

The Application of Blaise III to the Israel Labour Force Survey

Edith Noy and Gad Nathan, Central Bureau of Statistics, Jerusalem, Israel

1. Introduction

The Central Bureau of Statistics has only recently started to automate its collection processes. In the last two years the transfer of the telephone collection of monthly industrial establishment data to an automated process was implemented successfully, by means of locally developed software. This was an application to the continuous collection of data, at three regional centres, from a panel of respondents with a fixed questionnaire and it did not present any special problems. Following this success, a decision was reached to extend the automation of the Bureau's collection processes in two different areas. One is the collection, by CATI and CAPI, of price data from outlets for the consumer price index, presently under development. The second, to be described in the following, is the application of automation to the collection processes of the most important and central household survey system of the bureau - the Labour Force Surveys.

The Central Bureau of Statistics has carried out Labour Force Surveys since 1954. These surveys are the source of official government statistics on employment and unemployment. In addition, these surveys also provide data on demographic characteristics, such as age, sex, marital status, educational attainment, family relationship, occupation and industry. Occasionally, additional questions are asked on health, education, income, previous work experience and other topics. The statistics obtained from these questions are used to update similar information collected through the decennial census and from other sources. They are also regarded by government policy makers as important indicators of the nation's economic situation and are used for planning and evaluating many government programs.

A major redesign of the survey was initiated in 1995, primarily to improve the quality of the data derived from the survey. This process is being implemented in two stages:

- 1) Redesigning the questionnaire, primarily aimed to measure the official labour force concepts more precisely and to incorporate definition changes, according to the recommendations of the International Labour Organisation. s persons who had worked for at least one hour during the

determinant week, at any type of work, for pay, profit or other remuneration and Unemployed Persons as those who had not worked for even one hour during the determinant week and who had actively sought work during that week, by registering with the Labour Exchanges of the Employment Service or any other employment office. Finally all persons aged 15 and over, who were neither "employed" nor "unemployed" during the determinant week, are defined as Not Belonging to the Civilian Labour Force.

2. The Labour Force Survey - general description

2.1. Definitions

The sample population includes the permanent (de jure) population of Israel aged 15 and over, including the Jewish residents living in Judea, Samaria and the Gaza Area, potential immigrants and permanent residents living abroad for a period of less than one year. It does not include tourists and temporary residents, unless they have been living in Israel continuously for more than one year.

The Civilian Labour Force is defined as persons aged 15 and over who were "employed" or "unemployed" during the determinant week, Employed Persons are defined a.

Towards the next pilot test, we shall attempt to develop the treatment of partially filled out questionnaires and the transition from the questionnaire of one person to another one, in the same household. In addition, we shall work on developing the mode of interviewing separate household units in the second wave in cases where households have split or new households have joined since the first wave, when only a single unit was interviewed.

Since we have realised that the transition from the dialling stage to the interview was not flexible enough, we decided to change the standard mode suggested by Blaise and to adapt it to our needs.

In general, it is clear that there are significant potential benefits from moving to computer assisted methodologies, in terms of maximising the opportunities for better and faster data collection, and also possibly long-term cost benefits for a large survey, such as the Labour Force Survey. We expect that in the next few years, our office will be able to take full advantage of the CAI.

2.2. Sampling Methods

In the present format, four quarterly surveys are conducted each year, the interviewing for each quarterly survey being spread over the entire quarter. In each survey approximately 12,000 households are sampled and about 22,000 different households are interviewed over the year.

Two types of frames are used to select the sample: (a) a frame of localities, and (b) frames within localities. In urban localities and in some rural localities, the frames are the lists of dwellings in the municipal tax file.

In other localities the frames are lists of households, of dwelling units or of individuals.

The sample is drawn in two stages. In the first stage, localities are sampled from the Bureau's file of localities. In the second stage, in most localities, dwellings from the sampled localities and in some of the rural localities, households are sampled. The sampling is done in such a way that the final probability of being included in the sample is the same for every household in the population (slightly more than 0.7%).

The sample which is selected once a year, is divided into four groups or panels (waves). The interviewing of each of the four panels commences in consecutive quarters. Each panel is investigated four times - two investigations during two consecutive quarters and then, after a break of two quarters, there are two additional investigations during two consecutive quarters.

Thus, in each quarter, the sample is composed of four panels. The method of investigation by panel was devised for the purpose of providing good estimates for the differences between: consecutive years, consecutive quarters and parallel quarters from consecutive years.

