice these data may be used to assess areas like the non-R&D costs of innovation, the sources of and barriers to innovation, R&D co-operation and innovative strategies, etc. Because comparability between countries in some cases is low the data should not be used for detailed analysis of differences between countries that result in detailed policy advice on the (re)distribution of EU (research) resources between countries or regions, or initiatives aiming at harmonising structural and institutional factors across countries.

## 7.5 Recommendations for the first round of the CIS

In this section two actions are recommended which, we believe, could enhance - or at least assess properly - the quality of the data from the first CIS.

*Subsampling for non-response.*

By now sub-sampling for non-response have been performed in some countries, and if the relevant calculations are made this can help assess the quality of the data for these countries (cf. section 6.4.3 and 6.4.4). However, non-response analysis for the remaining countries (except Spain and the United Kingdom) should be performed.

Furthermore, since it is probable that the information on non-response can be used in an operational way in order to adjust results for non-response errors i.e. to enhance the quality of the data, a special project aiming at deciding whether this is the case should be set up immediately (the data assessed in this report were not sufficient to perform this analysis).

*Calculation of standard errors*

Calculation of standard errors has not been performed for any country even though, in order to assess the inevitable sampling errors, this is an indispensable part of a sample survey.

We recommend to ascertain means, variances and covariances within strata for some important characteristics as a basis for calculation of standard errors. Standard errors should then be calculated for proportions, total values and means of selected characteristics and also for relations of pairs of characteristics (e.g. innovation expenditure/turnover).

We suggest that this is done also for the subset of innovative enterprises to assess whether results for this subset of enterprises are more accurate.

## 7.6 Recommendations for future innovation surveys

These recommendations are based partly on the evaluation of what have been done on both EU level and national level and partly based on the subsequent statistical analysis of some of the realized samples in the first CIS.

However, also the previous experience of the evaluation team will, naturally, enter the recommendations - we cannot pretend to be unbiased by what we knew and thought about innovation surveys in advance of this evaluation.

### *7.6.1 General recommendations*

The evaluation has shown that co-ordination at the European level is essential if comparable results across countries are going to be created. Innovation surveys are a new initiative and are not yet backed by a solid experience as data on the main economic indicators. A common standard is still searched and to achieve this requires close co-operation among the various organisations which are involved in the field. It is the opinion of the evaluation team that such a satisfactory co-ordination can only be achieved through a legal basis and consequently it is recommended that a legal basis for innovation surveys is adopted. The evaluation of the first CIS points to that on the basis of a legal basis the Commission should:

- Co-ordinate the venture i.e. make sure that timing is appropriate.
- Pick the best possible national contractors (or contractor teams cf. section 7.6.2) i.e. contractors with relevant experience.
- Create a revised pre-tested questionnaire to be implemented in all countries.
- Work out detailed instructions of all aspects and levels of the implementation of the surveys on the national level, as displayed in Table 7.1 on the next page.

*Table 7.1 Implementation and statistical issues for future innovation surveys*

| | | |
|---|---|---|
| *Implementation issues* | Target population | Branches covered |
| | Contractor | |
| | Cut off point | Number of employees |
| | Timing | Time of start<br>Time of closing |
| | Reminder procedure | Number of reminders<br>Timing of reminders |
| | Follow up on responses | Quality checks of responses<br>Procedure for follow up |
| *Statistical issues* | Survey method | Questionnaire/interview |
| | Frame | |
| | Survey unit | |
| | Sample size | |
| | Sampling technique | Sampling method<br>Stratification method<br>Sampling fractions |
| | Subsampling for non-response | Sample size<br>Sampling method |
| | Imputation of missing data | Method |
| | Raising factors | Method |
| | Assessment of reliability | Estimation of standard errors<br>Estimation of bias of results |

These points are elaborated below.

In case these issues are too detailed to be included in a legal basis it is recommended that the legal basis is made as a frame which can be filled in by the Commission. It is recommended that the basis for creating the legal basis is the experience from this first round of the CIS.

