

32p
international union
for the scientific study
of population

union internationale
pour l'étude scientifique
de la population

COMMITTEE ON DATA COLLECTION AND PROCESSING IN LDCs



centro latinoamericano
de demografía

SEMINAR ON COLLECTION AND PROCESSING OF
DEMOGRAPHIC DATA IN LATIN AMERICA
May 23 to 27, 1988, Santiago

RECENT EXPERIENCES IN THE COLLECTION OF DEMOGRAPHIC DATA
IN LATIN AMERICAN POPULATION CENSUSES OF THE 1980s

Juan Chackiel
Carmen Arretx
Centro Latinoamericano de Demografía (CELADE)
Santiago de Chile

CELADE - SISTEMA DOCPAL
DOCUMENTACION

AMERICA LATINA



900057237 - BIBLIOTECA CEPAL

CELADE
Santiago
Chile

IUSSP - UIESP
34 rue des Augustins
B-4000 Liège (Belgium / Belgique)

Tel. : (041) 22 40 80
Telex : 42648 POPUN

RECENT EXPERIENCES IN THE COLLECTION OF DEMOGRAPHIC DATA IN LATIN AMERICAN
POPULATION CENSUSES OF THE 1980s.

Juan Chackiel
Carmen Arretx

CELADE, Casilla 91
Santiago de Chile

INTRODUCTION

If we consider the classical division of demographic data into "stocks" and "flows", population censuses have always been the natural and most important source of data on the former. Through these data, information is obtained on the total of the population and its principal characteristics at a given point in time.

Flow variables, which express the changes which occur in the population, such as fertility, mortality and migration, ought to be obtained from continuous registers, but the reasons why these suffer from serious deficiencies in most countries of the region are well known. As a consequence, indirect methods have been sought for the calculation of estimates of the components of demographic change, and the population censuses have become the principal source of data for this purpose.

In the last few decades, Latin American censuses have ever more frequently and successfully incorporated questions aimed at obtaining estimates of fertility, mortality and migration in the recent past. A summary of the questions included in the 1980 round of censuses in the countries of the region and brief critical references on this theme are given in Section 1 of this document.

One of the most important events which have occurred in the region in recent years, in the search for improving census data, is the carrying out of regional workshops on the censal experience of the 1980s and on the preparation of the next round of censuses. The Workshop on Analysis and Evaluation of the Population Censuses of the 1980s held in Buenos Aires in 1985 marks the first of the meetings of this type, which have included the participation of representatives of the statistical offices of the countries of the region, census subject matter specialists and the international organizations involved in technical co-operation. This document basically gathers together the ideas which CELADE has presented at the above-mentioned workshops, concerning the problems which have been encountered in the utilization of data from the 1980s censuses to estimate fertility, mortality and international migration. We shall discuss these problems and, whenever possible, identify some of their causes. This analysis may help those countries which have not yet conducted their censuses to adopt measures to avoid some of the difficulties which have arisen and, of course to take them into account in preparing the 1990's round of censuses.

Although every stage in a census is important and bears on the overall quality, we do not intend to make an exhaustive analysis of each stage; rather, we shall refer only to those stages and, in some cases, to specific aspects which directly affect the quality of the information to be used in estimate demographic variables.

One aspect which we do stress is the need for census organizers and users to be constantly in touch with each other. This is particularly important at the planning stage, especially as regards decisions on subjects to be studied,

design of the census questionnaire and at the data-processing stage. It is also important that other specialists, such as data-processing staff, samplers, cartographers, and others, also be involved in those stages which directly or indirectly affect their field work. An extensive literature on the subject may be found in the publications which contain the documents presented at the above-mentioned meetings and which are indicated in the bibliography of the present document.

To ensure the quality of information collected in a census, it is not enough to have good maps, a well-designed questionnaire form, proper instructions, and a logical data-processing plan; it is also essential that the field work be carried out conscientiously. Field work, in turn, depends on the other elements just mentioned, as well as on institutional support, local geographical conditions, the population's receptivity, the enumerators' ability to get information and the respondent's capacity to give correct answers. It can be said that the optimum conditions for conducting a census exist only as a theoretical abstraction, while in real life situations arise which are often beyond the control of census organizers or enumerators.

We shall now discuss problems regarding the collection of demographic data through questions included in the censuses, preparation of the questionnaire, omissions, non-response and use of field samples.

1. DEMOGRAPHIC TOPICS INCLUDED IN CENSUSES

Latin American censuses are increasingly using questions that allow the indirect estimation of fertility, mortality and international migration. Experience has shown that such information provides a basis for timely and relatively reliable demographic estimates in countries that do not have adequate vital statistics. It is also useful in countries which do have adequate data, both as a means of obtaining estimates from independent sources for evaluation purposes and for conducting special studies, especially concerning explanatory variables, which cannot be carried out with vital statistics.

The following information on demographic components is usually sought.

a) Fertility and childhood mortality.

- i) Total number of children ever born.
- ii) Of these, the number of children alive (or dead) at the time of the census.
- iii) Date of birth of the last child born alive or children born alive during the last year.
- iv) Survival of the last child born alive or of the children born during the past year.