2.3 Interviews and questionnaires

The collection is conducted by trained interviewers of the Central Bureau of Statistics who visit every dwelling in the sample. Dwellings not used for residential purposes, empty apartments, businesses etc., are considered as out of scope ("zero cases").

In residential dwellings, one of the household members is interviewed with respect to all members of the household. For households where no one is at home on the first visit, the interviewer usually makes two subsequent visits. If also on these visits no one is at home, special questionnaires are usually left to be filled in and returned by post. For each visit resulting in non-interview, the interviewer indicates the reason in a special questionnaire.

In urban localities the interviewers were asked to obtain the permission of households with telephones to conduct the second and third stage interviews by telephone. In 1995, about 52% of all interviews were conducted by telephone (about 87% of the interviews in the second and third investigations).

Interviewing is carried out continuously during each week for the entire three months of the survey. During each week, about 1/13 of the households included in the survey are interviewed. The "determinant week" in this period always refers to the "previous week," namely the week ending on the Saturday prior to the visit of the interviewer. The data obtained for any period (quarter, year, etc.) are intended to reflect the situation of an "average" week in this period.

The questionnaire used in the current survey is a paper and pencil questionnaire. For each household one questionnaire is completed

containing information pertaining to the entire household and one for each member aged 15 and over.

The questionnaire for an individual includes questions on work in the determinant week, number of work hours in general and in the determinant week, number of work hours less than usual in the determinant week, number of overtime hours, reasons for part-time work, reasons for absence from work during the entire determinant week or a part of it, place of work, geographical mobility of all employed persons, type of work and employment status. In addition, information is obtained on unemployed persons on search for work, search for full-time or part-time work, reasons for unemployment, whether the individual ever worked in Israel and when, and the previous occupation. For employed persons, the amount of time worked during the year and reasons for working only part of the year are also investigated. Those persons not in the civilian labour force but who had worked during the year preceding the interview, are asked about the last type of work they have done. Those not in the civilian labour force and who had not worked during the year preceding their interview, are asked why they did not work.

Apart from details about work, demographic information is also collected: age, sex, marital status, country of birth, period of immigration, level of education (number of years of schooling and type of last school attended). The household questionnaire includes information on the number of persons in the household, the number of rooms in the dwelling, the number of children in the household and the number of hours of work of any paid domestic help.

Besides the regular questions about work, and household and demographic characteristics, the questionnaire occasionally relates to other matters, such as housing conditions, domestic appliances owned by the household and various other subjects.

The inclusion of these subjects in the survey is intended to provide statistical material for a detailed investigation of the connection between work patterns of the household and other areas of behaviour. Also included occasionally in the survey are questions related to specific aspects of work such as seniority at work, labour mobility and so on. They are intended to provide further information on both household and individual behaviour in the labour sphere.

2.5 Processing the data and estimation

The collected data are checked, completed, edited, coded, punched and undergo logic editing - all at a central location.

The data are weighted, by age, sex and type of locality group, based on estimated total population of each group. The annual estimate is the average of the four quarterly estimates.

3. The redesign of the collection process

The preparations for the redesign of the data collection methods, began in 1996 and are presently continuing. The major feature of this redesign is the automation of the collection process. The revised questionnaire is now being designed for a computer assisted interview. The questionnaire, as it is now being designed, will most likely serve for a long period before it is redesigned. Therefore, a large amount of effort is being given to considering alternative questionnaire designs. Issues such as question phrasing, screen layout, computer assisted coding and editing methodology, must all be considered.

The possibility of developing a custom-made software for automated collection was considered, but quickly rejected. Given the late entry into the area of computer assisted interviewing and the high stage of development of ready-made software, development of our own software was considered as too costly and time-consuming. The only potential advantage of self-development was that it could potentially overcome more easily the inherent problems in the use of Hebrew (written from right to left) for interviewing. However after some initial testing of the Blaise system, the transfer to the use of a Semitic language was shown to be feasible, though not without its problems - see further detail below.

No systematic testing of the feasibility of various alternative software packages for computer assisted interviewing was carried out. However after a perusal of the literature and based on personal experience and demonstrations of the alternatives, the Blaise system was selected as a first choice for testing. Since there was no rigid deadline for automation, the perceived advantages of Blaise were considered sufficient to test for feasibility, with the possibility of turning to an alternative should it fail to perform as required.