### *7.6.2 Specific recommendations*

Which recommendations should be given to the issues in Table 7.1 will be discussed in the following, based on the experience from implementation at the national level. In other words, in this section we will attempt to develop some 'best practice' recommendations for future versions of the CIS.

*Target population*

It was not the intention in the CIS to cover the service sector. However, a few countries included the service sector in their target populations and it should be decided in the very beginning of a future CIS whether services are to be included, and if it is, which branches of the service sector should be included.

Since this issue was not a part of this evaluation we recommend that a project is set up with the specific aim of assessing the experience in these countries in order to decide whether (and how) services should be included in future surveys. If such a project recommends that services should be included it should also explore the advantages and disadvantages of creating a specific and customised questionnaire for the service industries.

Furthermore a decision has to be made on whether the target population is industry (and services) as a whole or whether it should be confined to the sub-set of innovating enterprises. Since it is vital to be able to assess national totals, it is highly recommended (following the Oslo manual) that surveys cover both innovative and non-innovative firms.

*National contractors*

The national contractors play a central role for the realisation of the surveys and thus for the quality of the national data. Therefore efforts should be made to pick the best possible national contractors or contractor teams. 'Best' means that national contractors should fulfil three requirements: 1. They should have access to a (sampling) frame of high quality (cf. later). 2. They should have statistical expertise in carrying out business surveys and 3. They should have the economic qualifications to carry out surveys on technological development. Furthermore, since the value of innovation data is multiplied if it can be combined with information on e.g. firm growth, productivity and investment it would be preferable if national contractors had access to this type of data on firm level.

Since these requirements may be hard to meet for a single contractor co-operation between the national statistical offices - that have the best available sample frames, the experience in carrying out business surveys and other relevant firm-level business statistics - and some national institutes specialised in the economics of innovation may be envisaged as the optimal solution.

*Pre-tests*

Pre-tests could play a greater role in improving the design of the questionnaire and the methodology of the survey. In future surveys, three different stages should be planned to fully exploit its lessons:

a) Pre-test of the harmonised questionnaire at the European level and subsequent modification of the questionnaire (as already done for the preparation of the harmonised questionnaire).

b) Distribution of the questionnaire to national contractors, with a recommendation to pre-test the questionnaire in the country if any modification is introduced.

c) National contractors report the conclusions of their pre-test at the EU. The lessons learned should be disseminated across all countries and, if needed, the harmonised questionnaire should be revised before the release of the final version of the questionnaire.

*Cut-off point*

A cut-off point of 20 employees cannot be recommended since it has been shown in other studies that much innovative activity goes on in firms below this size. However, for resource reasons a cut-off point should probably be employed, and our recommendation is to cut-off at 10 employees (i.e. to include all enterprises with more than 9 employees in the sampling frame).

*Timing and periodicity*

On the basis of the experience from the first CIS it is clear that timing has to be harmonized. The planning of a future CIS has to start early (at least one year ahead of the launch of the surveys) and surveys in all countries should be launched within a few months.

In most countries the first CIS was launched in a year where the bi-annual R&D survey was not launched. This has proven to be an advantage since no country reported on major problems with competing surveys. Thus, every two to four years may be the optimal periodicity. A method implementing a simple questionnaire with only quantitative questions bi-annually and supplementing with all the qualitative questions every four yeas was supported by most national contractors.

The best time of the year to launch the survey is probably early Spring when most firms have closed their accounts but still have the events of the previous year present in their memory. However, no connection between the time of survey launch and realised results (e.g. response rate or item non-response) was found during the evaluation.

*Reminder procedure*

This point does not relate to countries that use interviews.

It is difficult to give recommendations on the reminder procedure since this to some extent must depend on country specific factors. As described in section 7.3 Germany reached a response rate of just 22% even though the contractor made an intense response chase, whereas Denmark and Norway reached response rates of over 50% with just two written reminders.

