Scientific Reports

NUMBER 80 JULY 1985

BENSON C. MORAH

Evaluation of the Nigeria Fertility Survey 1981–2

INTERNATIONAL STATISTICAL INSTITUTE
Permanent Office. Director: E. Lunenberg

Mailing Address:
428 Prinses Beatrixlaan, PO Box 950
2270 AZ Voorburg
Netherlands

WORLD FERTILITY SURVEY
Project Director: Halvor Gille

The World Fertility Survey is an international research programme whose purpose is to assess the current state of human fertility throughout the world. This is being done principally through promoting and supporting nationally representative, internationally comparable, and scientifically designed and conducted sample surveys of fertility behaviour in as many countries as possible.
The WFS is being undertaken, with the collaboration of the United Nations, by the International Statistical Institute in cooperation with the International Union for the Scientific Study of Population. Financial support is provided principally by the United Nations Fund for Population Activities and the United States Agency for International Development.
This publication is part of the WFS Publications Programme which includes the WFS Basic Documentation, Occasional Papers and auxiliary publications. For further information on the WFS, write to the Information Office, International Statistical Institute, 428 Prinses Beatrixlaan, Voorburg, The Hague, Netherlands.

L'Enquête Mondiale sur la Fécondité (EMF) est un programme international de recherche dont le but est d'évaluer l'état actuel de la fécondité humaine dans le monde. Afin d'atteindre cet objectif, des enquêtes par sondage sur la fécondité sont mises en oeuvre et financées dans le plus grand nombre de pays possible, Ces études, élaborées et réalisées de façon scientifique, fournissent des données représentatives au niveau national et comparables au niveau international.
L'Institut International de Statistique avec l'appui des Nations Unies, a été chargé de la réalisation de ce projet en collaboration avec l'Union Internationale pour l'Etude Scientifique de la Population. Le financement est principalement assuré par le Fonds des Nations Unies pour les Activités en matière de Population et l'Agence pour le Développement International des Etats-Unis.
Cette publication fait partie du programme de publications de l'EMF, qui comprend la Documentation de base, les Documents Non-Périodiques et des publications auxiliaires. Pour tout renseignement complémentaire, s'adresser au Bureau d'Information, Institut International de Statistique, 428 Prinses Beatrixlaan, Voorburg, La Haye, Pays-Bas.

La Encuesta Mundial de Fecundidad (EMF) es un programa internacional de investigación cuyo propósito es determinar el estado actual de la fecundidad humana en el mundo. Para lograr este objetivo, se están promoviendo y financiando encuestas de fecundidad por muestreo en el mayor número posible de países. Estas encuestas son diseñadas y realizadas científicamente, nacionalmente representativas y comparables a nivel internacional.
El proyecto está a cargo del Instituto Internacional de Estadística en cooperación con la Unión Internacional para el Estudio Científico de la Población y con la colaboración de las Naciones Unidas. Es financiado principalmente por el Fondo de las Naciones Unidas para Actividades de Población y por la Agencia para el Desarrollo Internacional de los Estados Unidos.
Esta publicación ha sido editada por el Programa de Publicaciones de la EMF, el que incluye Documentación Básica, Publicaciones Ocasionales y publicaciones auxiliares. Puede obtenerse mayor información sobre la EMF escribiendo a la Oficina de Información, Instituto Internacional de Estadística, 428 Prinses Beatrixlaan, Voorburg-La Haya, Países Bajos.

Scientific Reports

Evaluation of the Nigeria Fertility Survey 1981–2

BENSON C. MORAH
National Population Bureau
Lagos, Nigeria

The recommended citation for this publication is:

Morah, Benson C. (1985). Evaluation of the Nigeria Fertility Survey 1981–2. *WFS Scientific Reports* no 80. Voorburg, Netherlands: International Statistical Institute

Typeset and printed in Great Britain by
H Charlesworth & Co Ltd, Huddersfield.