The major change in collection procedure envisaged is that interviewers will conduct the survey, either in person at the respondent's home, using laptop computers, or by telephone, calling from a centralised location, using personal computers. The Bureau is looking forward to the potential benefits of the computerisation, mentioned in the research literature, which provides information on computer assisted interviewing. Primarily, the move to computer-assisted collection is expected to yield the following advantages.

Currently we are burdened by the time-gap between the collection of the survey data and its release. We expect that the new system will help to reduce this lag considerably, since it simplifies the process of collecting and transmitting data. In addition, work that is currently carried out after the termination of the collection of data will be done during the data collection stage and expedited.

Also, we hope that use of this system will aid in attaining data that is higher in quality through various checks, incorporated into the interview process. Errors detected by the computer at this stage can be corrected with the aid of the respondent. This capability is made possible through built-in editing features, verification procedures and automatic consistency checks.

In addition, we expect the automation of the questionnaire to aid the interviewer in following the correct line of questioning, i.e., skipping when necessary, etc. and moving the respondent smoothly through the questionnaire. We are confident that computerising the questionnaire will aid in reducing misclassification and other response errors. Questions will be tailored automatically to the respondent through insertion of appropriate pronouns and changing verb tenses and also by omitting irrelevant questions on a case-by-case basis.

4. The implementation of the Blaise system

At the beginning of 1996, a four-member team was set up to design and to implement the strategy for developing and testing computer assisted interviewing for the Labour Force Survey. This team consists of one representative from the Labour Division, one representative from the Field Division and two programmers from the Information Systems Division. Its work is guided by a steering committee of senior staff from these divisions, headed by the Bureau's Chief Scientist.

The basic strategy consisted of a continuous programme for testing and implementing computer assisted interviewing for the Labour Force Survey collection, starting with CATI for the portion presently interviewed by telephone. In the process of converting the questionnaires to telephone computer assisted interview, the Blaise case management system for CATI will also be introduced. Once the CATI system becomes operational, the field interviewing will be transferred to CAPI, with experiments continuing in parallel with those for CATI.

The programmers were responsible for designing the Blaise application of the questionnaire and for implementing the CATI Management System. At the start of the implementation, they studied the Blaise system, (from manuals) and functioned with a certain measure of success. When they encountered problems, they sought assistance from Blaise advisors via e

mail. When problems arose more and more, solving them via e-mail became arduous and time consuming. As our timetable was already set for the first test of the Blaise system, we invited a Blaise advisor from Holland to provide staff support in promoting the implementation. The advisor also distributed an updated and more detailed, "Developer's Manual" which aided the programmers greatly in their work in the Blaise system.

Two main types of difficulties were encountered while designing the questionnaire - those of the programmer and those of the interviewer. Following are a few examples of the problems encountered, of each type.

1) Difficulties for the programmer :

Most of the programming difficulties related to the use of Hebrew. Thus, on the question screen, the default position for the code numbers and the text for the possible answers is left aligned (numbers first). In Hebrew, it is necessary to right-align, with the numbers first (on the right). It was not possible to change the position of the code numbers, but we could

manually align the text to the right. This caused the numbers to be at unequal distances from the text, which makes it difficult to choose the right answer.

The error screens, which are in English, had to be translated into Hebrew. To do this, our programmers had to delve into the system codes and change them one by one, which makes for excruciating work.

Many problems were associated with the use of tabular entry. For the interviewer's convenience, we would like to consolidate all the responses to questions on demographic subjects in one table. This way, all questions on each person in the household can be asked and recorded consecutively. Due to the size limitations of the screen, the responses scroll beyond the width of the first screen and it is not clear which response belongs to which household member. To overcome this problem, first names are presented again on the second screen. However, on the second screen, we were unable to determine which responses would appear or to align the names on the left side of the screen. Currently, some of the responses appear on both the first and second screen (before the name).

When a closed question is asked, the answer is numerical. When we wish to present the answers in a table, it is difficult for the interviewer to use the table with no text. If we wish to insert the text into the table, we must create an additional column. We cannot refrain from presenting the numerical answer because of Blaise limitations. This creates a problem because of the limited screen space for the chart.

By Blaise default, each question is in a separate screen. In certain cases, we need to organise more than one question on the same screen. We were told by the Blaise advisor that this could be achieved, but it would be a difficult task. The default layout of the screen does not always meet our demands. In order to adapt it to our needs we must create different layouts, which involves a great deal of work.