The first three of these questions have been proved to be very useful to improve the knowledge of fertility and mortality at early ages. There is abundant literature on procedures used to obtain demographic estimates from these data (Brass, 1974; United Nations, 1983). The last question, referring

to survival of the last child, which would allow the estimation of infant mortality, has not yet yielded satisfactory results, but demographers are continuing their efforts to find out what problems are involved in this and how to solve them, since this would be the way to obtain a current estimate of this indicator. In the last few years important progress has been made towards obtaining better results from this question, as may be seen from the experiences of the National Retrospective Demographic Survey of Guyana 1986 (Statistical Bureau and CELADE, 1987) and the Experimental Census of Junín de los Andes 1987 (Somoza, 1987).

b) Adult female mortality. Maternal orphanhood can be determined by asking the entire population whether the mother is still alive at the time of the census. In many countries, this information has provided the only available indication of adult mortality. Techniques for analyzing it may be found in the aforementioned publications (Brass, 1974; United Nations 1983).

c) Deaths in the household by sex and age. Brazil and Haiti have included a section on deaths in the household, by sex and age, during the year prior to the census; this information is used to estimate adult male and female mortality following a procedure developed by Brass (1977). Up to now the results obtained have not been evaluated.

d) International migration.

i) Country of birth for those born abroad.

ii) Year of arrival in the country, in the case of those born abroad.

These questions provide the basis for CELADE's research programme on International Migration of Latin Americans (IMILA), which includes a data bank containing information on nationals of each country who have been enumerated in a foreign country (CELADE, 1986).

e) International emigration. Studies on this subject -using indirect techniques- are still at the experimental stage and are based on indirect questions concerning residence in and outside the country of certain relatives of the persons enumerated (Somoza, 1980, Zaba, 1986). In five censuses in the region, international emigration has been studied by breaking down the data on surviving children according to whether they live in the country or abroad.

Table 1 shows which countries have included these items (a,b,c,d,e) in their 1980 censuses. The most important and encouraging conclusion is that the countries are fully aware of the usefulness of these questions. An example of this is the fact that every single country included the questions on children ever born and on children surviving, even those countries which had not formulated them in the 1970 round of censuses.

However, there remain significant omissions as regards the inclusion of questions aimed at measuring the above mentioned variables.

i) Many countries did not include the question "year of arrival" for persons born abroad. This is a serious omission, because it is not sufficient, in estimating the stock of migration, to find out how many foreigners there are in a given country. It is also important to determine the changes of this stock through time, so as to be able to project the trend for

the next few years. The question on the year of arrival was not included in the case of Brazil and Mexico, which received a considerable flow of migration, during 1970-1980. Thus, the lack of this information makes it difficult to determine trends of emigration from the countries of origin and complicates study of the phenomenon in the receiving countries.

ii) Some countries with inadequate vital statistics did not include the question on maternal orphanhood, which has proved to be very useful in estimating adult mortality. Since each question in a census must compete with others to be included in the questionnaire, a country with good vital statistics could afford to leave the question about the survival of the mother out of the census; however, it should be included, in countries which have incomplete vital statistics. This is the case, for instance, in Ecuador and Haiti, which nevertheless did not include the question in their 1980 censuses.

Although it is encouraging to note that censuses do include questions designed to gather information on population dynamics, the quality of the data collected is sometimes questionable. Information which had been found to be of good quality in the 1970 censuses appears to be less reliable in many of the current censuses; for example, there is some doubt about the estimates of childhood mortality based on the proportion of children dead among children ever born. Figure 1 shows the behavior of infant mortality trends estimated from various sources for some Latin American countries. Estimates obtained from the 1980 censuses for a given point in time are systematically lower than those obtained from previous censuses and from other sources. These differences are even greater in the case of estimates for certain subgroups of the population. These results give the impression that there has been a deterioration in the quality of data gathered for the most recent censuses.

2. PROBLEMS DURING PREPARATION OF THE CENSUS QUESTIONNAIRE

There are certain errors in the design of a census questionnaire which can lead to poor data quality, even if proper instructions are given to enumerators. We shall consider two aspects of this problem: the content of the questions, their wording, and their placement in the questionnaire.

a) Problems relating to the content of the questionnaire. To ensure that questions for the census are properly worded, it is essential that:

i) The objective of each question be perfectly clear, i.e., it must be clear why the question is included. Consequently, it is very important that the user take part in examining the questions, because he knows how the data will be used.

ii) Questions must be drafted by persons who are very familiar with local expressions, and questions which may be technically and grammatically correct, but which would not be understood by the local population must be avoided. This is not an easy task, because there may be significant cultural differences within the same country.

One example of how poorly designed questions can seriously limit the study of fertility and child mortality is to be found in the section on "children ever born and children surviving" in the Argentine census form.

FOR WOMEN AGED 14 OR OVER

16. Have you ever had children born alive?

Yes /___/ No /___/ Do not know /___/

17. If the answer is yes,

a) How many children have you had?

1	2	3	4	5	6	7	8 or more	Do not know
/___/	/___/	/___/	/___/	/___/	/___/	/___/	/___/	/___/

b) How many children are still alive?