Contents

FIGURES

Preface

One of the major objectives of the World Fertility Survey programme was to assist the participating countries in obtaining high quality data through national fertility surveys. The high standards set by the WFS were expected to yield better quality data than typically obtained in the past, but this expectation in no way obviated the need for a detailed assessment of the quality of the data. It is recognized that such an evaluation will not only alert the analysts by identifying defects, if any, in the data, but also may throw light on the shortcomings of the WFS approach, which can be taken into account in the design of future fertility surveys.

It is in this context that, as part of its analysis policy, the WFS initiated a systematic programme for a scientific assessment of the quality of the data from each survey. A series of data evaluation workshops was organized at the WFS London headquarters with the dual objective of expediting this part of the work and of providing training in techniques of analysis to researchers from the participating countries. Working in close collaboration with WFS staff and consultants, participants evaluated the data from their respective surveys after receiving formal training in the relevant demographic and data processing techniques.

The present document reports on the results of the evaluation of the data of the Nigeria Fertility Survey of 1981–2 and was prepared by Dr Benson C. Morah, who was the Survey Director of the NFS.

Dr Shea Oscar Rutstein, co-ordinator of these workshops and Enrique Carrasco (WFS co-ordinator for the NFS) assumed a major responsibility in the successful completion of the work. Nuri Ozsever provided much valuable assistance, while many other staff members also made significant contributions.

HALVOR GILLE
Project Director

1 Introduction

1.1 OBJECTIVES OF THE NIGERIAN FERTILITY SURVEY (NFS)

The Nigeria Fertility Survey (NFS) was conducted between October 1981 and August 1982 by the National Population Bureau (formerly National Population Commission) of the Federal Government of Nigeria and in collaboration with the World Fertility Survey (WFS). The survey was jointly financed by the Government of the Federal Republic of Nigeria and the United Nations Fund for Population Activities (UNFPA). The primary objectives of the survey are to obtain accurate and reliable information on the levels, patterns and trends of fertility in the country, to discern any differentials in fertility between the various subgroups of the population and to ascertain the extent of knowledge and practice of contraception. Because of the paucity of nationwide demographic data for the country, the survey was also to provide bench-mark data with which the results of subsequent surveys can be compared; it was also to serve as a model for future surveys on fertility or other demographic processes.

The initial results of the survey have already been published in the First Country Report on the project and in its Summary of Findings. For the benefit of readers who have not had access to either the report or the summary of findings, a brief description of the survey is represented here.

1.2 CHARACTERISTICS OF THE SURVEY

Data for the NFS were collected from a nationwide probability sample of 250 enumeration areas (EAs). These EAs were subsampled from the 912 EAs (48 from each of the 19 states in the country) that had constituted the sample for the National Demographic Sample Survey (NDSS) conducted by the National Population Bureau in 1980. The NFS subsample (182 in rural and 68 in urban areas) was drawn in such a way as to yield a PPS sample, ie with the probability of the selection of each EA made proportionate to a measure of size so as to yield a self-weighting sample. This attempt to make the sample self-weighting was, however, subsequently abandoned and the data had to be weighted to compensate for the unequal probabilities of selection in each EA.

The actual data collection consisted of two components: a household survey and a survey of women aged 15–49 (irrespective of marital status) who were *de facto* residents of surveyed households. The household survey was conducted by means of a household schedule in which all the members of sample households were listed. The following information was also collected on each listed member: name, relationship, residential status (*de jure* and *de facto*), sex, date of birth, age and eligibility for the individual interview. To be eligible for the individual interview, the household member had to be female, aged between 15 and 49 and must have slept in the household the night before the (household) interview. Further information was also collected on the environment of each household and the possession of selected consumer durable items. The individual survey was conducted by means of an individual questionnaire which was based on the WFS core questionnaire and incorporated the WFS module on factors other than contraception affecting fertility (FOTCAF) and elements of the family planning module. The individual questionnaire consisted of the following sections:

1 respondent's background;
2 maternity history;
3 birth intervals and breastfeeding;
4 marriage history;
5 contraceptive knowledge and use;
6 fertility preferences;
7 work history;
8 current (or last) husband's background.