However, national contractors should ensure that a statistically satisfactory response rate is achieved. The methods to ensure this will be different in each country, but since a high

response rate is essential to the success of the venture, the Commission should be involved in the arrangement of the national reminder procedure.

*Follow up on responses*

Quality checks and follow up on incoming questionnaires are vital to secure reliable responses and reduce item non-response. Whenever possible this follow up should be based not only on inconsistencies within the questionnaire but also the information should also be checked with existing data (especially data on R&D and turnover). The activities on follow up should be planned and budgeted at the beginning of the project. Follow up on questionnaires should be executed as soon as possible after receiving the questionnaire.

*Survey method*

Which survey method should be preferred varies across countries. For small countries census surveys are preferable whereas for resource reasons this is not the case for large countries. The same applies for whether surveys should be postal surveys or interview surveys.

Mandatory surveys are preferable since this (normally) entails a high response rate, but this has to be confined to countries with legal frameworks that offer this opportunity.

Thus survey methods may vary across countries. However, it is recommended that the Commission should be involved in the selection of survey methods to secure valid, reliable and comparable data.

*Frame*

For the quality of the realized data it is vital that frames are of a high quality. This involves that they should have a coverage near to 100% of the target population. Furthermore, to facilitate stratification at least data on size of enterprise and branch should be included in the frame, and for large countries also regional information should be included in the frame. Finally, the frame should be updated (at least) every one to two years to secure the highest possible coverage and to secure that the auxiliary information is up to date.

These issues point to that whenever possible the official business registers of the national statistical offices should be used for sampling. If other sampling frames are used - and this may be the case in some countries where the official data registers cannot be utilised by other institutions than the national statistical office - their quality should be carefully checked by comparisons with other available registers of firms.

*Survey unit and respondents*

To secure that comparable survey units are used across countries we recommend to use the enterprise as a legal unit in the survey procedure even though there may be some uncertainty about the suitability this unit for analysis (as discussed in section 5.3.3).

Identification of the right respondent within the survey unit is very important. Who (i.e. which type of manager: general manager, R&D manager, marketing manager, etc.) should be preferred as the 'best' respondent is not entirely clear. Therefore a very important action which can be made to contribute to high response rates and reliable answers is to pre-identify the relevant respondent in each participating enterprise. This pre-identified respondent may not be the person to fill in the questionnaire, but it should be the person empowered to ask other persons to fill it in.

*Sampling techniques*

Sampling techniques should be harmonised in order to reach the highest possible level of comparability. We suggest the following procedure:

- The stratification for the sampling procedure should be harmonised and planned with respect to its feasibility and efficiency.
- Not only the variability of some characteristics but also the propensity of non-response in the strata should be considered.
- The sample from each stratum should be made by means of the Deming-plan (Deming 1956). This method can:
    - Secure a random selection of units
    - Utilise regional information
    - Favour the selection of ear-marked units (e.g. for a panel of enterprises).

*Subsampling for non-response*

It is recommended that an good design of sub-sampling for non-response in the CIS is developed (see e.g. Madow and Olkin, 1983). On this basis of such a design a survey of non-response should be executed promptly when the main survey has been closed. The non-response survey should be planned and funded at the start of the project.

*Missing data and raising*

Harmonized methods to edit, check and weight the CIS data should be developed by Eurostat and implemented both at the national level and on the European level.

*Tests of reliability*

We recommend to ascertain means, variances and covariances within strata for some important characteristics as a basis for calculation of standard errors. Standard errors should then be calculated for proportions, total values and means of selected characteristics and also for relations of pairs of characteristics (e.g. innovation expenditure/turnover).

### *7.6.3 The learning effect*

One thing that have not been touched upon in this last chapter is the learning effects. These learning effects are considered immensely important by the evaluation team.