Additional difficulties arose in adapting the system to the transfer of data from an external source (the first round of interviewing - carried out by field interview) to the telephone interview, via MANIPULA. The difficulties were mostly technical in nature and related to problems such as missing system files and unexplained error messages.

2) Difficulties for the interviewer :

Again many of the problems are associated with the use of the Hebrew language, as Blaise was created for languages which are written from left to right, while Hebrew is written from right to left. By default, the text is in English. In order to change to Hebrew, Ctrl + Tab must be pressed. Numbers are considered as English text, so every time a number is typed, the program enters English mode. In order to continue typing in Hebrew the interviewer has to switch to Hebrew. The same problem occurs when the "Enter" key is pressed.

If an error in an answer to an open question must be corrected, it is not possible to switch to Hebrew with Ctrl + Tab. Instead, the keys Shift +

Backspace must be pressed. It is complicated for an interviewer to cope with these different functions while conducting the interview. In addition, complications arise when answering an open question in Hebrew, unless one first opens the answers window using the space-bar, which is not necessary in English.

5. The testing program:

In order to test the impact of transferring the Labour Force Survey to computer assisted interviewing, it is proposed to conduct a number of small scale tests prior to a large-scale test. The outcome of these tests, will provide the basis for the decision as to whether a parallel survey is necessary.

Beginning in January 1997, a number of small-scale tests are being conducted using CATI. The tests involve the transfer of the questionnaire onto the CATI instrument and the CATI management. A total of some 80 households is indicated for the tests. These households will be interviewed over a six month period. It is proposed to analyse the results from the test, every two months, to the extent that the system has been implemented at that stage, and based on the conclusions, to continue the process.

During the first wave (carried out at the end of 1996), the interviewers conducted face-to-face interviews using paper-and-pencil questionnaires. During the second wave, the same households will be interviewed again, by telephone interview, using computerised questionnaires. The telephone interview is conducted from one centralised location, using personal computers. The CATI case management has been introduced, but the interviewers still dial the phone numbers by themselves, due to lack of appropriate equipment. The first part of this experiment has been concluded successfully and, though indicating some problems, shows that overall the CATI system works.

A larger-scale operational test, with a sample of approximately 2,500 households, is planned for the end of 1997. This test will be based on a full probability sample, to allow comparison of its results with those of the current paper-and-pencil collection and a fully developed CATI instrument. In 1998-99, an additional test will follow to begin with face-to-face interviews, using laptop computers instead of paper-and-pencil questionnaires.

The larger-scale operational tests will provide the necessary criteria regarding the functionality of the Blaise system according to our requirements and will indicate if the system is easy to use by survey programmers, by developers and by interviewers. Other issues will also be checked, such as potential effects on interviewers and permanent staff, effects on respondents, response effects, data quality, timeliness, costs, etc.

A decision whether to conduct a parallel test will be reached later on. Testing experiences of the US Census Bureau have shown that CAI produces comparable results to paper-and-pencil collected data and the

agency suggests that statistical impacts of CAI alone are not significant. In our case, as the revised questionnaire was introduced prior to CAI, a serious break in series is not expected with the implementation of CAI. Given the likelihood that the impact to be measured would be small and the assessment that a parallel test would be very expensive, probably no parallel test will be conducted.

6. Conclusions

As mentioned before, our Bureau is only at the beginning of the computerisation of the data collection in the Labour Force Survey. Our experience is limited to the design of the questionnaire and the implementation of the CATI management. Although we have already performed a few real interviews, beginning in January 1997, they were done by the planning team and not by the survey interviewers; so that we do not have the interviewers' opinions on using the new system. As the interviews were conducted from the central office, the transmission of the collected data has yet to be checked.

Despite the early stage of our project, we are pleased to have already a basis for the computerised questionnaire, which became possible due to the Blaise system. We are confident that we will be able to overcome problems that will arise in the process of developing our project. Our main concern is the difficulties in the transition from English to Hebrew. We hope that these difficulties will be solved by the adaptation of the Blaise system to Hebrew, or by its introduction into Windows.

In order to partially overcome the problems arising from writing in Hebrew, we plan to reduce as much as possible the writing of text and the use of remarks. Also, there will be no writing in Hebrew and English texts together in one answer. This stage was completed in 1994 and the new version of the questionnaire is fully implemented in the field since January 1995, as an operational paper-and-pencil questionnaire.

2) Redesigning the collection methods, begun in 1996 with experiments in the use of computer assisted interviewing, initially for telephone collection, using Blaise III. This will be described more fully in the following.