



**Government Statistical Service**  
**Methodology Series No. [9]**

---

# **Reducing Statistical Burdens on Business**

Office for National Statistics

**April 1998**

© Crown copyright 1998

ISBN 1 85774 286 9

## **ABOUT THIS SERIES**

This report is the second in the new *GSS Methodology Series*. The series will comprise monographs with a substantial methodological interest. Findings can be included provided that they illustrate methodology, but the series is not for findings *per se*.

Contributors should send their submissions to the series editor at the address below. Before doing so, however, they should clear the work with their line manager. The submissions will then be subjected to peer review, external or internal, or both, as the series editor see fit. Authors can submit their work for peer reviewed journals as well.

### **Publication process**

The series is aimed at getting out results quickly and easily. It is the intention that authors will prepare the final documents themselves, on their own PCs. The style has therefore been kept simple. Style guidelines specifying the fonts, type size, margin settings, etc. are available from the series editor (at the address below). The series editor should need to do minimal editing.

### **Dissemination**

The series will be available to GSS staff free of charge.

### **Series editor**

John Charlton  
Office for National Statistics  
Room DG2/08  
1 Drummond Gate  
London SW1V 2QQ  
Tel 0171 533 6239  
email john.charlton@ons.gov.uk

# Contents

---

	<b>Page</b>
Abstract	5
Introduction	5
The nature of statistical burdens on business	7
The variety of measures for reducing burdens	9
The Osmotherly review	9
Using administrative data	11
Main administrative sources of information	12
Inter Departmental Business Register (IDBR)	13
Geographical data	14
Issues emerging from UK experience with administrative data	15
Making greater use of employment data from administrative sources	17
Electronic data collection and business software	18
The way forward	21
References	22
Annex 1: Ten ways of reducing statistical burdens on business	23

---



# Reducing statistical burdens on business

*by Andrew Machin, Office for National Statistics*

## Abstract

*Over recent years there has been a strong drive to ease the demands on the public resulting from government requirements generally. In particular there has been downward pressure on the compliance costs of business surveys. This paper discusses the nature of statistical burdens on business and the various methods to reduce them. It describes experience of measures in the UK, especially the use of existing administrative data sources, development of business registers and electronic data collection methods. A recent independent study identified savings in compliance costs to business of government surveys of up to 26 per cent without impairing the quality of essential statistics. Future possibilities for using existing data sources and collecting data directly from businesses' own information systems are described. Developments such as special modules in commercial business software packages are particularly important given the potential for extracting statistical data by electronic means with minimal burden on the companies concerned.*

*(Keywords: Government, statistics, surveys, reducing, burdens, business)*

## Introduction

1. Statistics of business are vital for running the economy and business generally, yet the burden of government surveys is relatively small overall. The compliance costs to UK business amount to some £60 million per year, 0.01 % of gross domestic product. The load of statistical surveys on business is no more than about one per cent of the costs of complying more generally with administrative requirements of government. The burdens are nevertheless perceived by some businesses to be considerable and thus merit serious attention. It is not only important to minimise the burdens from the business point of view but also to do so for the quality of the statistics. It is vital to focus on the respondents in order to get co-operation and thus obtain reliable and accurate data.

2. The Survey Control Unit (SCU) of the Office for National Statistics (ONS) was established in 1968. Since then successive Prime Ministers have issued instructions on the control of statistical surveys to Ministers in charge of departments from time to time, in particular to ensure rigorous assessment and review of surveys. The rules have tended progressively to become tighter, particularly in respect of surveys of business, as a result of deregulatory pressures. The SCU is a small unit which is responsible for ensuring the control procedures are followed and promoting best practice. It aims to promote necessary surveys of high quality, prevent bad or unnecessary surveys and ensure that burdens on data providers are minimised. In assessing survey proposals and reviews, the quest is to achieve a balance, looking after the interests of data providers and also working closely with survey sponsors. Key roles of the unit are to check for potential duplication of surveys, monitor and report on survey activity and the compliance costs to business, and to encourage respect for

data suppliers. The unit also co-ordinates particular initiatives and the development of policy where required.

3. There have been various initiatives to cut back on statistical surveys, notably during the 1980s. Some growth in survey activity returned in the early 1990s, partly to rectify some of the deficiencies created earlier and also to meet new demands. We have recently had another major review with some further consequent reduction in the burdens, though there is now more emphasis on maintaining the quality of essential statistics. The measures to reduce burdens have also become better focused as the monitoring of compliance costs has revealed how the burdens on business are concentrated among a few major surveys. The 25 largest regular business surveys, some 5 per cent by number, account for some 88 per cent of the overall compliance costs.

4. The burdens on business resulting from surveys are varied and complex and there is correspondingly a great variety of possible methods to reduce them, many of which do not threaten the quality of the statistics produced. Recent savings in compliance costs have been mostly achieved through the streamlining and simplification of surveys, more efficient sampling and the development of business registers which have enabled more effective use to be made of administrative sources of data already available to central government. Minimising the burdens on data suppliers has become a key objective of the ONS and the wider Government Statistical Service. Major surveys of business are now subject to compliance planning whereby compliance costs are kept within limits agreed each year with Ministers on a three-year rolling basis. There is also a drive to improve information about available government statistics, encourage wider access and greater use, and promote coherence and harmonisation of data, under the '*wider agenda*' programme. A key component is the development of a new integrated database (*StatBase*) of statistics drawn from the whole range produced by government. This should help to ensure that existing data sources are used to the maximum extent.

5. At present there appears to be limited scope for the further use of administrative sources for business statistics under current arrangements. Administrative data on businesses are however also held within the companies themselves i.e. financial accounts, production and delivery records, wages data. A key strategy in the UK which in due course is expected to yield considerable savings in compliance costs, and which may also lead to more accurate statistics, is to tap at source these data that are used to operate the business. The ONS is thus working with commercial software companies to develop report modules in packages that meet ONS survey requirements within the packages themselves.

