Cooperation between a statistical bureau and an academic department of statistics as a basis for teaching official statistics

Gad Nathan
Hebrew University, Department of Statistics
Mt. Scopus
91905 Jerusalem, Israel
E-mail: gad@huji.ac.il

1. Introduction

The teaching of official statistics was for long deemed to be outside the scope of academic curricula and was traditionally confined to the internal training programmes of the official statistical agencies themselves and to those of international agencies. The relevance of academic training in statistics for official statisticians was questioned, for instance, by Bishop (1964), and by Benjamin (1975), who emphasized the need for the training of official statisticians in administrative skills, rather than in statistics. In recent years, however, statistical sciences have developed considerably to cover a wide range of applications, each with its own specific methodology. Some applications, such as econometrics, psychometrics, biostatistics, operations research etc., have become disciplines in their own right and recognized as closely related to classical statistical sciences. The development of a vast range of statistical techniques, developed over the past fifty years, has resulted in a high degree of specialization and this is manifested in the diversity of graduate academic programmes and curricula in statistics.

On the other hand, the functions and modes of operation of governmental statistical agencies have changed considerably over the past fifty years. The developments in statistical methodology, especially in sampling techniques and analysis, have shifted the emphasis from a primarily descriptive mode, mostly concerned with the tabulation and simple description of administrative data, to very sophisticated methods of collection by sample surveys, sometimes linked to administrative data, and to advanced methods of analysis, such as time series analysis and categorical data models. The increasing use of statistical models, implicitly or explicitly, both in the analysis of official data and in a wide range of supplementary operations, such as editing and imputation, matching, data dissemination, statistical disclosure control and estimation for small areas, has resulted in requirements for a far higher degree of academic sophistication in the activities of official statistical agencies.

The result of these developments has been the increased cooperation between official statistical agencies and academia, in particular with departments of statistics. This has manifested itself in widespread consulting activities of academic statisticians with respect to many activities of statistical agencies, joint research activity, both in methodology and in subject-matter topics, and the enhancement of direct access of academic researchers to micro-data collected by national statistical institutes for detailed analysis and for teaching purposes. Related to these developments, a variety of academic-based courses in official statistics

and closely related topics are being provided by academic institutions and, in some cases, complete integrated study programmes in official statistics have been developed at universities.

In the following, we shall examine the conditions for successful academically based programmes, which demonstrate that strong cooperation between academic institutions and national statistical institutes is a fundamental requirement for quality training in official statistics. We consider some models for such cooperation, based on the experience in Israel and in other countries, and describe more fully the programme in official statistics at the Department of Statistics at the Hebrew University, which is based on a high degree of cooperation between the two institutions, with courses often given jointly with CBS staff. The aims, structure and specifics of the programme are discussed, as well as its achievements and failures. Finally, conclusions and recommendations for models of inter-institutional cooperation, as a basis for quality training in official statistics, are presented.

2. The Israel Central Bureau of Statistics (CBS)

The CBS was founded, together with the foundation of the State of Israel in 1948, under the first Government Statistician, Prof. Roberto Bachi, as a central independent government agency, with the aims of providing the government and its agencies and the general public with the data required for policymaking, planning and research. Its functions and operations are legally determined by the Statistical Ordinance, 1972, which ensures its centralized position and independence. Thus, the head of the CBS, the Government Statistician, is appointed by the Prime Minister and must "in carrying out his functions, act on the basis of scientific considerations". A Public Council for Statistics, representing state agencies, institutions of higher education and research, public bodies and independent experts, is appointed by the Prime Minister to advise him and the Government Statistician on the bureau's programmes and operations. The CBS has sole responsibility for the collection and dissemination of all national statistical data and any statistical activity of other government agencies can only be carried out after consultation with the Government Statistician. The CBS is responsible for carrying out the census of population and housing, other censuses and a variety of sample surveys, as well as the collection of data from administrative sources. The data collected and disseminated cover practically all aspects of social and economic activity, such as population and households, labour, public and private expenditure and consumption, health, education, transportation, leisure, price indices, national accounts, environmental protection etc.