Both the household schedule and the individual questionnaire were translated into six of the Nigerian languages namely Hausa, Yoruba, Ibo, Efik, Nupe and Kanuri.

The survey was designed to collect information from approximately 10 000 eligible women. The fieldstaff were specially recruited and trained for the exercise and were grouped into 30 teams, each consisting of a supervisor (male), an editor and 4–5 interviewers (females). Each team worked in the state from which its members had been recruited.

The application of the survey design in the field yielded a sample of 9361 households for the households survey; 9236 of these households were occupied at the time of interview. Out of these occupied sample households, interviews were successfully completed with 8624, yielding a response rate of 93.4 per cent. The number of women eligible for individual interview in the successfully interviewed households was 10 134; out of these, interviews were successfully completed with 9727, yielding a response rate of 96.0 per cent.

1.3 OBJECTIVE OF THIS DOCUMENT

The WFS, as a matter of policy, typically encourages further analysis of data collected in particpating countries after the publication of the First Country Report. The evaluation of the quality of data collected in each

survey is given high priority among such subsequent analyses. This is the main objective of this document. The quality of the data collected in the NFS is evaluated here with the objective of discovering the types of errors found in the data, the sources of the errors and their effects on the demographic estimates as published in the First Country Report and the Summary of Findings. An evaluation of this type will be of invaluable help to researchers who want to do the necessary further analyses of the data and to planners and policy makers who would wish to know how reliable the published findings are. Such an evaluation typically involves internal checks on the consistency of the data, internal comparison of the data from different survey instruments (or parts of the same instrument) and comparisons of the data with external sources or theoretical models. Since external data sources on a national scale are not available for the country (except the 1963 population census), the evaluation of the NFS data to be undertaken here will unavoidably be confined to internal consistency checks and comparisons or checks with any adequate and available theoretical models.

The actual evaluation is, however, preceded in the next chapter by a brief discussion of possible sources of error in survey data of the NFS type and typical consequences of such errors.

2 Quality of Reporting

2.1 INTRODUCTION

Information collected in fertility surveys of the NFS type is typically prone to a number of errors and biases which may affect its quality and thus the reliability and accuracy of estimates derived from it. Such errors and biases may occur right from the planning to the implementation and data processing stages. There could be errors in the design of the sample and in the field application of that design. The selected and surveyed sample may not be truly representative of the universe of interest; some subgroups may be over or under-represented. The extent of coverage may not be the same for all the sampling units. The net effect of these may not be universally applicable or true for the entire universe or subgroups of the universe. Errors could also arise from the procedures for selecting the ultimate sampling units or respondents and from non-response. Non-respondents may differ from respondents with respect to the information of interest; when this is actually the case, biases may be introduced in the obtained data. There may also be errors in the design of the questionnaire or other instruments of data collection that may introduce systematic biases or errors in the information collected. Much of the information collected in these surveys is retrospective. Such information is particularly prone to misreporting. Dates of events may be wrong or the events completely omitted. Finally, errors and biases may be introduced to the data at the coding, editing, transcription or processing stages.

More details on a few of the above sources and types of error, particularly those that are crucial to the evaluation of the NFS data, are briefly presented below. These include errors that could arise from the selection procedures adopted and in the dating and omission of events covered in the survey.

2.2 SELECTION PROCEDURES

Fertility surveys differ in the definition of women eligible for interview. In some, all women of childbearing age (typically defined as aged between 15 and 49 years) in the sample areas are covered. In others only such women in legal (or consensual) unions are included. Further restrictions such as *de facto* or *de jure* residence may be introduced in some surveys. In the NFS, women eligible for individual interview were defined as those of childbearing age (15–49) who were *de facto* residents of sample households. *De facto* residence was defined as having spent the night before the (household) interview within the household.