None /___/

1	2	3	4	5	6	7	8 or more	Do not know
/___/	/___/	/___/	/___/	/___/	/___/	/___/	/___/	/___/

In order to estimate the number of children ever born (and the number still alive at the moment of the census) especially to women aged 30 and over, we are obliged to invent the number of children for the category "8 or more" and this will affect our estimates of fertility and childhood mortality. In defence of the form used, it might be argued that the problem does not arise in the case of Argentina because the average number of children per woman is relatively low, and it could be assumed that very few women have more than eight children. Nevertheless, some analyses that have already been made of census data for certain regions of the country show that the information is incomplete and considerably limits the possibilities for effectively use of the data (García, 1984). This is so because in some regions of Argentina, the average number of children ever born per woman is much higher than the average for the country. The category "8 or more" was used because the idea was to pre-code the answers with a single digit.

The information desired should never be sacrificed when there are simple solutions which do not substantially change the questions and the working criteria used (in this case, pre-coding). The approach taken in Venezuela and the Dominican Republic, was to add a first line with code numbers 0, 1 and 2, making it possible to obtain information from women with as many as 29 children:

Cuántos hijos nacidos vivos ha tenido en total?

0	1	2								
/	/	/								
0	1	2	3	4	5	6	7	8	9	
/	/	/	/	/	/	/	/	/	/	/

One problem that is very difficult to assess is whether or not respondents have interpreted a question correctly. For example in answering the question about how many children a woman has borne, respondents may tend to think that the reference is to "male children" (in Spanish, "children" is the masculine noun "hijos"), in which case, fertility would be underestimated; however, it is very difficult to prove this if the question does not make a distinction between male and female children. Often, in an effort to ensure that the concept behind the question is clearly understood, long and complicated sentences are drafted which only confuse matters further. The ideal is to have a brief and simple question which does not give rise to any ambiguity or misinterpretation.

b) Problems relating to the format of the questionnaire. This matter could be studied at length from several points of view: size of questionnaire, horizontal or vertical layout, order of questions, or type of pre-coding. Here, we shall discuss certain negative aspects of the forms used by some countries.

i) Transcription of data. In the Haitian census, women were asked whether or not they had borne children (the filter question) and, if the answer was yes, their names and numbers were transcribed to the last page of the questionnaire, where the questions of fertility, childhood mortality and international migration of surviving children were placed. In general, it is not advisable to design questionnaire requiring a transcription from one place to another on the questionnaire. Transcription makes the questionnaire more difficult to handle and can introduce errors.

ii) Placement of questions. The placement of questions on maternal orphanhood and international immigration generally is not problematic, since these questions are directed to the total population and usually appear among those referring to the general characteristics of the population. Because the question on orphanhood status is very brief and takes up very little space, it may tend to be overlooked if the questionnaire is too complicated.

No easy solution has yet been found to the problem of where to place questions on how many children a woman has borne and how many are still living, directed at women above the age in which the reproductive period is presumed to begin. The fact that they are usually placed at the end of the questionnaire following questions on the labour force (LF), may explain why they are often left unanswered. In at least one country, enumerators were instructed to draw a diagonal line across the LF questions when the women were

economically non-active; on almost every questionnaire, the line was extended right to the end, thus also striking out the questions about fertility and child mortality. This is an important point, since these questions appear in this position in twelve out of sixteen censuses taken.

If these questions are to be placed in a different position, it is important to ensure that they follow the logical sequence of the other questions in the schedule, so that the universe being studied is narrowed down as the interview progresses. It should begin with general questions directed at the entire population, followed by questions on education for those aged 5, 6 or 7 and older, then the questions on LF for those older than 10, 12 or 15 (or a similar age), and finally, the questions on fertility and child survival for women over an age of around 15. If a different placement is to be used, the questionnaire should indicate very clearly to whom the questions are directed and it should follow a coherent design. In Paraguay, for example, questions for women aged 14 years and over were placed within the block of questions directed at the general population, which is not logical.

One solution is to place the questions for these women at the end, as is done at present, but to clearly identify the universe to which each block of questions refers, being careful to indicate this in the instructions for enumerators, and to ensure that field work is closely supervised.

iii) Complicated questionnaires or questions. It is generally true that the simpler the questionnaire used as regards the number of questions and their wording, the lower will be the risk of error, particularly related to non response or wrong notation on the questionnaire. Organizers of population

censuses are often pressured by users in different fields to include highly specialized questions, the argument being that the census provides a unique opportunity to gather such information.

Some of the questionnaires used in the latest round of censuses were overly complicated; for example, the questionnaire used by the Dominican Republic had 72 questions for the population and Brazil's expanded form had 57. Most countries included around 20 questions.

iv) Wording of questions. The quality of the data collected depends, to a large extent, on the wording of questions. Thus, for example, better answers were obtained to the question "is your mother alive?", when the answers "yes" and "no" were replaced by the expressions "alive" and "dead", probably because this avoided confusion between the answers "no" and "do not know". Nevertheless, none of the censuses adopted this minor change, probably because of the lack of communication between users and organizers of censuses, as mentioned earlier. The fertility questions, directed at women of around age 15 and over present a more complex situation. It is not clear what would be the best wording for these questions. In the latest censuses, 15 countries did not include a filter question about whether or not the women had children born alive; this was done in response to a recommendation aimed at avoiding non-response, particularly among women who have no children. The reason for not including a filter question is that it seems better to ask the women directly "How many children have you ever borne alive?", in order to ensure that the enumerator is obliged to write in the number, even if it is zero. Those in favor of including the filter question argue that it is very rude to ask young

unmarried women directly how many children they have had, particularly if their mothers are respondents.