6. There may also be further opportunities for the sharing of data more generally in the future which would not necessarily involve extracting information from different government computers. This could be done by co-ordinating the (electronic) collection of the overall data required by government from business for all purposes. The UK Government intends to implement a strategy for the sharing of information needed by more than one government department, where legally permissible, by means of a new infrastructure. This would link each department's information systems, which should remain separate, with individual businesses (or individuals). There are useful opportunities here for the development of statistics, although clearly some difficult issues concerning the quality of the data will need to be considered. The means of preserving the confidentiality of the data will be especially important.

## The nature of statistical burdens on business

7. The burdens on business, which depend on both the costs and uses of statistics, vary enormously among businesses and are influenced by a range of factors:

- the nature and size of an enterprise
- the resources and expertise available
- the widely different perceptions of how useful the statistics are.

8. It is difficult to balance the burdens against the benefits. Few people are in a position to fully appreciate the needs of both users and providers of data but it is the statistician's job to do so.

9. It is useful to distinguish the two terms 'load' and 'burden'. The 'net burden' of a survey for business can be defined as follows:

$$\begin{array}{l} \text{Net burden} \quad = \quad \text{cost to enterprises of} \quad - \quad \text{value to enterprises of uses of} \\ \quad \quad \quad \quad \quad \quad \text{collecting the statistics} \quad \quad \quad \text{the statistics} \\ \quad \quad \quad \quad \quad \quad \text{(the load or gross} \quad \quad \quad \text{(the benefits)} \\ \quad \quad \quad \quad \quad \quad \text{burden)} \end{array}$$

10. In practice however, whether businesses see a survey as a burden or not will be a matter of perception which may not be in accordance with the actual costs and benefits involved. Thus:

$$\text{Perceived burden} = \text{perceived load} \quad - \quad \text{perceived benefits}$$

11. While it is important to minimise the actual burdens, it is also vital to address the problem of the perceived burdens. They may have more impact on the quality of statistics produced (through response rates, reliability of replies etc.) as a result of the effect on the willingness to co-operate with the survey.

12. The loads on business, which in the UK we refer to as compliance costs, often have to be approximated in practice but can be calculated in principle according to:

- the time taken to complete questionnaires
- the sample size and the frequency of the survey
- the salary costs of people providing the information
- other costs involved in organising the task, assembling data, computing and finally despatching the data required, responding to queries on the data. (Extra costs may include those of setting up systems and directing the questionnaire to the appropriate person.)

13. These costs can be monitored, at least approximately, and a lot of attention has been given to developing such measurement in the UK in recent years and inevitably to the reduction of the gross burden measured in this way. The benefits, which need to be included

in the equation to calculate the net burden of surveys, are far more difficult to measure and apportion to businesses. It is nevertheless just as important to take account of these, from the direct point of view of business as well as government decisions. Wider considerations also apply based on the net burden or benefit to society as a whole, not just to business.

14. The imbalances between costs and benefits or the perceptions of these, and their variations, mean that a proportion of enterprises, especially smaller ones, can be very reluctant to answer statistical questionnaires. Firms often see no benefit from statistics on business activity, especially if they are in local 'niche' markets or their products are highly specialised. Larger firms are more likely to benefit if they are competing in major national or international markets as they will naturally have more interest in the statistical information about the general market trends.

15. There are many factors which may affect the perception of the burden. If the statistics are clearly justified, people may feel positive about contributing to a worth-while cause, even though their business may only benefit indirectly. However the uses of statistics and their wider benefits, for economic policy etc., are often not easily appreciated by respondents. Purposes may not be clearly established and well known or the reasons for complexity may not be obvious. Apparent or actual duplication of requests for the same information elsewhere can be a particular source of irritation. Companies may also have concerns about the protection of confidential information given in surveys. They may not wish to share their data, or the risks they see of information being divulged may add to the perceived burden.

16. Statistical surveys can thus cause special concern to respondents when they have no interest in the results or the purposes are not understood. In contrast some administrative processes may be costly for business but they will be more readily accepted since the respondents will often have a clear interest in the process e.g. tax returns and claims for grants. Even where companies can benefit directly from the statistics, the perceived burden of a survey may be high because the respondent does not know about the benefits. Those using the statistics and those providing them within companies will often not be the same people. A company may commission some market research or employ a consultant to produce a forecast but be completely unaware that these exercises use the same statistics to which the company has actually contributed.

17. Changes to surveys may affect the balance of costs and benefits for enterprises in different ways. For example a survey may be simplified to reduce the load. This may be helpful for some companies but others who use the figures may find them useless because they are no longer available in the required detail. Samples may be rotated to prevent burdens always falling on the same respondents but the accuracy of results for assessing trends may be reduced unless the sample size is also increased. Both loads and benefits may be affected by changes to nomenclatures and the speed and frequency of their introduction. The balance of costs and benefits may also be affected by the quality of the statistics produced in other ways, for example the speed with which the figures become available.

## **The variety of measures for reducing burdens**

18. Just as the statistical burdens on business are complex and are influenced by a variety of factors, so they can be reduced in many different ways. Methods may involve:

- reducing compliance costs by collecting fewer or simpler data
- reducing compliance costs by making it easier to provide the data
- raising the quality and thus the benefits of the statistics
- enhancing perceptions of the statistics and their value

19. Only in the first case are actions likely to impair the quality of data and then damage to statistics can only occur where the data are essential. While it is by no means easy to keep statistics users, providers and producers all happy, much can be achieved by focusing on the data providers as the most important people in a survey, understanding their perspective and making their task as easy as possible to get a timely and reliable response. Much action to reduce burdens - especially when aimed at improving the dialogue with the respondents - is consistent with the aim of obtaining good quality statistics and ensuring that the objectives of surveys are successfully and efficiently met.

20. Some key themes are :

- providing clear justification for data and consulting users and providers effectively
- integration of procedures and co-ordination of classifications
- improving communications, from simplifying questionnaires to developing electronic methods

21. A full range of measures for reducing statistical burdens is listed in Annex 1. Many of these are already well known and applied widely as part of standard survey design although, in the experience of the Survey Control Unit, they are not all universally applied and particular aspects can often be overlooked. It is helpful to have a checklist when considering strategies for reducing burdens.

