

DISCUSSION

J. KANTNER UNIVERSITY OF WESTERN ONTARIO

The three papers presented at this session suggest a sequential order of business. First there is the evaluation of the 1966 Census; second a discussion of plans for the coming decennial census of 1971; and finally an account of the work being done to increase the utility of that census through the introduction of geocoding. I shall take the papers in this order. In addition, since the title of the session "Aspects of the 1966 and 1967 Census Programmes in Canada" tempts a discussant to reflect on those aspects not mentioned by the authors, I do so briefly despite the acknowledged unfairness of such behavior.

Krotki, Muirhead and Platek "Evaluation Programme of the 1966 Census of Canada".

This paper discusses a series of projects designed to evaluate the results of the 1966 Census. These include:

(1) Reverse Record Check -- an impressive effort about which I would raise only two minor questions:

a) What is the logic and purpose of the double coverage of persons under six months of age in 1961? The RRC cannot evaluate coverage of this age group in 1966 and this may as well be accepted as a deficiency of the method. Nothing is gained by the double coverage technique.

b) The case for the application of the Chandra-Deming correction for events missed both in the Census and in other records is not convincing since the independence of the two approaches -- in terms of the types of errors to which each is prone -- has not been demonstrated.

(2) Agricultural Quality Check -- two points which the authors do not discuss are of some interest:

a) the direction of the 'error' is the same for all items compared. Thus the errors would not appear to be due to ignorance about agricultural holdings, as is suggested, but rather, to omissions that can be filled in by additional probing.

b) the net error for number of farms is larger relative to its estimated standard error than is the relative error for total acreage. This suggests something about the location of the error.

(3) Demographic checks -- the authors candidly acknowledge their difficulties in applying standard methodology. It would be ungenerous to add to their woes. There are, however, one or two minor points:

a) the continued high survival ratios for persons 10-14 and 30-34 are noted (p.15) as evidence that there has been little improvement in the underenumeration of children and persons 20-24. Apart from the fact that survival ratios are somewhat dull tools for the analysis of underenumeration, the exposition at this point is confusing perhaps because of the overly condensed references to dates and intervals.

b) in view of all the methodological uncertainty to which the authors confess and the fact that the matching tests have not yet been completed, how do the authors conclude, as on page 15, that "it can be reported that there has been no obvious improvement in the apparent net undercount of, say 2.5 per cent over all ages ..."? Such an important conclusion deserves fuller documentation.

(4) Household check -- theoretically such a check is useful but in practice leads to some ambiguity of interpretation since a good Address Register shows up deficiencies in the Census and vice versa. For the three cities in which the household check was performed the number of households on the A.Rs exceeded the number on the V.Rs by more than 1 percent. The authors seem inclined to dismiss the results for two of these cities but would, I expect, take the view that the A.R. is generally more complete than the V.R. and yet the postal check added to the A.R. by 2 per cent but only by $\frac{1}{2}$ percent to the V.R. It is hard to reconcile this with the idea that the A.R. is the more inclusive

list. Since the A.R. is made up from several sources and involves different conceptual approaches to the same "place" e.g. as an addressable structure and as an assessable property, might there not be an inflationary tendency in the A.R. of serious enough magnitude to compromise this particular validation technique. The very substantial deletions from the A.R. made on the basis of the postal check raises a similar question and also introduces an element of judgement that may be difficult to assess.

(5) Postal-Change-of-Address-Cards -- this postal check, instituted in Ottawa-Hull, is designed to measure underenumeration among persons who changed their residence around Census time. The test of underenumeration consisted of determining whether moves were enumerated either at their old or new address. That is to say, the matching of records involved searching Census records for the addresses in question and then ascertaining whether the family enumerated at that address matched the one which according to the change-of-address-card should have been there. At least this appears to have been the procedure from the description given. A further test taking account of the moving date was used to gauge the extent to which those who were enumerated were assigned the appropriate address as of the critical date of the Census i.e. June 1. From the first test it was learned that "the under-enumeration rate among movers was more than five times greater than for the population as a whole..." This is an important finding but possibly fallacious since in many cases indicated change-of-address may not represent a bona fide residence but merely a place, perhaps the home of a friend or relative, where mail is picked up until the relocation is completed. Unless the entire universe of census documents is searched, a formidable undertaking, the extent of underenumeration of movers could easily be overstated. Perhaps a greater danger is that I have misunderstood the procedure from the abbreviated description given in the paper.

Fellegi and Krotki "The Testing Programme for the 1971 Census in Canada".