Some countries have included a question about survival of the last born child or of children born during the year preceding the census, in order to estimate infant mortality. So far, that question has not led to satisfactory results, probably because when a woman gives the date of the last live birth, she misunderstands the question and answers it in respect of the last surviving child, leading to an underestimation of infant mortality. Some experiments have been done changing in the order of the questions, hoping that this may help avoid such misunderstandings. The general thrust of these efforts is to find out: first, whether the last child born alive is still living and then to ask the date of birth.

Another aspect of questionnaire design is that of pre-coding. Beyond considerations related to data processing (Silva, 1985), it is important to ensure that pre-coding does not become an end in itself, above and beyond the concern for obtaining data. As we have seen in the example of Argentina, the question on the number of children ever born elicited incomplete information. Colombia included only pre-coded questions, even to the point of omitting the question concerning occupation of the LF. This entailed the loss of some of the most important data obtained from censuses for manpower studies.

All aspects of the census questionnaire should be fully tested in pilot and experimental censuses, as should all stages of the population census. Pilot or experimental censuses are often carried out more out of a concern for complying with a check-list of activities than out of a concern for

testing all aspects of the census and seeking better ways to obtain the information desired. In other cases, these operations are utilized only to draw conclusions about administrative problems. However the data are not processed or analyzed, both necessary steps in gauging the effectiveness of the questionnaire. Only in a few cases are written reports available on the pilot census and in even fewer cases are efforts made to discuss the experience with organizers and users.

3. CENSUS UNDERCOVERAGE

One of the most used indicators of the quality of a census is the percentage of undercoverage, i.e., the numerical and relative size of the population that was not enumerated. Omission affects demographic estimates in several ways:

i) First, census omission has a direct effect on absolute measurements, starting with the magnitude of the total population, the number of immigrants obtained from the question on place of birth and the number of emigrants obtained from the question on residence of surviving children. In these cases, omission would affect the total estimated number of migrants.

ii) Secondly, census omission affects demographic indicators, such as mean parity, death probabilities, migration rates, etc., because omission will particularly affect certain sub-populations with differential demographic behaviour.

Table 2 shows the percentages of undercoverage in the 1970 and 1980 censuses in several countries. The figure given is the percentage implicit in population projections prepared by CELADE in collaboration with national institutions (CELADE, 1987). These figures indicate that omission has risen in eight countries and in one of them it is close to 14 per cent, a figure which is higher than any recorded in previous decades.

Although these figures represent national averages, there are substantial differences within each country. Omissions are more likely to occur in certain regions or sectors of the population. Consequently, the exact internal composition of the enumerated population is not known and any conclusions drawn from analyses of these data may be distorted unless differential omissions are taken into account.

4. ITEM NON-RESPONSE

We shall consider the problems created by non-response to some of the questions included in the census schedule. Here we are not dealing with the omission of a person or group of persons, but with the gathering of incomplete information on persons who have been enumerated. This is what is usually labelled as "information not specified" or "information not known" or "information not reported", and is known as item non-response. It is important to establish the difference between this concept, in which no notation is made on the questionnaire, and the "does not know" answer which is assigned a special code to indicate that the respondent does not know the answer. Although such a response may be dealt with in the same way as non-response, it is in itself a useful piece of information for purposes of data analysis.

Census planners almost always anticipate some item non-response, which occurs with some variables more often than with others. If the percentage of non-response for a given variable is very high, this is a clear indication that at some stage the census work was not carried out correctly.

When the percentage of item non-response is very low, the problem is not serious and it is less important how the situation is solved; it may even be useful to keep the category in tabulations, since the low percentage is an indicator of a high-quality census. If a value is assigned to these cases, it should be done automatically by means of computer programs using logical criteria.

Unfortunately, in the 1980 censuses, the occurrence of non-response for some demographic questions was very high, raising questions about the reliability of the results. This was the case, in several instances, with information regarding fertility and child mortality. In some censuses, one-fourth of the answers to questions about children ever born and surviving children were left blank, and almost half were left blank in the case very young women (see the example of Ecuador in Table 3).

There is a widespread temptation to try to solve the problem of non-response by automatically assigning values for these cases by using computer programs especially designed for this purpose. We believe it is important to sound a warning about the risk involved in imputing such high percentages of information, inasmuch as this may introduce new biases in the data.

In more than one country of the region, it was possible to establish, by reviewing the schedules, that many enumerators never made notations for the category zero children in the question "How many live-births have you ever had?". Doubtless in the great majority of cases of non-response, the women concerned had no children; perhaps in a few cases, non-response was due to refusal or some other reason. The number of children per woman and hence the estimation of fertility, would be exaggerated if the number of children were assigned using as a reference the women who had reported the information, even if the reference women matched on all other characteristics.