## **The Osmotherly review**

22. An independent study for the UK Government led by Edward Osmotherly was conducted in 1996 to look at how the burdens of government surveys on business might be reduced, especially for small and medium-sized enterprises. This followed a recommendation in 1995 by the former Deregulation Task Force (a Government-appointed group mainly representing business interests). All 12 recommendations of the Osmotherly Steering Group report<sup>1</sup> were accepted by the Government, subject to further work in some areas. It is now being implemented.

23. The Osmotherly report identified savings in compliance costs to business of up to 26 per cent (£17.6 million at 1995 prices) which are expected to be achieved by the year 2000 without impairing the quality of essential statistics. Some initiatives are already in progress, for example measures to streamline surveys, reduce overlaps between them, sample more efficiently and make more use of electronic data collection methods. A number of

---

<sup>1</sup> "Statistical Surveys: Easing the Burden on Business" and "Government Response to the Osmotherly Steering Group Report on Statistical Surveys"

recommendations are being followed up to simplify specific surveys including (subject to negotiation) Intrastat, the European Union system for measuring cross-border trade.

24. A considerable part of the estimated savings is to be achieved, in partnership with software houses, through the development of commercial accounting packages to extract reliable statistical data for survey purposes automatically from existing company records. Trials have been set up. A high-level team, including several representatives of business, has been formed to see what can be done to speed up progress with such initiatives and has begun to identify issues which need to be tackled.

25. A major recommendation of the Osmotherly report is that samples for business surveys should, wherever practical, be drawn from the Inter-Departmental Business Register (IDBR) <sup>2</sup>. The IDBR holds identification details and basic statistical information such as employment, turnover and industrial classification of individual businesses and company structures drawn mainly from administrative sources, in particular from the Value Added Tax (VAT) register, the Pay As You Earn (PAYE) employees' income tax and the company registration systems. The IDBR is an important tool generally enabling more efficient and more co-ordinated sampling and it avoids overlaps in the collection of basic statistical information. It enables survey burdens to be shared more equitably between businesses as well as helping towards compiling more reliable and consistent economic statistics.

26. Use of the IDBR will be vital in working towards the introduction of guaranteed '*survey holidays*' for small firms as the Osmotherly report also recommended. The ONS has already introduced such a guarantee from 1997 for firms employing fewer than ten people, whereby if they are included in a statistical survey they are told in advance for how long they will be expected to participate. They will not subsequently need to take part in another ONS survey for at least a further three years. This guarantee is included in a Business Charter<sup>3</sup> which explains what service business can expect from the ONS. This type of approach does not reduce the overall burdens on business but it is very important in the interests of fairness and in terms of perceptions of business. Other departments have been considering their own guarantees where possible in line with the Osmotherly recommendations. An inter-departmental committee has been considering how to maximise the extent to which this approach can be co-ordinated across all government surveys and extended to cover firms employing up to 25 people as far as practicable. Such developments present a number of challenges for government departments.

27. Another important development recommended by the Osmotherly report, is that annual compliance plans (as recently introduced in ONS<sup>4</sup>) are now being produced by all departments with major responsibility for business surveys. These enable the compliance costs to business to be kept within limits agreed with Ministers. These plans and regular reviews of major surveys, required for survey control purposes, involve independent outside observers. This means that more attention should now be given to maintaining an effective overview and control of overall compliance costs. This should improve the priorities given to different measures to minimise the burdens with less emphasis on looking at individual surveys in isolation as with the traditional methods of survey control.

---

<sup>2</sup> See e.g. "The inter-departmental business register"- John Perry, *Economic Trends*, November 1995.

<sup>3</sup> "ONS Business Charter: a partnership with business"

<sup>4</sup> See e.g. "ONS Compliance Plan 1997-1999"

## Using administrative data

28. Using administrative data can reduce burdens on respondents to government surveys and help produce statistics more cheaply. This is possible either by providing complete alternative sources to surveys or more generally by helping - through a combination of sources - to collect statistics more efficiently. The opportunities to use large amounts of data already stored in increasingly advanced computer systems must be continually evaluated.

29. Administrative data can provide a number of benefits for quality of statistics, for example the lack of sampling errors, non-response biases and recall errors which can occur in survey results. Analysis may be available more frequently and in more detail than from surveys. However, the use of administrative sources can also lead to problems of quality and coverage. In particular there may be difficulties over the timeliness and consistency of the information provided. If the obstacles to access can be overcome, these sources may provide detailed statistics not available elsewhere, but the accuracy of a set of data will depend critically on its use within the administrative system, how important it is, and the checking mechanisms in place. There is no point in gaining access to the data if there is insufficient knowledge or influence over the quality to turn the data into useful information and deal with potentially distorting changes to the administrative system.

30. Partly as a result of decentralised organisation of government statistical work in the UK, with statisticians working closely with administrators in government departments, a range of administrative data is successfully used for statistical purposes. The main administrative sources used for the purpose of business statistics are the VAT register, the PAYE income tax system and company registrations which feed into the IDBR.

31. There are inevitably some limitations in the coverage and continuity of the IDBR but it does cover the businesses of interest for most surveys and is valuable for drawing samples. The strength of the IDBR lies in the integration of data from the various sources including supplementary survey information, especially on employment in local areas. It enhances the range of sub-national data available. However, a drawback is the lack of a common reference number system for businesses. In linking up the separate details from different sources it is necessary to resort to name matching which is not an ideal method.

## Main administrative sources of information

32. In more detail, the main sources of administrative data used for business statistics in the UK are as follows.

### VAT register

33. 1.6 million traders in the UK are registered for VAT. There is legislation to permit the use of details of the businesses, including the identification (name address and VAT

number), turnover and industrial classification, for statistical purposes. The department responsible, Customs and Excise, send these details of turnover to the ONS annually. New registrations and changes to existing registrations are sent weekly.

