

SYMPOSIUM STATISTICS FOR POLICY AND RESEARCH

THE LINKS BETWEEN STATISTICS, RESEARCH AND POLICYMAKING

JOHN BLACKWELL

*Resource and Environmental Policy Centre,
University College, Dublin*

(Read before the Society, 7th November, 1985)

INTRODUCTION

The recent Government paper on the Central Statistics Office[1] and the accompanying report of the Statistical Council are to be welcomed. This is partly because their specific proposals would lead to a more efficient system of providing official statistics. Moreover, the very publication of these papers, together with the Statistical Users' Seminar of June 1984, have meant that the provision of statistics has come alive as a policy issue. One hopes that this symposium can give a further impetus to this policy debate. This Society is a particularly appropriate host for this set of contributions, as it is the only regular forum which brings together statisticians, social scientists and Government officials.

For too long, the provision of statistics has been seen as either something which was of little importance or else the concern of a small group of "experts". It is tempting to decry such notions by assertion. There is, however, a danger that at a Society such as this, the importance of statistics will be taken as self-evident. This is understandable. After all, the national income accounts have been described as the staff of life for the macro-economist, and similar things could be said about the links which bind policy analysis and statistics provision in so many areas. At an even broader level, information has been described as the currency of democracy. But in the harsh world which we are facing, the importance of statistics will not be taken as axiomatic.

This is especially the case in view of the many competing claims for public expenditure. Hence, this paper begins with the links between statistics provision, policy making and research. It goes on to point out a number of ways in which the provision of statistics falls short of potential. Some key constraints on the achievement of this potential are briefly enumerated. Finally, a number of matters which might usefully be considered by the new National Statistics Board are discussed.

STATISTICS, RESEARCH AND POLICY MAKING

A number of points can be made at the outset about the links between policy making, research and the provision of statistics. First, much of the recent discussion of statistics provision has revolved around the "needs" (or is it demands?) of users and the problems which face producers of statistics. However, such a division into producers and users misses out on some of the channels through which statistics are used. There are producers of statistics, researchers and policy analysts, public administrators and private sector users. One cannot appraise the provision of statistics without taking account of the way in which social science research occurs and is used. In particular, how is this research funded, to what extent is it used in policy making, and how do the activities of research, administration and statistics provision bear on one another?

To take one instance, the time horizon of policy makers may often differ from that of data producers. There are usually time pressures on policy makers. Research findings and data on specific policy issues must often be speedily available or they will not be used at all. And when the data and research findings are produced, the provider may find to his or her chagrin, that policy makers have moved on to some other topic. Hence, the production of statistics may continue to lag behind the changing focus of policy making. To take another instance, there is little evidence that the findings of social science research are used effectively in policy making, and there has certainly been too little contact between policy makers, researchers, private sector users, and producers of data.

A final instance there have been cases where, at

departmental level, statistical systems have been put in place without close links to the planning functions within these departments - not to speak of links to analysis outside the respective departments. Nor has there been a firm policy on dissemination in these instances. Not surprisingly, existing data bases at departmental level are woefully underutilised for policy analysis.

In the time available, there is no opportunity to cover private sector users of statistics. It should be noted, though, that there is a variegated group of users, and that the activities of intermediaries in data provision are likely to increase. These take CSO and other data and package the data in a way which meets the demands of their clients. This will raise questions in the future about the terms of access to CSO data for the purpose of secondary data provision.

Second, to focus for a moment on social science research, most of this research in Ireland occurs through the research institutes, through a funding system which was laid down 25-30 years ago. In the universities the model which applies is the humanities one of an individual researcher engaged in some form of desk work. For the social sciences this has become quite inappropriate, especially in areas such as macro-economic models or analysis of longitudinal data sets for social policy work. In these cases there is a need both for large data sets and team work. At least outside the research institutes, Ireland has fallen well behind the best European practice in this work. The pattern of research will influence, to say the least, the demands for data provision which are made.

But the implications are more far-reaching than this. Ireland is one of the few European countries which lacks a central funding body for social science research. If there were such a body, it could play a useful role, not only as a bridge between the academic community and the providers of data, but in other ways. It could, for example, seek to ensure the best use of the social survey research capacity in the country. It would also be the obvious body to pursue the idea of a data archive, which is taken up below.