In table 3, mean parity was estimated by considering, firstly, that the women who did not report had no children, and secondly, that they had the same fertility as those who did reply. The differences are significant, particularly for young women. The parity of the 15-19 age group is practically double when the denominator consists only of women who replied.

To analyze this problem and find a solution that will make it possible to salvage the information required, demographers use a procedure devised by El Badry (United Nations, 1983) which can be employed when there is a clear pattern of non-response according to the women's ages.

The advisability of assigning values automatically will depend partly, as mentioned above, on the percentage of data lacking, but also on the variable being studied. Although decisions on whether to assign values, as well as the criteria to be used for assignment, will depend on the case at hand, some general suggestions may be useful:

i) The original information should always be kept, without any assignment or correction. The original information is often a much more useful source for analysis, inasmuch as it allows a better evaluation of the quality of data and for the detection of patterns of error. The criteria for assignment can be applied at any time; however, if the information as originally gathered in the field is lost, it can never be recovered.

ii) Cases of non-response must be evaluated in order to determine their extent and possible causes before establishing assignment criteria.

iii) Except for the actual programming, this task should be carried out along with those persons who have taken part in other stages of the census, as they may be able to help explain the reasons for the problem. The participation of demographers and other data users is obviously important, since they can help decide whether an assignment should be made and, if so, they can help to establish the criteria.

iv) Institutional and political considerations or concern for elegant presentation of the results must not take priority over the purpose for which the information has been sought, especially when that could lead to a distortion of the results. Even if only as an anecdote, it is worthwhile mentioning that one country used, as a criterion for assigning the variable on economic activity status, a reproduction of the rates which had been estimated by means of projections made before the census. Assigning values in order to obtain a pre-established result is not only incorrect, it could also be interpreted as intentional manipulation of the information.

v) The criteria used in making an assignment must be well documented in order that the users may know clearly how the data were manipulated. Unfortunately, in most cases, it is impossible to find out how an assignment has been made.

Automatic assignment is an important aspect of census data processing. It is essential that census organizers realize that modern technologies such as computerized imputation, which allow rapid processing of large volumes of information, cannot improve the basic data available. All they can do is help facilitate their use and interpretation. Automatic assignment should never be used to conceal or disguise a problem, which would indeed lead to a distortion of the facts, with all the complications that entails.

5. DEMOGRAPHIC QUESTIONS AND THE USE OF SAMPLING IN CENSUSES

In this section , we shall consider come aspects of the use of sampling in population censuses and its effect on results refering to the components of demographic growth.

We shall analyse three uses of sampling which are commonly employed in population censuses, i.e., field sampling, advance sampling of results and post-census surveys.

a) Field sampling

To reduce costs, some countries (Argentina, Brazil, Colombia and Peru) have used sampling in their censuses of the last decade. In general this involves using two schedules, a short form which applies to the universe and a long form which applies only to a sample. The questions discussed in this paper (Table 1) -except the one on place of birth, in Peru- were included in the form used for the sample but not in the one used for the universe. Hence, sampling has a significant impact on demographic estimates resulting from these data. In the case of Peru, the application of field sampling complicated the data processing and does not seem to have been justified, since both forms used were of the long type: even the short form included many questions (9 on housing and 18 on population).

When sampling is used, the questionnaire applied to the universe should include only basic characteristics of the population; hence, one would assume that the demographic questions should be included in the long form. It does not seem advisable, however, to include in this category the questions relating to international migration, because even in countries with substantial numbers of immigrants, this is a relatively scarce phenomenon, particularly in specific regions, and thus there may not be enough information in the sample to give reliable estimates.

Argentina and Brazil, which are countries that have received a large number of migrants, included the questions on international migration in the sample, thus making it difficult to assess the impact of immigration on those countries or on the other countries of the region which provide the emigrant population.

A major risk involved in sampling has to do with how representative the sample is. One frequent tendency, as was noted in the case of Brazil, is that small households are over-represented in the sample; this has led, among other things, to an underestimation of fertility (Arretx, 1984). When enumerators have to decide in the field to which households they will apply the long sample questionnaire, they may tend to choose small families, especially if they are being paid according to the number of schedules they fill in. This type of error results in a heavy underestimation of fertility and childhood mortality. In fact, it may also introduce biases in the other variables studied, such as age-sex composition, economic characteristics, etc.

b) Advance sampling

It is useful to generate advance samples of census results, since this makes it possible to obtain timely data while the complete census is being processed. Although this might appear to be a simple task, it has proved difficult in many countries of the region. It has been said, somewhat ironically, that when a country conducts a census, it has two new sources of information: the census itself and the advance sample. In many cases the demographic estimates obtained from the sample and those obtained from the final census data have turned out to be quite different.

In one country, for example, two samples were taken and neither one could be used, as both produced deficient results. What was the reason for this? The problem might be cleared up by studying each specific case, but, at any rate, it is important that all concerned should take the greatest care to avoid such situations.