34. Many businesses do not need to register for VAT although some firms with a turnover below the VAT threshold choose to do so. Some entire industries are exempt. Despite the exclusions the VAT register is a very useful source although there can be some errors in the turnover figures. Group registrations cause some difficulties for industrial or geographical analysis.

35. A recent improvement has been to adopt the Standard Industrial Classification (SIC(92)) (which is compatible with, but more detailed than, the European 'NACE' classification). This makes the VAT records a potentially much more useful statistical source than previously. The possibilities of direct use of these data on their own for statistical purposes, however, appear limited, for example for monitoring trends in turnover. A particular problem is that the turnover data do not all relate to suitable time periods required, with only a third of the figures for 3-month periods coinciding with the calendar quarter. The annual updating of the VAT threshold and the likelihood of distorting influences from occasional more fundamental changes to the system also mean that surveys need to continue as the prime sources of the statistics. The administrative data, however, do offer a valuable means of calibrating and validating survey information and of sampling more efficiently i.e. in sectors where the VAT turnover information is more reliable it is possible to sample fewer units.

36. Figures on total imports and exports to the EU are also available from VAT returns and are updated annually. For smaller traders these supplement the information from Intrastat forms which are only issued to the larger traders. Wider use is, however, limited since the data cover only EU trade.

### **Income tax and other records**

37. Details of the 0.9 million employers with employees paying income tax under the PAYE scheme are also made available for statistical purposes under legislation. The Inland Revenue supplies to ONS identification details (including the PAYE reference number), industrial classification data of limited quality and the number of employees every quarter. The industrial classification used is unfortunately not fully aligned with SIC(92) although it has recently been improved. This means the data are of lower quality and the costs to government and business are higher than they would otherwise be. Better quality classification of the data would be especially valuable for statistics of the smaller businesses.

38. PAYE data have been especially useful in helping to reduce the compliance costs of collecting statistics of employment for local areas.

39. *Other data from income tax records* of the self-employed (i.e. "Schedule D" which covers nearly 3 million sole proprietors and partnerships) are not accessible by ONS for statistical analysis except in anonymised form. Similarly information on company profits based on the corporate tax system (for some 0.5 million companies) are only provided to ONS in aggregate form as these data are highly sensitive. The figures tend to be out of date and volatile although, if the data could be suitably developed, they could be useful for example in calibrating profit inquiries.

40. The ONS also receives *annual data on personal income* from employment based on a 1 per cent sample of employees from individual National Insurance records held by the Department of Social Security. The annual New Earnings Survey which provides detailed statistics (from businesses) on the levels of earnings by occupation is also based on the same 1 per cent sample.

41. *Customs declarations* also provide another useful source of data although there are some problems. For example the location on the form may not be where a good was made. It could be the head office, depot or the address of an agent.

### **Company registrations**

42. Some 1 million incorporated companies are registered at Companies House, although about half of these are inactive. There is a lot of information although it is not all easily accessible. A useful development of this source is that the SIC(92) has recently been adopted for classifying companies by industry.

43. There are presently two main obstacles to using annual accounts of companies lodged at Companies House: the timeliness of the data and the definitions used which are different from those required for the national accounts. Information may include international activities and may not provide a suitable split by industry. The problems caused by the length of time it takes to produce company accounts may diminish with the advance of accounting computer systems though the differences in definition are more difficult to resolve. The development of accountancy packages or other computer software to extract required statistical information direct from company records also offers considerable potential and this is discussed later.

### **Inter Departmental Business Register (IDBR)**

44. The IDBR is a database of businesses run by the ONS which is used for selecting samples for most of the main business inquiries. It is also used increasingly for analysis in its own right. The register meets EU legislative requirements (Council Regulation (EEC) No 2186/93 of July 1993 on Community co-ordination in drawing up business registers for statistical purposes.)

45. In addition to names, addresses and reference numbers, the variables stored on the register include annual turnover (VAT), employment(PAYE), the SIC(92) industrial classification and information on the structure of the enterprise.

46. A major drawback is that all the different component sources for the register use different reference number systems. ( Only corporate businesses register at Companies House and neither VAT nor PAYE cover all such businesses. Firms may register on none, one, two or all three of these systems.) Automated name matching techniques are used to link the records. Duplicates occur, affecting the accuracy of population totals, and cannot be easily detected. In the absence of a common identifier, linkage could be improved for example by putting the VAT number on the PAYE return but it has not previously been possible to find an agreeable solution.

47. The company registration number, once it is found by name matching, is used to determine the enterprise group structures by reference to files of data produced by private sector suppliers, Dun and Bradstreet supported by their CD-ROM publications (“*Who owns whom*” and “*Key British Enterprises*”). ONS is now moving to a ‘*Worldbase*’ system which is replacing these products. Business structures can change rapidly with acquisitions and mergers and it is important to keep this information up-to-date.

48. While there are an estimated 3.8 million active enterprises, 1.9 million are on the IDBR. But the missing 1.9 million account for only about 1 to 2 per cent of GDP. The IDBR is now used for selecting the samples for most of the major business surveys. It is expected to help considerably towards obtaining more reliable and consistent economic statistics. While there are some weaknesses in the various components of data, as described above, the strength of the register lies in using the combination of all the data sources. The information from administrative sources is also supplemented by some data from surveys, in particular from the Annual Employment Survey which gives information about employment in local units. There is also a programme of “register proving” to check information. This is important but can sometimes be seen as an unnecessary repetition by businesses and does impose some additional compliance costs. The checks are restricted to firms employing ten or more people and the efforts concentrated on the largest businesses.

49. The register enables efficient stratification and selection of samples with minimal burden on businesses, especially small single unit enterprises, and also helps to avoid duplication of requests for information. It records the surveys each business is involved in and provides a valuable tool for sharing the burdens equitably and preventing individual small firms receiving an undue number of survey forms. As mentioned earlier the IDBR will be important in implementing the recommendations of the Osmotherly review concerning survey holidays for very small businesses. There are inevitably some problems to resolve and there may be limits to what may eventually be achieved but the IDBR will be a vital tool in working towards this goal.