The testing programme for 1971 thus far has been confined largely to the Test Census conducted in London in the fall of 1967. Although to some

observers the most interesting aspects of this test are the content innovations, it is clear from the remarks of DBS officials as well as from the design of the testing program, that Bureau interest centers in the feasibility of self-enumeration and mail-out-mail-back procedures. The debate over the advantages and disadvantages still rages (this may not be too strong a term) in the United States where the general educational level is significantly higher than it is in Canada. Thus, the DBS concern on this score is fully justified although London, a relatively sophisticated commercial city with a large university and closer to the urban-industrial heartland of the United States than most states, can hardly be regarded as the place for an acid test. London is "average" or "typical" with respect to a number of Census variables - and it is for this reason that it was selected - but it is probably well above average relative with respect to many of the considerations that make for successful self-enumeration. This is not to criticize DBS for its choice of London as a test site but merely to say something about the interpretation of results. If the London test "fails", which appears quite unlikely, the outlook for self-enumeration in 1971 would be bleak. If it "succeeds", well

The authors in stating the a priori case for self enumeration, overstate it to some extent by claiming advantages which logically could be had also in connection with conventional enumeration. They site the advantages of (1) early mailed returns (2) reduction of coverage errors through the preparation of address registers (3) questionnaire probes to improve coverage (4) concentrated publicity (5) geographic coding and (6) respondent specificity. The last mentioned which refers to the greater likelihood under self enumeration of getting information from the best qualified respondent in the household is perhaps an advantage that can be granted although there is nothing in the test program as here described that directly tests this point. As for the other points, none seems overwhelmingly wedded to self enumeration. The mails are available to interviewers as well as to other citizens so that field editing of mailed-in daily batches would seem to be a possibility; address registers are better than poor listings but nothing prevents using A.R.s in conjunction with regular enumeration; interviewers can probe as well, possibly,

as printed instructions; publicity campaigns may be easier to bring to a pitch on a given census-day but to conclude that this condition is the sine qua non of an effective publicity campaign perhaps gives too little credit to an industry that sells soap and automobiles all year round. The link between self enumeration and geocoding is a derivative of the earlier statements about the advantage of Address Registers. As an argument for self-enumeration it is, by itself, a non sequitur.

The test provisions themselves deal with a limited number of questions. the editing test, the Post Enumeration Survey and the computer programming are confined to the items on the short form questionnaire. Presumably several interesting content innovations, most of which are found on the long form questionnaire, are to be examined in later Census Tests.

The London Test questionnaire carried new items designed to give the Census greater penetration in the areas of education, fertility, language, foreign background, mobility and household structure. There does not seem to be any ready way to validate these new items, at least in the London test. The Post Enumeration Survey would have been the logical vehicle but, as already noted, the P.E.S. was limited to the small list of items collected on a 100 percent basis. It is possible to evaluate the relative troublesomeness of the new items by tallies of omissions, imputations, required follow-up, etc. but no tests of validity are presented. One hopes that the fate of these items will not be left solely to an impressionistic evaluation.

Fellegi and Weldon "Computer Methods for Geographical Coding and Retrieval of Data in the Dominion Bureau of Statistics"

This paper describes an exciting new development for the custom assembly of data by prescribed areal units. If successful in operation, one of the major sources of discontent between producers and consumers of census data will be eliminated. The chief virtues of computerized geocoding to which the paper gives attention are the ability to specify suitable areal units and the efficient access to data stored according to block face coordinates.

The application as well as the problems of this new technique are only

dawning. It can be expected that subsequent application will go beyond problems of areal assembly to applications that would treat location as an individual, household or family attribute in deriving journey-to-work configuration, pattern of intercensal residential mobility and so on. If spatial information other than residence and place of work, say the schools attended by children or the place the wife purchased her last basket of groceries were added, we would be well along toward a spatial representation of urban metabolism as it applies to individuals and households. As described in the present paper, the technique appears to be much more supple in arranging space than in filling the resulting parallelograms with information of more than one dimension. One can be confident that the competent ingenuity that has brought the technique this far will soon take it beyond its present limitations.

Returning to what might be regarded as the implicit theme of this session - the capacity of the Census to meet those data requirements for which it is the appropriate vehicle - it is difficult not to remark on the lack of attention on an occasion such as this to the role of the Census in a system of statistical indicators. The Census, as it is now, is the outcome of a large number of games being played by governmental, professional, academic, commercial and industrial gamesmen each with his own requirements for data. There is a growing insistence that this is not good enough in the face of the complex business of diagnosing and prescribing for the ills of a modern society. The development of a comprehensive and meaningful system of social indicators is one of the major tasks before our statistical agencies over the next decade. Significant improvement in the battery of economic indicators is to be hoped for also and indeed the momentum in this field of measurement makes progress highly likely. The succession of social crises that continually assault our Panglossian perception of the world can be expected to spur the development of social indicators in much the same way as the problems of depression and wartime mobilization forced us to a more sophisticated monitoring of the economy.

One should not expect Census personnel to do this job single handedly. This is a task requiring effort and broad commitment from the academic,

professional and governmental communities. It can be anticipated however, that just as in the development of economic series, our Census

colleagues will contribute in a major way. Let us hope that they will find time soon to give this "aspect" the attention it deserves.