Advance samples should have a simple design -systematic and self-weighting- and, of course, they should be large enough, covering around 500 000 persons, in order to allow for demographic estimates to be made for the country as a whole, for selected groups of the population, for certain geographical regions and for other strata, as deemed advisable.

Whether to generate advance samples in the future should be considered in the light of each specific case, taking into account the size of the population and the countries capacity to utilize the techniques of modern data processing.

c. Post-census surveys

Most of the Latin American countries have carried out post-census surveys in connection with the 1980 census. It may not be advisable to conduct for following reasons, among others:

i. These surveys are very costly and usually require external financing. Up to now, the high cost has been out of proportion to the usefulness of the results obtained.

ii. Part of the administrative and executive staff who should be devoting all their time to task relating to the census itself are obliged to put additional time and effort into surveys, thus jeopardizing the success of the census.

iii. There is often a long time span between the date of the census and that of the post-census survey; because of population mobility, this is a serious disadvantage.

iv. Very few technical reports describing the methods and results of these surveys are available. Studies published on the matter are even scarcer.

v. With regard to the above, it should be noted that in most cases, the census users do not take advantage of these studies to analyse the quality of the census.

vi. When the results on census omissions have been available, the estimated values have often turned out to be lower than those which could be calculated by using indirect analytical procedures based on data taken from the census itself and from other available sources.

CONCLUSIONS

Until a few years ago, it was thought that the coverage and quality of information provided by censuses necessarily improved over time, as a result of technological advances and the economic, social and cultural development achieved by the countries of the region. This, however, has not been the case with the censuses of the last decade; on the contrary, in some countries, there has been a deterioration. The economic crisis and political conditions prevailing in recent years in Latin America may be partly responsible for this deterioration. However, a study of how the crisis may be affecting censuses and other statistical sources is beyond the scope of this paper.

Despite the situation mentioned above, the countries of the region have been aware of the usefulness of including questions aimed at obtaining information about the demographic variables determining the growth changes and structure of the population.

Although this report refers fundamentally to problems related to the collection of demographic information, particularly on fertility, mortality and

international migration, it raises issues that are of more universal concern, since it discusses problems which have arisen in various stages of the census operation.

An analysis of the errors that can occur in the course of the census exercise show that at all stages significant problems are present due to a lack of co-ordination among the different specialists participating in the census and to the introduction of new technologies. Thus, the use of new tools, such as computerization (automatic assignment, optical reading, etc.) and sampling, which might be expected to bring significant benefits, have actually, on many occasions, caused errors or complications in the census operation.

Finally, it seems wise to stress the importance of conducting a thorough study of all these matters. Such a study should include the participation of specialists from the different fields of activity involved in censuses, with a view to finding more adequate solutions so that these difficulties may be avoided in future censuses.

BIBLIOGRAPHY

- Arretx, C. (1984). Proyección de la población de Brasil, por sexo y grupos quinquenales de edad. 1950-2150. In Métodos para Proyecciones Demográficas. CELADE, Serie E, N° 1003, San José, Costa Rica.
- Brass, W. (1974). Métodos para Estimar la Fecundidad y la Mortalidad en Poblaciones con Datos Limitados. CELADE, Serie E, N° 14, Santiago, Chile.
- Brass, W. 1977. Cuatro Lecciones de William Brass. CELADE, Serie D, N° 91, Santiago, Chile.
- CELADE. (1986). Boletín Demográfico N° 37. Santiago, Chile.
- CELADE. (1987). Boletín Demográfico N° 40. Santiago, Chile.
- CELADE and INDEC (1985). Los Censos de Población del 80. Taller de Análisis y Evaluación. Buenos Aires, Argentina.
- CELADE, INDEC AND CENEP (1987). Los Censos del 90. Características Económicas de la Población. Buenos Aires, Argentina.
- García, Estela. (1984). "Desigualdades Socio-económicas y Mortalidad Infantil. Santa Fé, Argentina". CELADE, Santiago, Chile. (Unpublished thesis).
- INEGI (1986). Taller Nacional de Evaluación del Censo General de Población y Vivienda, 1980. San Juan del Río, Querétaro, México. Volúmen 1 y 2.
- INEGI and UNFPA (1987). Memoria del Taller Interamericano de Planeación de Censos de Población y Vivienda. Cuernavaca, México.
- Silva, A. (1985). Procesamiento de Datos Censales: Lecciones de la Experiencia Latinoamericana. CELADE. Serie A No. 170, Santiago de Chile.
- Somoza, J. (1980). Estimaciones indirectas de la emigración. Aplicaciones de los procedimientos basados en información sobre la residencia de hijos y de hermanos. In Notas de Población N°23. CELADE, Santiago.
- Somoza, J. (1987). Censo Experimental de Junin de los Andes. Resultados y Análisis. CELADE, Fundación Cruzada Patagónica and ACDI-Canadá. Santiago de Chile.
- United Nations. (1983). Manual X Indirect Techniques for Demographic Estimation. New York.
- Zaba, B (1986). Measurement of Emigration using indirect Techniques. Ordina Editions, Liège, Belgique.