The bureau's subject matter divisions are in charge of planning the collection and dissemination of data in their specific domains, after extensive consultation with data users in government and public agencies, in the research community and in the private sector. Central methodological units are responsible for developing and applying advanced statistical methods to the bureau's operations. Their activities include sample design, development of editing and imputation methods, weighting and estimation, methods of file linkage and matching, census methods, statistical disclosure control, time series analysis and application of models for the advanced analysis of data. The methodological activities have developed considerably over the years, via cooperation and consultation with other national and international statistical organizations and with academic researchers in Israel and abroad (see further details below). The bureau's activities in dissemination of data are extensive and diverse. Traditional paper publications are gradually being phased out, to be replaced by a wide range of automated methods of dissemination. These include, for most

aggregate data, direct internet access, via PDF or Excel files, CD's, an RSS service for direct transfer to personal computers, a Geographical Information System, table generators and a time series archive. A vast array of micro data sets is available as Public Use Files (PUF), as Microdata Under Contract (MUC) or by access at a Research Room on the premises of the CBS office, with strict adherence to individual confidentiality requirements. In addition, the Israel Social Sciences Data Center (ISDC) holds and preserves the most important historical CBS data sets and serves also as an authorized distributor of the bureau's micro data.

As a result of all the above, the quality of the bureau's methods and the data it disseminates has received widespread recognition, both locally and internationally. The degree of public confidence in the information the CBS publishes and distributes is very high, both of the general public and of academic researchers, policy makers and planners. The respect the CBS has attained, by safeguarding its independence and integrity, is unprecedented in Israel, especially when compared to public confidence in other government activities. The data it publishes are publicly accepted without question and, for example, its consumer price index is of prime importance to the national economy in the linkage of loans, salaries etc., while the national accounts data are the major factor in measuring the performance of the economy.

The professional staff of the CBS is primarily recruited from first degree graduates of statistics, economics, sociology and other social sciences from the Hebrew University of Jerusalem and from other academic institutions in Israel. Some 20% have advanced degrees (MA or PhD), mostly in statistics and economics. However most of their training in official statistics has been in-house and on-the-job. An extensive training programme includes in-house training courses and seminar series in a variety of topics, many taught by external lecturers from academic institutions, external courses in management and specialized subjects, and encouragement to undertake further academic studies on a part-time basis, either single courses as external students or advanced degrees. Participation in tuition fees and paid or unpaid time off for study are often provided. Over two thirds of the bureau's staff participates in some type of structured training each year. However, until recently, most such training based on academic courses, was not specifically related to official statistics. The aim of setting up the Hebrew University's programme in official statistics, to be described in the following, was to rectify this situation.

3. The Department of Statistics at the Hebrew University of Jerusalem

The Department of Statistics was set up over fifty years ago in the Faculty of Social Sciences at the Hebrew University, under the leadership of Professor Roberto Bachi, the first Government Statistician and also one of the first deans of the Faculty of Social Sciences. Initially, the emphasis was on descriptive statistics and demography, but the development of research and teaching curricula in statistical theory and inference quickly became dominant and a separate Department of Demography was created. From its beginning, the department has had teaching programmes in both the Faculties of Social Sciences and in the Faculty of Science and several of the staff have held joint appointments with the Department of Mathematics and with other departments in social and natural sciences. The department comprises some 20 faculty members with a wide range of research interests, and administers Bachelor's, Master's and PhD programs in applied statistics, mathematical statistics, biostatistics and operations research. In addition, the department provides the vast majority of service courses in statistics at the university, primarily in the Faculty of Social

Sciences, but also in other university units.

The research record of the department's faculty is quite outstanding and it ranks very high by international standards of scientific publication. Its research interests cover a wide range of topics, including, for instance, sequential analysis, change point problems, process quality control, stopping rules and 'prophet inequalities', categorical data analysis, biostatistical methodology, longitudinal data analysis, survival analysis, analysis of correlated binary data, design of clinical trials, sampling methodology and nonsampling errors, analytical inference from complex samples under informative sampling, estimation of trends and seasonal effects and deterministic and stochastic models in operations research.

The three year undergraduate programme in statistics has some 200 students of the Faculty of Social Sciences, who take statistics as a single major or in conjunction with studies in other social sciences, while some 150 students of the Faculty of Science take statistics in conjunction with mathematics, computer science or other natural sciences. First and second year studies include required courses in introductory statistics and probability, statistical theory, supplementary mathematics, computing, a statistical laboratory and regression theory. Third year studies include required courses in sampling and statistical models, an independent research project in applied or theoretical statistics and a wide range of electives, such as: stochastic models, limit theorems, Markov chains, categorical data analysis, time series analysis, advanced sampling, introduction to Operations Research, sequential analysis and design of experiments.