## **Geographical data**

50. Except for single site businesses, none of the administrative sources above cater adequately for sub-national statistics because they do not generally distinguish between local units in multiple-unit enterprises. For example the VAT and PAYE registers are based on legal units or groups of legal units. They may have different local units in different parts of the country. However, the register data do provide a powerful source of local statistics when combined with and supplemented by survey information, for example employment data from the Annual Employment Survey. This covers the great majority of multi-site businesses every four years, the largest businesses every year. The compliance costs to business of collecting the employment and other data have been reduced considerably as a result of this arrangement. Moreover, the IDBR now offers the potential for enhancing the range of sub-national economic statistics available without additional data collection.

## **Issues emerging from UK experience with administrative data**

51. Listed below are some general observations from experience in the use of administrative data in the UK.

### **Access to administrative data**

52. There are often constraints on the use of data, e.g. legal obstacles or general safeguards of confidentiality. Where the data are sensitive, access for statistical purposes may be limited to anonymised forms of data and it is then impossible to link with other variables at the level of individual companies. However, acceptable arrangements for access may often be found and even legal constraints may be removed if a good case can be made to allow for statistical use of the data.

### **Integration of administrative and survey sources**

53. Good use of administrative sources is made where the data are combined with survey information. Statistics can be considerably enhanced in terms of the detail available, e.g. for local analysis. The data from different sources are often very useful for cross-checking to improve the quality of the statistics overall. This is clearly most effective if data can be linked at individual company level.

### **Reference number systems**

54. Good use of administrative data also depends on using a good business register. While in the UK there is a very useful combination of sources of data, the lack of a common reference numbering system is a serious drawback and there appears to be no immediate way round this problem. This limits the quality and range of the resulting statistics and the opportunities to enhance the information obtained.

### **Classifications**

55. To make the most of administrative data, it is important to standardise the definitions and classifications used. There have been improvements in the industrial classifications with the adoption of SIC(92) for all the data feeding into the IDBR database although this is still not used fully in the coding of PAYE income tax data. While NACE is compulsory for the purposes of providing data to Eurostat, administrative sources are not obliged to use NACE.

### **Definitions and coverage**

56. Definitions used for administrative data and the coverage are not usually ideal for statistical purposes. It is often difficult to make comparisons with other statistics and the different sources of similar data may tell different stories. Administrative data may often need to be supplemented by other sources and much effort is often required to make the data useable e.g. in making estimates to fill gaps or adjusting for time lags in the administrative system.

### **Accuracy**

57. There is usually a need to check accuracy of the records. The accuracy of administrative data depends on their use and the consequent feedback obtained. If administrators don't see the data as important they may make little effort to control quality. Consequently administrative data may require checking and editing before they can be used for statistics. However it is important to remember that administrative sources do not suffer

from some of the drawbacks of sample surveys. For example there are no sampling errors or non-response biases. It would be wrong to assume that the answers given in surveys are always the right ones.

### **Timeliness**

58. The late availability of administrative data can sometimes be a significant drawback preventing their use for statistical purposes. Examples, as described above, are the data published in companies' annual accounts and information on company profits from the corporate tax system. Some improvements are being made and it seems likely that delays in the availability of the information may be a less important factor in future as a result of improvements in computer systems.

### **Control and knowledge of systems**

59. One difficulty is that statisticians may not have much influence over administrative systems. It is essential at least to be well informed of developments affecting systems and any forthcoming changes. It may be necessary to develop contingency plans or have other sources of data ready to deal with likely changes to administrative systems which could distort the statistics or even interrupt their supply.

### **The need for an overview of all the costs and benefits**

60. In considering the relative merits of administrative or statistical sources of data it is important to assess the effect on all the costs (both to government and to business) and the benefits of alternative proposals. For example, saving on a survey, using an administrative data source, may require a question to be added to an administrative form. The costs of this would need to be weighed against the cost of the sample survey and the relative benefits, e.g. in terms of the likely accuracy of the figures. The balance of relative costs, and the benefits, may also change over time. The best solution for processing and transferring data today may not be the best in, say, five years' time, especially considering the dramatic changes taking place in the use of information technology.

### **Other means of controlling burdens on business**

61. While it is clear that the use of administrative sources for statistics can be an effective means of cutting burdens on business it is important to consider the variety of other means available e.g. reducing overlaps between surveys, improving their design and aligning nomenclatures more closely with those already familiar to respondents. Work is being done to adopt definitions that are in common use in commercial accounts and make more use of language and terms which are used by businesses in particular sectors. Electronic methods of data collection now seem to offer the most potential for reducing costs and burdens.

### **Accounting and other business software**

62. A key development in the UK, expected to yield substantial cost savings over the next few years, is an initiative to automate the provision of statistical information by business. This is discussed in more detail later in the section on electronic data collection and business software. As mentioned in the context of the Osmotherly review, the ONS is developing, in partnership with accountancy software firms, modules in packages which would extract the required information for surveys relatively easily from company records through the use of standard reports, thus virtually eliminating the costs of data extraction within the enterprise. There is also potential for collection of employment and earnings information direct from

company payroll systems. Certainly in terms of the savings in costs, and probably in terms of the accuracy of figures (given their more direct route), such developments seem likely to become more important than initiatives to make more use of administrative data held by other government departments.

63. For some of the larger companies, which are more likely to have their own bespoke computer systems, the use of electronic questionnaires and Electronic Data Interchange (EDI) may be more important, although the costs of obtaining and compiling data in the required form may generally need more consideration than those of data transfer.

### **Accounting standards**

64. There may be some possibility, by developing closer links with the appropriate bodies, to influence accounting standards and make company accounts more useful as a source of statistics. It would seem unrealistic to expect to bring accounting and government statistical concepts entirely into line but at least it may be possible to eliminate some differences in definitions where these are not essential.