Table 1

QUESTIONS ON THE COMPONENTS OF DEMOGRAPHIC CHANGE INCLUDED IN
1980 CENSUSES IN LATIN AMERICA

	Adult mortality		Fertility and child mortality				International migration		
	Maternal Orphan- hood	Deaths in last 12 months	Children ever born	Children still living	Births in the last year	Surviving children born in the last year	Surviving children living abroad	Place of birth	Year of arrival in country
Argentina * (1980)	-	-	x	x	x(b)	-	-	x	x
Bolivia (1976)	-	-	x	x	x(a)	-	-	x	-
Brazil * (1980)	x	x	x	x	x(a)	-	-	x	-
Chile (1982)	-	-	x	x	x(b)	-	-	x	-
Colombia * (1985)	x	-	x	x	x(a)	x	x	x	-
Costa Rica (1984)	-	-	x	x	-	-	-	x	x
Cuba (1981)	-	-	x	x	-	-	-	-	-
Ecuador (1982)	-	-	x	x	x(a)	x	-	x	-
Guatemala (1981)	x	-	x	x	x(a)	x	-	x	x
Haiti (1982)	-	x	x	x	x(a)	-	x	x	-
Mexico (1980)	-	-	x	x	x(a)	-	-	x	-
Panama (1980)	x	-	x	x	x(a)	x	-	x	-
Paraguay (1982)	x	-	x	x	x(a)	x	x	x	x
Peru (1981)	x *	-	x *	x *	x(a)	x *	-	x	-
Dominican Rep. (1981)	x	-	x	x	x(a)	x	x	x	x
Uruguay (1985)	-	-	x	x	x(b)	-	x	x	x
Venezuela (1981)	-	-	x	x	x(b)	-	-	x	x

* Study by sample

(a) Date of birth of last child

(b) Birth during the year prior to census

Figure 1

INFANT MORATLITY RATES (IMR) ESTIMATES IN SELECTED LATINA AMERICAN COUNTRIES,

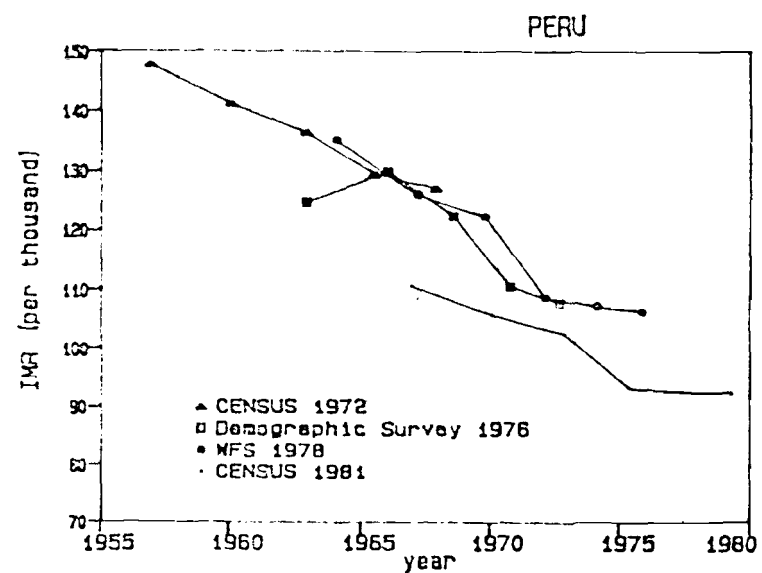
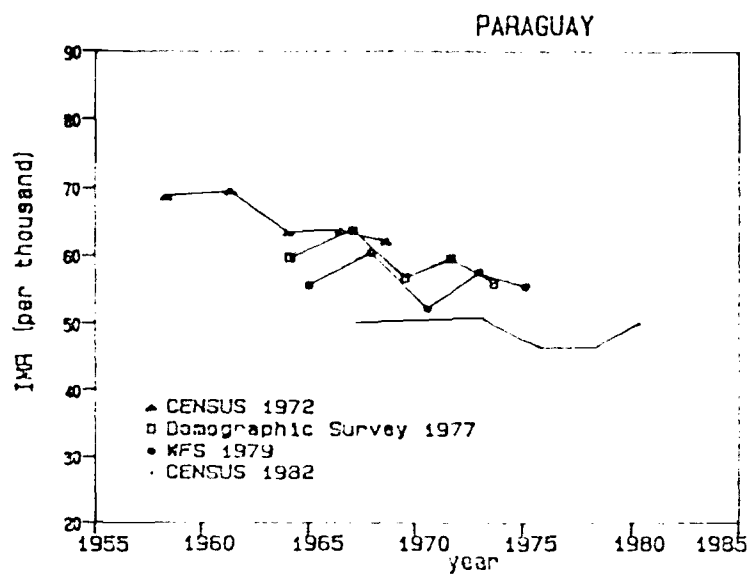
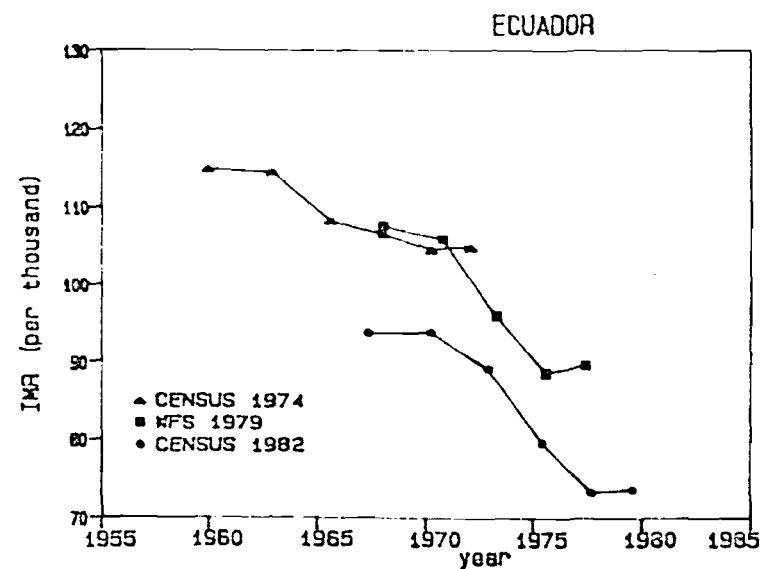
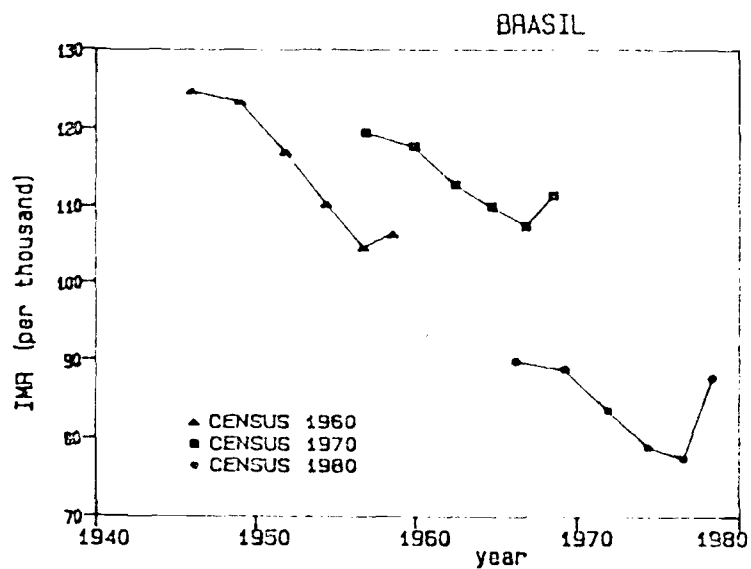