The MA or MSc programme, which has some forty students, can be undertaken in one of five tracks: Applied Statistics, Theoretical Statistics, Biostatistics, Operations Research and Official Statistics (see details below). Joint programs with other departments of the Faculty of Social Sciences are possible. A thesis is required for the MSc and for the research option of the MA degree. Required core courses (depending on track) include statistical decision theory, statistical inference, multivariate analysis, random processes, Markov processes, linear and non-linear programming. At least two Master's seminars are required, and students in the Statistical Methods track must participate for one year in the Statistical Consulting Workshop, run in conjunction with the Applied Statistics Laboratory. The Ph.D. program is primarily thesis-based and there are some twenty doctoral and post-doctoral students.

4. Cooperation between the CBS and the Department of Statistics at the Hebrew University

Both the CBS and the Department of Statistics have benefited over the years from a high degree of cooperation between the two bodies. This cooperation manifests itself in a variety of modes. On the personal level, all the four Government Statisticians and some of the other senior staff of the CBS (e.g. the Chief Scientists) have had some academic connection with the Department of Statistics or with other departments of the Faculty of Social Sciences. This has been in the form of joint appointments, temporary appointments to the CBS of university staff on leave of absence and adjunct appointments.

The most important aspect of this cooperation has been in joint research of staff of the CBS with academic researchers from the department, sometimes with the participation of researchers from other departments or from abroad. The research has been on topics of interest to and for direct application by the CBS. In some cases, formal research projects were set up, sometimes partially funded by local or foreign research funds. For instance, a research agreement on surveys with multiplicity, initiated by the U.S. National Center for Health Statistics, carried out jointly by the Bureau's staff and academic researchers, was

funded by the United States–Israel Binational Science Foundation - Nathan, Schmelz and Kenvin (1977). Following are some other examples of such joint research, which has often resulted in publications in refereed journals or in papers presented at conferences. Pfeffermann and Ben-Tuvia (1985) proposed a statistical method for assigning priorities to road sections for resurfacing; Kantorowitz and Nathan (1987) consider the estimation of response error micro-effects from repeated surveys; a method for robust small area estimation, combining time series and cross-sectional data, is proposed by Pfeffermann and Burck (1990); Pfeffermann and Glickman (2004) consider a mean square error approximation in small area estimation by use of parametric and nonparametric bootstrap; Rinott and Shlomo (2007) study a smoothing model for sample disclosure risk estimation as part of a joint research agreement on statistical disclosure control between the CBS and a researcher from the department.

Other important modes of cooperation include consulting to the CBS by academic staff and their participation in advisory committees. Many formal and informal arrangements for consulting by academic researchers to the bureau's staff are implemented on a current basis. These are primarily in the areas of statistical methodology, such as sample design and data analysis. The CBS also utilizes expert knowledge in subject-matter areas and in methodological aspects of its work by means of a network of advisory committees. These are sometimes appointed directly by the CBS and sometimes as sub-committees of the Public Council of Statistics. In addition to data users, experts from other government agencies and research organizations, they usually include academic researchers from universities.

Teaching and training is another area of cooperation between the CBS and the Department of Statistics. Besides the cooperation of the CBS with the department's programme in Official Statistics, to be discussed below, the regular in-house CBS training programmes often incorporate lecturers from the department and members of the CBS professional staff are invited to give courses at the university, as adjunct lecturers, or to give single lectures, on their area of expertise.

Finally the CBS provides important data sets, primarily via the Israel Social Sciences Data Center (ISDC), at the Hebrew University, for use of students in course work, projects, seminars and theses. The possibility of easy access to real data by statistics and other students has been emphasized as of prime importance in teaching of statistics in a way which is relevant to social science applications – see Gal (2003) and Ho (2005).

5. Models for training government statisticians

There is no doubt that in-house and on-the-job training is the predominant method of training government statisticians at national statistical institutes. Much of this training is informal but many statistical agencies have formal training frameworks and programmes of studies in official statistics, which in some cases may accept students from outside the agency or offer special programs for foreign students, e.g., the Statistics Canada Training Institute and the U.S. Bureau of the Census International Programs Center. Examples of intensive internal programmes are those of the Israel CBS, mentioned above, of India's Central Statistical Organisation Training Division and the training programme of the Hungarian Central Statistical Office. International and regional training centres are another important supplementary resource and of prime importance for agencies that do not have the resources or manpower to provide specialized training in official statistics internally. Examples are the UN Economic Commission for Africa East African Statistical

Training Centre in Tanzania and the INSEE Training Center in Libourne (CEFIL).