### **Making greater use of employment data from administrative sources**

65. It is always worth considering the possibilities of making greater use of administrative data sources. For example in ONS we have been considering how to make better use of employment data from PAYE income tax records to improve the precision of short-term (monthly and quarterly) employment series. Currently the PAYE data on the IDBR are mainly used in the identification of births and deaths of enterprises. The PAYE employees figures are kept on the business register only to fill certain gaps in the 'register employment' data, where there are no suitable survey data, and to help gross up the less frequent detailed employment figures. The PAYE data are not used directly themselves towards the short-term statistics of employment.

66. The short-term employment estimates are obtained by drawing samples of businesses from the IDBR, stratified by industry and by 'register employment'. The latter is based mainly on previous employment surveys or other estimates available from other sources including other surveys conducted by ONS. The estimates of total employment are formed using a ratio estimator with 'register employment' as the auxiliary variable for the purposes of grossing up the survey results

67. There are some differences in the definition of employment between the PAYE administrative source and the survey estimates. In particular the PAYE employees figures only include those employees with earnings high enough to be potentially liable to income tax, i.e. some low paid or part-time workers are excluded. There are also lags in the administrative system and the problems over the quality of the industrial classifications used mentioned earlier. Nevertheless the PAYE data have the distinct advantage of being available quickly following the reference date.

68. It was recently found that the correlation between PAYE jobs figures and the employment figures 'returned' in surveys is closer than the correlation between 'register employment' and 'returned employment'. There is, however, no correlation between *changes* in PAYE information and changes in 'returned employment' due to the lags and idiosyncrasies of the administrative process, and some complications arising from business structures, which tend to distort measurement of the trend. It is therefore not possible to use the administrative data to estimate employment changes directly, in place of survey data.

69. However it was found that there was a marked reduction (of the order of 20 to 30 per cent) in the variance of employment estimates created using PAYE employees as the auxiliary variable in the ratio estimation compared with using register employment as the auxiliary variable, as at present.

70. As a result of this study, it is now proposed to switch to using PAYE employment both as a stratification variable and as the auxiliary variable in the estimation process for the short term employment series. More work is required to check that the quality will be good enough and to investigate the position in greater depth for different industries. However, it seems likely that this work will lead to greater precision of estimates of *changes* in employment. Some concerns over the relationship between PAYE units and enterprises need to be resolved. For example employees of a particular enterprise may belong to a PAYE scheme serving several enterprises. In addition possibilities of producing total turnover and employment estimates using regression techniques are worth considering.

71. This example illustrates that while administrative data may not always be suitable as a direct substitute for survey data they may help considerably to improve the efficiency of estimates from surveys and thus may still help to reduce compliance costs.

## **Electronic data collection and business software**

72. There is a wide range of possible electronic methods for collecting statistics, from simple electronic questionnaires submitted on floppy disk to the fully automatic extraction of information from business computer systems which can then be transmitted electronically to the statistical agency. The ONS has used various forms of electronic data collection for business surveys, so far mostly on an experimental basis. The main focus is now on developing collection using specific reports included as separate modules in commercial accounting software packages. The use of EDI ( Electronic Data Interchange) is well established for the collection of trade statistics (Intrastat) by Customs and Excise.

73. Full EDI replaces four elements in the process of supplying statistical data. It eliminates the need to :

- obtain a physical report from the business' system
- transfer the relevant data to the statistical form
- transmit or post the form back to the statistical agency
- enter the data into the statistical system

74. From the contributor's viewpoint it is the first two of these that provide the major savings with little or no benefit arising from replacing the postal system as the means of

returning the data. From the viewpoint of the statistical agency, however, it is replacing the second, third and fourth of these processes that would bring the greatest immediate savings. Early work in the UK thus concentrated on the method of transmission as the key problem to be resolved but it is now realised that the automatic extraction of data from existing company records should be the priority.

75. Very few businesses in the UK are so far able to send data by full EDI and few are willing to adapt systems just for the occasional few surveys. A major obstacle to the use of EDI is that it requires contributors and/or the statistical agency to invest in appropriate software and access to a communications network. Without sufficient incentives or the promise of significant savings of effort it is difficult for the contributor to justify such investment unless they are already using EDI for other business purposes.

76. Intrastat, presently accounts for the bulk of statistical data collected using EDI in the UK. Nevertheless the volume of data keyed from paper remains significant and has increased. Initially the rejection rate of incoming EDI messages was high. One difficulty was that the system relied on commercially developed software and, as a result, over a hundred different packages were produced. There was no official certification of the software and therefore no control over its quality. The emphasis was on allowing choice and reducing the burdens on business. The position has, however, now much improved through education of traders and the development of a system for the prompt identification and correction of errors. Various strategies have been deployed to make it easier for traders to use EDI for Intrastat rather than submit paper returns. Help has been provided in setting up systems, providing communication links and evaluating and approving software packages produced for Intrastat use.

77. ONS have had some success with electronic questionnaires which involves providing business with some software. It is a big step for a business to move from paper based system to full EDI. Once business has seen electronic questionnaires working satisfactorily, this may help them to see the benefits of automatic extraction of statistical data from their own databases. However it seems unlikely that they will be universally used by all contributors. Also it is awkward and expensive to send to businesses the original software for electronic questionnaires, usually on floppy disks.

78. It seems that the future of the collection of business data lies in making use of the information that businesses require to operate either directly for their own purposes or to meet administrative requirements. While much administrative data can be captured from government sources and has a crucial role in the maintenance of business registers, the process can sometimes be too slow or inefficient to meet all requirements. ONS is therefore working with commercial software companies to develop modules in accounting software packages that meet ONS survey requirements using standard reports. The first software is now due to be launched.

79. There are three major advantages of this approach:

- there is no need to provide the software needed to provide the data
- the data comes directly from the business internal systems with the resulting benefits to quality and compliance costs

- as the recurrent burden is substantially reduced , it may be possible for the statistical office to collect more information more frequently than would be acceptable using conventional methods.