Third, one is struck by the extent to which official

data provision and data providers have occupied a relatively humble, low-key role in public administration. This is recognised in the Government paper where it states that in certain other countries

the importance of statistics is recognised both within the public service and nationally to a greater degree than in Ireland, and this recognition is reflected in the status of the statistical offices in these countries (para 16)

Indeed, one might contrast the perceived roles of statisticians and of social scientists in Ireland in somewhat the same terms as those of Bulmer

There is a practical corollary of the definition of social research as fact-gathering. If that is the case, then the status of those engaged upon the task is a relatively lowly one. The position of social scientists within government owes a great deal to the way in which their role is conceived of by their masters. The role of "fact-gatherer" is a humble one. Fact-gatherers would not presume to offer advice about policy, nor would they be seen as fit persons to do so. Their role is in the background and it should remain there [2]

There is a lot to be gained from bringing statistics and the statisticians out from the back room which they have occupied for too long. The Government proposals, which would enhance the role of official statistics, are to be welcomed. In addition, the comments of the Statistical Council about the need for a broadening of the range of graduate recruits in CSO, have a good deal of cogency. There is also need for closer links between the work of statisticians and policy making in the public sector which in the professional sense has been dominated by economists and other social scientists. Such a demarcation between social scientists and statisticians has not been good for the social scientists or for the statisticians themselves or for policy making as a whole. It has also contributed to some of the difficulties in ensuring that the output of statistics moves in line with

the changing nature of user requirements (In saying this, one needs to point out that the users themselves have not exactly been blameless -)

As well as broader recruitment and improved secondment arrangements, suggested by the Council, exchanges could occur between statisticians in the CSO and social scientists in other parts of the Government service and in the private sector. The result of all of this should be closer links between the work of the CSO and the policy process, a greater degree of mutual understanding about the problems which producers and users of statistics face, and an augmentation of the skills of both statisticians in CSO and social scientists outside the CSO.

EFFICIENCY IN STATISTICS PROVISION ACROSS THE PUBLIC SECTOR

We now consider to what extent there could be improved efficiency in the production and dissemination of statistics across the public sector as a whole. As the report of the Statistical Council and the Government paper deal with many of the organisational and management issues which relate to the CSO, I will concentrate on the issues which relate to data providers in total. In recent years there has been an increasing provision of data by departments and state agencies outside the CSO, which, in the main, complement the existing range of data from the CSO. In addition, the data of some of these bodies could be used as raw material for the data gathering of CSO or as a means of verifying data. For this, though, data would have to be cleaned up and kept up-to-date with suitable designatory codes.

A number of problems have arisen with regard to this statistical activity.

- (1) No explicit co-ordination of this output of statistics occurs, nor is there even any indication of the amount of resources which go into this activity. The CSO itself has had no "leverage" over the gathering of

statistics by other public sector bodies

- (11) There is no quality control with regard to the non-CSO statistical output
- (111) There are no principles of access to non-CSO data For instance, while there has been some dissemination of the Industrial Development Authority and Department of the Public Service data, this has been as a by-product of secondary analysis by outsiders
- (iv) No users' guide to published and unpublished (but potentially available) data is available
- (v) While there are *ad hoc* surveys done by departments and other bodies, by The Economic and Social Research Institute, by other research institutes and by individual researchers, and a smaller number of regular surveys, none of this body of material is pulled together in one place and catalogued
- (vi) A number of departments and agencies are expanding their information systems, using computerised data But there is a danger that this computerisation will result in further fragmentation in data designation Ideally, in this process, the CSO should be consulted and compatible codes for occupation and other entities used

Some of the constraints to the dissemination of these data reflect the features of Irish public administration In a world where the particular Minister is ultimately responsible for all actions of the department, the latter may be reluctant to release a report or survey which may be used by outside commentators to suggest that all is not well

with current policies. Other constraints relate to the particular incentives facing departments or agencies. These statistics often began as a by-product of filling internal planning needs, or of implementing administrative schemes. The IDA employment statistics would be a case in point. Departments and agencies will typically have a different set of objectives with regard to such data, from those of outside analysts or statisticians. In other instances, there are no particular incentives for a department or agency to consider compatibility with CSO categories or codes, as it begins data collection.

One particular issue which involves departments and public agencies is the extent to which administrative records could be used as a data source. Particular cases in point would be Revenue Commissioners' data and Department of Social Welfare data. The potential for making greater use of administrative records will increase as data become more computerised. There are a number of constraints which explain why, up to now, much of this potential has not been achieved.

First, the only aspects of administrative records which are reliable are those which are regularly used by the authority in question and are, therefore, kept accurate and up-to-date. Considerable clerical resources may be required to maintain administrative records for the purposes of statistical analysis. For instance, designatory details have to be kept up-to-date, duplication has to be avoided. Second, the agencies will typically have priorities other than data provision. The Revenue Commissioners regard their main purpose as getting in money from the tax system and regard data provision as, at best, a subsidiary objective and at worst having no justification at all. Third, much of policy analysis requires linkages - either to CSO data or to departmental data. An example would be the relationship between social welfare payments and patterns of work. Such linkages are either quite difficult to effect or can raise problems of confidentiality.