Although academic staff may participate in teaching in these modes of training, the academic basis of these programmes is, in general, minor. This is due to the fact that the training of official statisticians by the statistical agencies themselves is often divorced from methodological or subject-matter research. Only a few universities offer separate courses in official statistics within departments of statistics or other departments, outside complete programmes in official statistics. Examples are courses in official statistics in the Department of Statistics of University College, Dublin (temporarily suspended) – Murphy (2002) - and at the School of Economic & Business Sciences, University of the Witwatersrand, Johannesburg. Some commercial on-line teaching programmes also relate to topics of official statistics.

An integrated programme of teaching and research in official statistics is provided in the French system of "Grandes Ecoles" by the National School of Statistics and Economics (ENSEA) and the National School of Statistics and Information Science (ENSAI) – see Tassi and Trognon (1983). They are part of the GENES group which has its own Research Centre in Statistical Sciences (CREST). Both schools offer graduate programmes in statistics, especially aimed at training official statisticians, but also in the fields of economics, actuarial sciences, management and computer sciences. The teaching staff is of a high academic standing and is actively engaged in research. Practically all the professional staff of INSEE and other French statistical agencies are graduates of one of these schools. The level of academic study is equivalent to a first year Ph.D. programme at a good US university.

Similarly the Brazilian national statistical office (IBGE) has a "School of Statistical Sciences", which offers a four year undergraduate degree (BSc) in Statistics and a post-graduate programme offering MSc degrees in Population Studies and in Social Surveys, which includes coursework and a dissertation.

While the French and Brazilian models for training official statisticians are certainly desirable ones, it is doubtful whether they can be successfully applied outside their respective higher education systems. The graduate programmes in official statistics of the Hebrew University's Department of Statistics (to be described in the next section) and that of the Department of Social Statistics at the University of Southampton, UK, seem to be more suitable models for university based systems of higher education. The Diploma and M.Sc. Programmes in Official Statistics at the University of Southampton, first established in 1999, have been developed jointly with UK National Statistics to cover the core skills and knowledge needed by professional government statisticians. It is based on the Department's position as one of the world's leading research centres in official statistics and in the statistical methodology of sample surveys and censuses. Much of the research has been in collaboration with staff of National Statistics. The curriculum includes coursework of 11 compulsory units and 5 optional ones during three to four years, allowing those working to do their studies on a part-time basis, and a dissertation, required for those completing a master's degree. Courses are held in London at the ONS and at the university. The course topics include survey and sampling methods and estimation, data analysis and index numbers, statistical disclosure control, demographic methods, national accounts and small area estimation. Many of the students are employees of National Statistics or its candidates for employment.

Another complete academic programme aimed, at least partially, to train official statisticians, is the Joint Program in Survey Methodology (JPSM) at the University of Maryland, founded in 1993 and sponsored by the Federal Interagency Consortium on Statistical Policy. Its aim is primarily to offer graduate

training in the principles and practices of survey research. It offers PhD, MS and certificate programmes and a large proportion of its graduates end up working in government agencies or already work there.

6. The Official Statistics Programme at the Hebrew University

The Department of Statistics, in conjunction with the CBS, initiated the programme in 1996, originally as a Diploma Programme and, since 2001, as an Official Statistics Track for the MA Programme in Statistics. It aims at imparting knowledge and expertise in the fields of collection and analysis of data and in statistical methods which characterise the statistical work performed in both the governmental and public sectors. The programme is available to all those accepted for the Master's programme in statistics, in general those who have completed undergraduate studies in statistics with a sufficiently high final grade. The studies comprise basic courses of the Master's program, specific courses for the track and electives from the general Master's program. Compulsory specific courses are: introduction to statistical methods in official statistics; seminar in statistical methods in official statistics; and advanced sampling methods. Elective courses specific to the track are: time series analysis; small area estimation; innovative statistical methods for censuses; models and methods for the study of populations, price indexes, statistical disclosure control, analysis of large data sets, national accounts and presentation of data. Other relevant electives from the general Master's program include: resampling methods; frequency data analysis; and inference on models from complex sample surveys. Some courses are taught by CBS staff, who have the appropriate credentials, sometimes jointly with a faculty member. The CBS also supports the studies of many students, some of whom are employees of the CBS.