Table 2

PERCENTAGES OF OMISSION IN 1970 AND 1980 CENSUSES IN
SOME LATIN AMERICAN COUNTRIES

Country	Percentage of omission	
	1970 censuses	1980 censuses
Argentina	2.8	1.5
Brazil	2.8	1.8
Costa Rica	- 0.5	6.0
Dominican Republic	8.2	2.1
Ecuador	4.4	7.5
Guatemala	10.4	13.8
Haití	6.7	11.4
Mexico	7.9	5.2
Panama	3.6	4.7
Paraguay	5.2	10.8
Peru	2.3	4.0
Venezuela	3.5	6.7

Table 3

ECUADOR, 1982 CENSUS: REPORTING OF NUMBER OF CHILDREN EVER BORN
AND ESTIMATION OF AVERAGE NUMBER OF CHILDREN EVER BORN PER
WOMAN (MEAN PARITY)

	15-19	20-24	25-29	30-34	35-39
<u>Women</u>					
Total	440 225	394 682	316 908	252 622	204 310
Reporting	231 873	292 814	274 933	231 213	189 393
Not reporting	208 382	101 868	41 975	21 409	14 917
Percentage not reporting	47,3	25,8	13,2	8,5	7,3
Children reported	80 200	432 796	730 320	885 141	943 003
<u>Mean parity</u>					
Total women in denominator (a)	0,18	1,10	2,31	3,50	4,62
Women reporting in denominator (b)	0,35	1,48	2,66	3,83	4,98

(a) This is equivalent to assuming 0 children for women not reporting.

(b) This is equivalent to assuming that women not reporting have the same fertility as women reporting.

RESUME

Cet article se réfère aux expériences des pays de l'Amérique Latine concernant les problèmes de collecte de l'information démographique qui découlent des différentes étapes de l'opération censale.

On considère les problèmes de collecte de données sur la fécondité, la mortalité et la migration internationale au travers des rubriques implicitement ouvertes par les questions incluses dans les recensements, la conception du questionnaire, les omissions, la non-réponse et l'échantillonnage sur le terrain.

L'accent est mis sur la nécessité de collaboration entre l'ensemble des participants (organismes du recensement, utilisateurs et d'autres spécialistes), surtout au moment de la planification, dans la mesure où des obstacles peuvent se présenter en raison du manque de coordination et de l'introduction inadéquate de nouvelles technologies.

Certaines conclusions sont tirées de ces expériences d'où dérivent des recommandations proposées pour les recensements futurs.