Firstly there are 'internal' learning effects. In repeating this venture surely the Commission will have learned a lot from this first CIS. Also, if the same national contractors are picked for a new round of the CIS they will have learned much from this first venture. Finally, if the survey is repeated the respondents who filled in this survey or who are able to find the person who filled in this survey will have a clearer feeling of what innovation surveys are about. This has proven to be the case with R&D surveys.

Secondly, and equally important, there are 'interactive' learning effects. These learning effects should be not be understood as a vertical process only (i.e. between the Commission and the national contractors) but also as a horizontal process in which national contractors are learning from each other. The workshops that has been organized by the Commission in the first CIS already played this role. We recommend that future innovation surveys exploit this possibility further and arrange several interactive workshops. Furthermore, it may be beneficial for small groups of national contractors to meet and discuss specific parts of the innovation surveys.

If the venture is repeated several times, the learning effect will ultimately mean that innovation surveys will provide more valid, more reliable and more comparable data. This information will be invaluable for the design of technology policy, both at the national level and at the EU level.

## 7.7 Issues for further research

Embarking in a new venture as the CIS will naturally generate new issues for research since so many fields have been touched upon. Throughout this report, some issues for further research have been raised. All these issues are outside the commission of this evaluation and therefore they are presented here as issues for further research.

We present the issues in key-words since they are presented in detail in the report.

Some of the issues concern directly the quality of the data from the first CIS and Eurostat has reported that most of them will be dealt with in the final compilation of the CIS data. These are:

- Calculation of sampling errors for the realised samples and calculation of sampling errors for the sub-set of innovative enterprises (cf. section 6.6.2 and 7.5).
- The possibility of using information on non-response in the weighting procedure to improve the quality of the data (cf. sections 6.6.2 and 7.5).
- Production of a 'users guide' to the CIS database: Which variables can be compared across which countries (cf. section 7.5 - the first steps in this project have already been taken in Eurostat (1993)).
- Analysis of reliability of individual data records (analysis of the Netherlands information of the accuracy of results and analysis of double responses) (cf. sections 5.4.1 and 5.5.4).

Two other issues concern future innovation surveys and they should be investigated before a possible next round of the CIS:

- The possibility of making innovation studies in services (cf. sections 2.3 and 5.8).
- The advantages and disadvantages of using customised questionnaires (in services/ manufacturing, in small/large firms) (cf. section 5.5.1).

The two last issues are broader in their scope. The first is a general issue that it would be useful to investigate for the benefit of all questionnaire based surveys and the second is an issue that has a clear policy relevance:

- The influence of the survey method used in non-response analysis (cf. section 6.6).
- Review of non-EC innovation surveys (USA, Canada, Japan, Australia, etc.) to assess the level of comparability with the CIS data.

# References


Acs, Z.J. and D.B. Audretsch (1988), 'Innovation and Firm Size in Manufacturing', *Technovation,* Vol. 7, No. 3.

Acs, Z.J and D.B.Audretsch (1990), *Innovation and Small Firms,* MIT Press, Cambridge, Massachusetts.

Akerblom, M. (1994), *Proposed Compilation Methods for and Preliminary Set of Indicators Based on National Innovation Surveys Results*, Canberra, OECD-NESTI, 18-20 April.

Archibugi, D. (1988), 'In Search of a Useful Measure of Technological Innovation', *Technological Forecasting and Social Change,* Vol. 34, No. 3.

Archibugi, D. et al. (1987), 'Innovative Activity, R&D and Patenting: The Evidence of the Survey on Innovation Diffusion in Italy', *Science Technology Industry Review,* No. 2.

Archibugi, D. et al. (1991), 'Sources of Innovative Activities and Industrial Organisation in Italy, *Research Policy,* Vol. 20.

Archibugi, D. and M. Pianta (1992), 'Specialisation and Size in the Technological Activities of Industrial Countries', *Research Policy,* Vol. 21 No. 1.

Archibugi, D. et al. (1994), 'Concentration, Firm Size and Innovation. Evidence form Innovation Costs, *Technovation*, Forthcoming.