80. However, not all the information required for the statistics is available within a package and the information that is available may not meet the precise definitions required. Another consideration is the need for the statisticians to be more active in anticipating the changing requirements of users as the time taken to amend systems to collect more or different information will be much greater compared with using traditional paper forms. It is necessary to be involved with accounting standards bodies in the UK in order to influence the definitions used in commercial accounts. A more consistent and coherent strategy for data collection is also required, moving further away from the traditional piece-meal approach which may involve individual surveys asking for similar information in a slightly different way.

81. The initial report module for ONS statistics that has been developed is very simple. It is hoped that sufficient businesses will participate in trials in order to identify problems and future enhancements but it is possible that the take-up could be low.

82. It is clear that a number of issues will need to be tackled:

- Dealing with the flexibility in UK accounting systems. The lack of a standard general ledger structure makes it difficult to develop software modules that would be effective without customisation for individual businesses. One possibility is that ONS might provide guidance to companies on how to produce what is needed from what they already have using existing report generators in software modules. It seems essential that ONS report modules that are developed allow businesses to control and change the way data are input. (Whilst this would provide business with an opportunity to make mistakes, this is no worse than the present position with paper forms.)
- Finding ways to generate interest and make future initiatives commercially attractive to software houses and participating businesses, in particular to persuade accounting software companies to modify packages for the relatively few users who get particular government survey forms. (The relatively low cost to business of providing statistical information makes them reluctant to invest in systems to support electronic statistical returns.)
- The need for an integrated approach to meet a variety of government requirements, not just for statistics, avoiding duplication of systems collecting similar information
- The need to extend data collection to cover information wider than that available from company accounts. Access to other systems, such as production, delivery and purchasing modules, payroll and personnel systems also need to be considered.
- So far , the reluctance of many businesses to deal with government electronically and the lack of a widely available, secure and reliable electronic means of communicating in the business environment.

## **The way forward**

83. Much of the reduction in compliance costs identified by the Osmotherly review is coming from the streamlining of surveys, sampling more efficiently, reducing overlaps between surveys and the development of business software. The Osmotherly study concluded there was some scope for the greater use of data from administrative sources, e.g. in the use of turnover information on VAT returns, but this was not seen as a major means of reducing burdens. There is the possibility as mentioned earlier of making greater use of the employment data from PAYE income tax records to improve the efficiency of survey estimates. However it seems significant further progress in the UK in the use of administrative data for the purposes of business statistics would require more timely data, greater co-ordination of definitions and classifications and further integration of the data sources. A common, or linked, reference numbering system for businesses would help considerably. But the opportunities under present arrangements seem limited. To go much further would need more statistical influence over administrative systems and accounting standards etc. A key strategy is therefore to tap at source the administrative data that are used to operate the business.

84. Looking further ahead, there may be scope for more sharing of data required for other purposes but, as with the idea of using standard business software, this would not necessarily mean intercepting administrative data stored on other government departments' computers but making the most of the data already available within businesses in a strongly co-ordinated manner to meet a range of administrative and statistical requirements. A wide view would need to be taken of the quality of information required for all users.

85. The UK Government intends to implement a strategy for the sharing of information needed from businesses (or individuals) by more than one government department where this is legally permissible<sup>5</sup>. There is no intention of merging all the sensitive information held on a business (or individual) into a single database. The idea is to create a new infrastructure to provide the link between the departments' information systems - which should remain separate- and individual businesses (or people). The full implications of this for statistics are yet to be assessed. There are clearly many important issues to be considered such as the safeguards of confidentiality, suitable definitions and the quality standards for shared data, the mechanisms for sample checks etc. But such a development may present considerable opportunities for statisticians to make more use of information already available for other purposes both to enhance the range and quality of statistics and to produce them more efficiently with minimal burdens on business.

---

<sup>5</sup> "Government.direct - A prospectus for the Electronic Delivery of Government Services"

## References

Machin, A (1997) Use of administrative data to reduce statistical burdens on business: Experience in the United Kingdom, Proceedings of the seminar on the use of administrative sources for statistical purposes, Statistical Office of the European Communities

Office for National Statistics (1996) Business Charter: a partnership with business

Office for National Statistics (1997) Compliance Plan 1997-1999

Perry, J (November,1995) , The Inter-Departmental Business Register, Economic Trends

UK Government (September 1996), Statistical Surveys: Easing the Burden on Business

UK Government (September 1996), Government Response to the Osmotherly Steering Group Report on Statistical Surveys

UK Government Green Paper(1996), Government.direct - A prospectus for the Electronic Delivery of Government Services

Walker, G (1997) Collecting data using EDI: Experience in the United Kingdom, Proceedings of the seminar on the use of administrative sources for statistical purposes, Statistical Office of the European Communities

### Ten ways of reducing statistical burdens on business

There is a wide range of measures which may be taken to ease the burden on business. They may involve either reducing the costs to business, increasing the benefits, improving the image of the statistics or a combination of these. The measures fall into ten broad categories:

#### **(1) Justification of surveys to ensure they are necessary and alternatives are not overlooked, co-ordination and avoiding overlaps in data collection**

- **The justification of data requirements and the methods of a survey should be as clear as possible** not only to safeguard against unnecessary surveys but also to ensure that the reasons for the survey can be conveyed to the provider. It should be clear what the benefits of the survey are.
- **Ensure alternatives to surveys are fully considered. Maximise the use of data already collected and avoid overlaps with other sources.** It may be possible to fulfil requirements fully or partly by combining and linking other sources, or by adding questions to other surveys, for example. Use archives (both for storing new data for possible re-use and for searching for existing data). Check on the availability of other surveys collecting similar data. Maximise the use of auxiliary information from existing data sources. Make data easily accessible to other users (who might otherwise wish to conduct similar surveys of their own).
- **Co-ordinate surveys which collect similar information. Ensure that surveys are not conducted any more frequently than necessary.** Where possible avoid asking the same business more than once for similar information.
- **Use existing sources of administrative data wherever possible.** While there can be some drawbacks there are many advantages in terms of costs and absence of sampling and non-response errors etc.