In the light of all of this, the Government proposals to give the Central Statistics Office the central role for co-ordination of official statistics are timely. However, a number of queries can be raised with regard to this proposed

co-ordination It is not clear from the Government paper (paras 36-38) to what extent the CSO will be expected to use its co-ordination role in relation to non-departmental bodies such as State-sponsored bodies (For example, a number of these bodies compile data on specific entities from firms and in many cases have a client relationship with those firms) Nor is it clear what its relations with the Revenue Commissioners will be Furthermore, it will be easier to use administrative records at a time when new schemes are being put in place Unless resources can be found, it is likely that existing records - which need a once-for-all *plus* an ongoing commitment to data provision - will remain unused

SOME ISSUES FOR ATTENTION

Finally, some critical issues which might be addressed by the Statistics Board and by the CSO under its new structure, can be outlined

First, there is need to ensure an effective division of labour between the CSO and other bodies with regard to survey capacity and survey work Certain research centres, and in particular The Economic and Social Research Institute, engage in a good deal of survey work which has often provided invaluable primary data The Department of Labour Survey of School-leavers is a case in point This material could form a basis for a data archive, which is considered below There are questions to be raised about the terms under which access to such data could be obtained in the absence of an archive These are analogous to the issues which the Statistical Council (para 26) raises about user access to data from large-scale surveys of the CSO And is the existing division of labour between the survey capabilities of the CSO and those of outside research centres a sensible one?

Second, there is need for a hard look at the increasing proliferation of household surveys by public sector agencies Without any effort, one can count up surveys of housing conditions, health status, smoking, car ownership and use, and home holiday taking which have been taken or planned in recent years These are in addition to the

quarterly EEC Consumer Survey undertaken by An Foras Taluntais and the annual Farm Management Survey This raises the potential usefulness of a general household survey There has been a good deal of reliance on the Household Budget Survey by users in recent years, in the absence of a general purpose household survey which would collect data at the household level on living standards, education, health and housing The Household Budget Survey has become the most important single source of information on household well-being and on the redistributive consequences of Government actions However, the next such survey, following on the large-scale survey of 1980, will not be taken before 1987 It is also overloaded and cannot take further expansion In order to answer many policy questions, the ideal solution would be to move towards a general household survey which would ask questions on household composition, labour force status, employment, education, housing conditions, use of health services, among other things It could be used to illuminate a wide variety of policy issues, such as family formation, social mobility, health service usage by socio-economic group, use of education services, absence from work and the take-up of social welfare benefits in cash and in kind In such a survey, there could be a basic set of "core" questions, with a further group which varied from survey to survey (with custom-built questions such as, for example, on the incidence of smoking) Such a survey could be quite cost-effective, if one good general survey could substitute for a number of individual surveys One of its advantages is that departments could obtain data at relatively short notice without having to mount their own *ad hoc* surveys

Third, there is the lamentable lack of updated statistics on earnings and its correlates across sectors and occupations The 1979 Earnings Survey gave some invaluable data, although it lacked questions on prior education and training and trade union membership The key constraint here is one of resources There seems little chance of obtaining EEC funding in the short term

Fourth, there is need to explore the feasibility of a data archive A data archive would serve as a repository of machine-readable social science data from academic, commercial and Government sources and make them available

for secondary analysis. This would involve the acquisition, "cleaning" and cataloguing of these data. Such an archive would facilitate the secondary analysis of data. One of the useful functions which would ensue would be that of a broker between producers and users of such data. The analysis by users of large data sets could be eased by, for example, the development of software packages. It is possible that an archive could end up having a role in the dissemination of large Government data sets such as the Labour Force Survey and the Household Budget Survey.

Fifth, and related to the foregoing, there is need to "work" existing data as hard as possible. This can be done by the facilitation of cross-tabulations, by casting data on families and on individuals in ways which (however imperfectly) substitute for purpose-built longitudinal data, and by using rotating sub-samples in the Labour Force Survey with a sub-group of households being re-interviewed in consecutive years.

Finally, it may be that private sector firms would be prepared to share in the costs of engaging in certain censuses and surveys. The Census of Distribution, last taken in 1977, is a case in point. This provides valuable information for planners and for private firms. The cover charge of the hard-copy publication is far outweighed by the value of this Census to individual firms and to local authorities. There could be an opportunity to get a financial contribution from firms and local authorities ahead of the census. This could have the added benefit that a greater interest on the part of private firms in using official statistics could result.

FOOTNOTES

[1] *A New Institutional Structure for the Central Statistics Office*, October 1985

[2] Martin Bulmer, "Using Social Science Research in Policy-Making: Why are the Obstacles So Formidable?", *Public Administration Bulletin* No. 43 (December 1983), p. 41