Arrow, K. (1962), *Economic Welfare and the Allocation of Resources for Invention.*

Arthur, B. (1985), 'Competing Technologies and Lock-in by Historical Events', *Paper 43*, Center for Economic Policy Research, Stanford.

Cesaratto, S. et al. (1991), 'The Innovative Behaviour of Italian Firms: A Survey on Technological Innovation and R&D', *Scientometrics*, Vol. 21 No. 1

Cesaratto, S. and S. Mangano (1992), 'Technological Profiles and Economic Performance in the Italian Manufacturing Sector', *Economics of Innovation and New Technology,* Vol. 2.

Chadwick, B.A. et al. (1984), *Social Science Research Methods,* Prentice Hall, Englewood Cliffs

David, P. (1975), *Technical Choice, Innovation and Economic Growth: Essays on American and British Experience in the Nineteenth Century*, Cambridge University Press, London and New York.

David, P. (1988), 'Path-Dependence: Putting the Past into the Future of Economics', *Technical Report* No. 533, Institute for Mathematical Studies in the Social Sciences, Stanford University.

DeBresson, C. (1986), *Conceptual Notes on the Measurement of Innovation,* OECD, Paris

Deming, W.E. (1956), 'On Simplifications of Sampling Design Through Replication with Equal Probabilities and Without Stages', *Journal of the American Statistical Association,,* 51.

Dosi, G. (1988), 'The Nature of the Innovative Process' in Dosi et al. (eds), *Technical Change and Economic Theory*, Pinter Publ., Exeter.

EIMS (1994), *The Community Innovation Survey - Status and Perspectives.*

EIMS/Eurostat (1994), *The Comparability of the CIS Questionnaires and Analytical Projects*, Paper presented for Working Group on Innovation and Technology Policy, OECD, Paris.

Eurostat (1993), *Data Structure for the Community Innovation Surveys 1992/1993*, Joint EC/OECD Seminar on Innovation, Paris, OECD.

Eurostat (1994), *A Comparability Study between the CIS and the EU National Questionnaires*, Table Document, First EEA Working Party Meeting, R&D and Innovation Statistics, Luxembourg.

Felder, J.et al. (1993), *Results of the Pilot Survey for the Mannheim Innovation Panel*, Mannheim, ZEW.

Freeman, C. (1982), *The Economics of Industrial Innovation,* Pinter Publishers, London.

Hansen, J.A. (1992), *New Indicators of Industrial Innovation in Six Countries. A Comparative Analysis*, Division of Science Resources Studies, U.S. National Science Foundation, State University of New York, College at Fredonia.

Jaffe, A. (1986), 'Technological Opportunity and Spillovers of R&D: Evidence form Firms' Patents, Profits and Market Value', *American Economic Review*, Vol. 76, No. 5.

Kaminski, P. (1993), *Comparison of Innovation Survey Findings*, OECD, Paris

Kleinknecht, A. (1993), 'Testing Innovation Indicators for Postal Surveys: Results from a Five-country Project,' in Kleinknecht, A. and Bain, D. (eds), *New Concepts in Innovation Output Measurement*, Houndmills, Macmillan, 1993.

Kleinknecht, A. and J.O.N. Reijnen (1991), 'More Evidence on the Undercounting of Small Firm R&D', *Research Policy*, vol. 20, no. 6.

Kline, S.J. and N. Rosenberg (1986), 'An Overview of Innovation', in Landau and Rosenberg, *The Positive Sum Strategy*, National Academy Press, Washington D.C.

Kristensen, A. (1993), 'Innovation Structures and Performance in Nordic Manufacturing Industry, Forthcoming EIMS Report.

Levin, R.C. et al. (1987), 'Appropriating the Returns from Industrial Research and Development', *Brookings Papers on Economic Activity, Microeconomics,* No. 3.

Madow, W.G. and I. Olkin (1983), *Incomplete Data in Surveys,* Academic Press, New York and London.