#### **(2) Efficient use of registers and sampling frames avoiding excessive or unfair distributions of burdens and facilitating the sharing of data**

- **Be clear about the accuracy of data required and sample as efficiently as possible** using a good, accurate and up-to-date sampling frame.
- **Keep sample sizes to the minimum necessary** and be prepared to vary these and the content of surveys over time. For example it may not always be necessary to collect the same degree of detailed data. Two-stage sampling is often helpful as information

from the first stage can be used to narrow down the target population or reduce the sample size required for the second stage.

- **Use common registers of businesses as far as possible** (in the UK the Inter-Departmental Business Register is generally recommended) as a means of facilitating the sharing of data to meet different requirements and avoiding duplication of requests for basic information.
- **Use rotating samples to avoid an unfair distribution of burdens among businesses** where this is sensible. For smaller businesses, reduce the probability of sample selection according to participation in previous surveys. Consider schemes offering “survey holidays”. Make sure businesses are aware of the steps taken to treat them fairly. ( Such measures do not reduce compliance costs but reduce the number of enterprises experiencing an intolerable burden and improve their perception of the burden.)
- **Use other means of lightening the load on smaller businesses** e.g. use thresholds below which firms are excluded or at least sampled more lightly or given simpler questionnaires. Special estimation techniques may be used for the missing detail.

### **(3) Evaluation of compliance costs and benefits at an early stage. Effective dialogue with data providers. Consultation, testing, piloting and monitoring errors.**

- **Evaluate as soon as possible the compliance costs, the likely benefits of collecting the statistics and the effect on the businesses** participating in a survey.
- **Maintain effective dialogue with respondents as well as users. Thorough consultation is required.** Ensure that businesses have an input to decisions about surveys affecting them.
- **Ensure that surveys are always carefully tested and piloted** (ideally this should include direct observation of a sample of respondents completing questionnaires).
- **Ensure that errors or gaps in completed forms are carefully monitored** to help make communication with the respondent as effective as it can be. This not only avoids bad data but helps to avoid annoying respondents with confusing or ambiguous questions.

### **(4) Alignment of language, classifications and definitions with those used by the businesses surveyed and harmonisation of the concepts used for different surveys**

- **Use language and terms which are meaningful to respondents.**

- **Develop and promote standards for data collection and analysis.** Harmonise and standardise classifications, definitions and questionnaire formats as far as possible **in alignment with systems of businesses surveyed** and related administrative systems. Co-ordination of the outputs as well as the inputs helps to ensure that results and comparisons based on them are useful to the companies providing data. It will also help to avoid duplication and make data collection easier. Ensure that standards are clearly and easily accessible. As well as being co-ordinated, standards need to be adaptable to changing requirements in a rapidly changing world. At the same time it is important to maintain a degree of continuity in the concepts and definitions used in surveys and not to change these too frequently. This helps to keep down the costs to companies of adapting their information systems to keep up with changes to questionnaires and the classifications used.

## **(5) Effective and co-operative approaches to businesses**

- **Ensure that surveys are aimed at the appropriate person** at a suitable level of management in an enterprise.
- **Mind the timing of surveys. Give advance warning** (especially when they are infrequent) so that respondents have time to collect the required records. Make the timing of surveys as convenient as possible e.g. with regard to the preparation of company accounts.
- **Give helpful and positive introductions to surveys.** Give a full explanation of the purpose of a survey to respondents. Make it clear what is required, for whom and why. Distinguish clearly which surveys are compulsory (with reasons) and which are voluntary.
- **Give assurances over the confidentiality of the data collected.**
- **Allow businesses to provide estimates** where this is acceptable and the exact information is difficult to provide.
- **Provide incentives to participate in surveys where possible.** Stress any other direct advantages of the survey to the respondents, invite them to contact the office, **make summaries of results readily available to them** where the statistics are likely to be of genuine interest. (Providing unsolicited statistics which are of no direct interest to the business can be counter-productive.)
- **Minimise the number of contact points in the statistical office for an individual business.** Co-ordinate and streamline the approaches where possible.

## **(6) Good questionnaire and interview design**

- **Take care over questionnaire and interview design to ensure clear and effective communication.** Choose the appropriate technique ( e.g. use telephone surveys only when the required information is likely to be readily available). Good design helps to

avoid bias (non-response and other non-sampling errors) and heavy costs due to checking and correcting data. **Be relevant, positive, clear, brief and logical.**

- **Give clear concise guidance on the completion of questionnaires.**
- **Keep surveys as simple as possible**, bearing in mind the balance of costs and uses. (Oversimplification may render a survey useless for some purposes although compliance costs may be reduced.)

## **(7) Adapting approaches to suit individual companies**

- **Adapt and "personalise" questionnaires to individual businesses as far as possible.** Pre-printed forms to include basic information already available can be helpful. Take account of records systems used by firms. Extract data directly from company records where feasible. Offer respondents alternative ways of supplying data. Provide help as necessary e.g. by visits from field staff.
- **Payroll bureaux or other agencies may provide useful points to collect data** rather than companies directly themselves.

## **(8) Use of electronic methods for collecting data**

- **Use electronic data interchange (EDI) or other computerised methods for collecting data wherever possible.** Promote standards for storing and transmitting the data.
- **Use or provide computer software to interface with existing business packages used by companies.** (To be effective the software should be easily accessible.)

## **(9) Advice to companies on uses of statistics, maximising their usefulness**

- **Ensure that the statistics are as useful as possible to businesses providing the data** e.g. in respect of timeliness, the detail or other qualities of the figures provided.
- **Promote the benefits of the data.** Advise business on how to obtain statistics and on how to use them, especially those which they are involved in providing.
- **Encourage links between users and providers of statistics** in companies where they are not the same people.

## **(10) Controls of survey activity and regular review.**

- (If all the above steps were always followed this last step would not be necessary!)  
**Follow survey control procedures and advice on good practice. Regularly review all existing data requirements to ensure that they are still relevant and**

**worthwhile.** Monitor the burdens and publish the information to demonstrate to business that the issues are taken seriously. Publish plans for surveys and the associated compliance costs and the quality of the statistics.