7. Conclusions and discussion

We consider the multi-faceted inter-institutional cooperation, described above, as an important basis for quality training in official statistics at an academic institution. The cooperation in research activities ensures that the programme will be based on up-to-date research results, as required for all good academic teaching and not limited to professional training. On the other hand the active involvement of academic university staff in consulting and advising to the statistical agency's activities ensures that the teaching programme will not become too theoretical or divorced from the application requirements of the statistical agency. It should be emphasized that, although we regard integral university programmes in official statistics, such as the Southampton and the Israeli ones described above, as the ideal method for teaching official statistics, at least for statisticians, the programmes must be supplemented by in-house and on-the-job training at the statistical agencies.

Finally, some problems relating to the University programmes must be mentioned. Relatively stringent entry requirements for Master's degrees –the equivalent of a good undergraduate honours degree in statistics or in another discipline, if it includes a sufficient background in statistics - are required both at Southampton and at the Hebrew University. While there is no justification to relax these requirements, they do limit the intake of students, especially taking into account the other options available for Master's degree specializations. The supply of a variety of courses in the programme is obviously a function of student numbers, but also depends on the availability of sufficient academic staff who specialize in topics relating to official statistics. While teaching can be augmented by adjunct appointments and incorporating teaching by

suitably qualified staff from the national statistical institute, it is obvious that such a programme strongly depends on a sufficiently extensive core of regular academic staff at the university department, with research interests in topics related to official statistics.

REFERENCES

Benjamin, B. (1972). Teaching statistics and operational research to civil servants. *The American Statistician*, **26 (4)**, 23-26

Bishop, H. E. (1964). The training of government statisticians. *Journal of the Royal Statistical Society*, **A127**, 211-215.

Gal, Iddo (2003). Teaching for statistical literacy and services of statistics agencies. *The American Statistician*, **57**, 80-84

Ho, F. W. H. (2005). The role of official statistics agencies in the promotion of statistical literacy among students. In: *Proceedings of the 55th Session of the International Statistical Institute. Sydney 5-12 April 2005.* Voorburg, The Netherlands: International Statistical Institute.

Kantorowitz, M. and Nathan, G. (1987). The estimation of response error micro- effects from repeated surveys for invariant characteristics. In: *Proceedings of the Third Annual Research Conference*. Washington DC: Bureau of the Census, pp. 359- 390.

Ljones, O. (2003). Implementation of a certificate in official statistics. In: *Conference on Human Resource Management in National Statistical Institutes, 23-25 February 2003.* Thessaloniki, Greece

Murphy, P. (2002). Teaching official statistics in an Irish university statistics department. In: *Proceedings of The Sixth International Conference on Teaching Statistics 7 - 12 July* 2002. Cape Town, South Africa.

Nathan, G., Schmelz, U. O. and Kenvin, J. (1977). *Multiplicity study of births and marriages in Israel*. National Center for Health Statistics, Ser. 2., No. 70, Washington, D.C.

Pfeffermann, D. and Ben-Tuvia, S. (1985). A statistical method for assigning priorities to road sections for resurfacing. *Technometrics*, **27**, 75-79

Pfeffermann, D. and Burck, L. (1990) .Robust small area estimation combining time series and cross-sectional data. *Survey Methodology*, **16**, 217-237

Pfeffermann, D. and Glickman, H. (2004). Mean square error approximation in small area estimation by use of parametric and nonparametric bootstrap. *American Statistical Association, Proceedings of the Joint Statistical Meetings*, 4167-4178

Rinott, Y. and Shlomo, N. (2007). A smoothing model for sample disclosure risk estimation. In: *Complex Datasets and Inverse Problems: Tomography, Networks and Beyond*, IMS Lecture Notes Monograph Series, Vol. 54, pp. 161-171.

Tassi, P. and Trognon, A. (1983). Teaching of statistics at ENSAE. In: *Proceedings of the First International Conference on Teaching Statistics*, Eds. Grey, D. R., Holmes, P., Barnett, V. and Constable, G. M., University of Sheffield, UK, pp. 415-430.