Ministère de l'Industrie (1994), *L'innovation technologique dans l'industrie*, Paris, Dunod.

OECD (1981), *The Measurement of Scientific and Technical Activities, - Frascati manual,* Paris, OECD.

OECD-NESTI (1994), *Description of New Harmonised Innovation Surveys Carried Out in OECD Member Countries*, Canberra, 18-20 April.

Pavitt, K. (1984), 'Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory', *Research Policy*, No. 13.

Pavitt, K. et al. (1985), 'The Size Distribution of Innovating Firms in the UK: 1945-1983. Conference on Innovation Diffusion, Venice.

Piatier, A. (1984), *L'Innovation dans L'industrie; les Enseignements de Quelques Enquêtes*, CETEM, Paris.

Nelson, R. and S. Winter (1982), *An Evolutionary Theory of Economic Change,* Belknap Press of Harvard University Press, Cambridge, Mass. and London

Rosenberg, N. (1976), *Perspectives on Technology,* Cambridge University Press, Cambridge.

Rosenberg, N. (1982), *Inside the Black Box: Technology and Economics,* Cambridge University Press, Cambridge.

Scherer, F.M. (1983), 'The Propensity to Patent', *International Journal of Industrial Organisation,* No. 1.

Scholz, L. (1992), 'Innovation Surveys and the Changing Structure of Investment in Different Industries in Germany', *STI Review*, No. 11.

*Science and Public Policy* (1992), Vol. 19 No. 6.

*Science Technology Industry Review* (1992), No. 11.

Smith, K. (1992), 'Technological Innovation Indicators: Experience and Prospects', *Science and Public Policy*, vol. 19, no. 6.

Townsend J. et al. (1981), Innovations in Britain since 1945, *Occasional Papers* No. 16.

Van Raan A.J.F. (ed.) (1988), *Handbook of Quantitative Studies of Science and Technology.* Elsvier, Amsterdam.

Whelan, B.J. (1993), *Report on Pilot Survey of Innovation in Irish Industry*, Dublin, Economic and Social Research Institute.

von Hippel, E. (1988), *The Sources of Innovation,* Oxford University Press, New York & Oxford.

Wyatt, S. et al. (1985), 'Patents and Multinational Corporations', *World Patent Information*, No. 7.

# EIMS DOCUMENTS

| INNOVATION POLICY | Publication N° |
|---|---|
| An Integrated Approach to European Innovation and Technology Diffusion Policy: a Maastricht Memorandum, 1993 | |
| Public Measures supporting new technology based firms: Workshop Proceedings, 1994 | 7 |
| Policies to support Tacit knowledge Transfer: Workshop Proceedings , 1993 | 8 |
| Public Measures to support the Clustering and Networks of Innovative SMEs: Workshop Proceedings, 1995 | 16 |
| Public Schemes Promoting Active Involvement of Employees in Innovation, 1995<br>Synthesis Report (10 pages)<br>Study (190 pages) | 22<br>22-S |

| EMPIRICAL STUDIES | Publication N° |
|---|---|
| Innovation Activities and Industrial Structure: Industry and R&D in a Comparative Context, 1993 | 1 |
| Investment, Innovation and Competitiveness: Sectoral Performance within the Triad, 1993 | 2 |
| Patterns of Innovation in Italian Industry, 1993 | 3 |
| Innovation Structures and Performance in Nordic Manufacturing Industry, 1993 | 4 |
| Technological Diffusion, Productivity and Competitiveness: An Empirical Analysis for 10 Countries - Part 1: Technology Diffusion Patterns, 1993 | 13 |

| INNOVATION IN THE FIRMS | Publication N° |
|---|---|
| Knowledge-Intensive Business Services. Users, Carriers and Sources of Innovation, 1995 | 15 |
| Innovation Strategies of Europe's Largest Industrial Firms, June 95 | 23 |
| **FINANCE** | **Publication N°** |
| European Second-Tier Markets for NTBFs, 1994 | ISBN 1-898975-02-7 (*) |
| Securitisation of Guaranteed SME Loans in Europe and Finance for Innovation, 1995 | ISBN 1-898975-04-3 (*) |
| Potential Market for Initial Public Offerings (IPOs), 1995 | |
| **REGIONAL ASPECTS OF INNOVATION** | **Publication N°** |
| Surveys of Regional Innovation? A Feasibility Study for Europe, 1994 | 9 |
| Analysis of SME Needs: Methodology in Design, Construction and Operation of Regional Technology Frameworks, 1996 | 18 |
| Assessment of the Regional Innovation Support Infrastructure: Methodology in Design, Construction and Operation of Regional Technology Frameworks, 1996 | 19 |
| Means of Obtaining and Exploiting Information on Main Industrial and Technology Trends: Methodology in Design, Construction and Operation of Regional Technology Frameworks, 1996 | 20 |
| Innovative Regions? A Comparative Review of Methods of Evaluation of Regional Innovation Potential, 1995 | 21 |
| **EVALUATION** | **Publication N°** |
| Evaluation of the Community Innovation Survey (CIS) - Phase I, 1995 | 11 |

| INNOVATION AND TECHNOLOGY TRANSFER<br>SUPPORTING INFRASTRUCTURES | Publication N° |
|---|---|
| Quality Promotion in Europe. A Review of European Community Member States' National and Regional Schemes and Measures in the Field of Quality, 1994 | ISBN 0-566-07512-1 (*) |
| Quality Networking in Europe | ISBN 1-85972-364-0 (*) |
| The future of research and technology organisations in Europe: Conference Proceedings, 1994 | ISBN 92-826-8451-2 (**) |
| Technology Brokers in Europe, 1995<br>Vol. 1/3: Technology Brokers in Europe, Summary<br>Vol. 2/3: Technology Brokers in Europe, Summary and Country Reports<br>Vol. 3/3: Directory | 10 |
| Technology Demonstration and Application Centres in the European Union, 1995<br>Vol. 1/2: Empirical survey 1994 and Workshop proceedings<br>Vol. 2/2: Country reports EU, USA and Japan | 14 |
| Science Park Networks,<br>Vol. 1/2: Development, characterisation and role, Jan. 95<br>Vol. 2/2: 12 Country reports, May 94 | |
| Consulting Engineering Services in Europe, March 96 | 17 |
| Good Practice in Managing Transnational Technology Transfer Networks, 1995<br>Volume 1: Subject Papers / Volume 2: Case Histories / Summary | 24 |
| Survey of the Innovation Infrastructure in Central and Eastern Europe, Nov. 94 | 25 |
| Good Practice in the Transfer of University Technology to Industry, May 1996<br>Vol. 1/2: Good Practice Guide<br>Vol. 2/2: Case Studies | 26 |

(*) Only available in book shops
(**) Available in the Office for Official Publications

Copy of all the reports, except those where the ISBN n° is given, can be ordered free of charge to :
DG XIII/D/4 - Rue Alcide de Gasperi EUFO 2254 L-2920 LUXEMBOURG
Fax N° +352 4301 34544

**EUROPEAN COMMISSION**

DG XIII-D-4 EUFO 2254 Rue Alcide de Gasperi L-2920 LUXEMBOURG

☎+352 4301 32321 Fax +352 4301 34544 e-mail: digna.amil@lux.dg13.cec.be Contact: D. Amil

# Request for transmission of EIMS STUDIES

# of the INNOVATION PROGRAMME

**I would like to receive one copy of the following EIMS studies/publications:**

| Publication N° | Title |
|---|---|
| | |
| | |
| | |
| | |

**to be sent to the address below:**

Name: _____

Firm / institution : _____

Address _____

_____

Postal Code: _____ City: _____ Country: _____

☎ _____ Fax _____ E-mail _____

Date: _____ Signature : _____

esn4