The field application of this definition might lead to erroneous exclusion or inclusion of some women. Since eligibility in the NFS was determined from the list and characteristics of household members during the household survey, wrong ages of women given by the respondent to the household survey (who may or may not have been the women themselves) or wrong residential status may have led to wrongful exclusions or inclusions. This would particularly affect possible eligible respondents at the lower or upper age boundaries especially in households with many members and in instances where age had to be estimated. Wrong application of the definition of the household could also lead to such erroneous inclusions or exclusions.

2.3 DATING OF EVENTS

Another potential source of error is the dating of events collected in fertility surveys, especially in cultures where not much importance is attached to dates. Misreporting of dates of such events as the birth of the respondent, the birth of her children, marriages and remarriages often lead to distortions in the data collected and in age or period specific rates based on the data.

Respondent's age or date of birth

Misreporting the age (or date of birth) of respondents often results from actual ignorance of the age or date of birth and the preference or avoidance of reporting dates or age in particular digits. Typically the preference is for reporting age in numbers ending in 0 and 5 or in even numbers, and for the avoidance of the other odd numbers. This leads to the concentration of respondents in the preferred digits (heaping). Even if the respondent had given her date of birth instead of age, there is also the tendency to establish that date by relating it to the interview date so that age also ends in rounded numbers. Misreporting is worse when age has to be estimated, either by the respondent herself or by the interviewer. Estimates based on physical appearance or on the demographic characteristics of the respondent (such as her number of children) are particularly prone to error and often end in preferred digits. The net effect of these errors in the reporting of age or date of birth is that the curve of the obtained age distribution is no longer smooth but assumes a saw-edged pattern with the peaks corresponding either to rounded ages or rounded durations from the date of interview.

Misreporting dates has a serious effect on estimated fertility (or other data) based on the age of the women. The misreporting leads to transference of age in which some women report lower or higher ages than they really are and may be classified in either higher or lower age groups than the ones to which they actually belong. The

effect of such transference on estimated fertility rates depends on the direction of the transference (to lower or higher age groups) and to whether it is selective with respect to fertility. For instance, if some women actually aged 40–44 at the time of interview were reported as aged 45–49, this transference would bias cumulative fertility of the 45–49 age group downwards since older women tend to have given birth to more children than younger women. Their transference would also bias estimates of current fertility for the 45–49 age group upwards since current fertility is higher for younger than for older women. The reverse would be the case if some women actually aged 45–49 were reported as 40–44. The direction of these biases will hold true for all women over 30 years old whose ages have been transferred upwards or downwards. For women aged 15–24, the situation is reversed while for women aged 25–29, the direction of the bias in current fertility will be indeterminate. The overall effect of these transferences on estimates will depend on the extent, direction and cancelling effects of the transference. Further distortions may arise depending on whether or not such a transference is also associated with misreporting dates of live births and the direction of such misreporting.

Dating of live births

The accuracy of fertility estimates also depends on the correct reporting of the dates of birth of all children recorded in the birth or pregnancy history section of NFS type surveys. In the NFS, all the children a woman had given birth to were listed in chronological order starting with the first and the date (or years ago) of birth obtained, in addition to sex, whether the children had survived and if not the date (or age) of death. Such a history of births is subject to misreporting dates; such misreporting may be systematic, eg transferring the dates of birth nearer or further away from the interview date. This may be more characteristic of births that occurred in the distant past or very recently. For instance, working with data from the Bangladesh Fertility Survey, Brass (1978) found that some births which occurred in the five-year period before the survey had been transferred to the previous (five–nine-year) period, and that some births which had taken place in periods further away were brought forward to the same (five–nine-year) period. Such forward and backward displacement of dates of births distort the analysis of fertility trends and produce a spurious decline or exaggerated decline in fertility levels. An earlier study with data from West New Guinea (Brass 1974) had found evidence of a shift in fertility to periods further removed from the date of survey, presumably because interviewers had assumed that women started childbearing at an earlier age. Fertility for the earlier periods was consequently overestimated and a false decline in recent fertility reported.

Potter (1977a) had developed a simulation model to discover the extent to which such forward and backward displacement of births would affect fertility estimates obtained from birth data. The model assumed that the more distant the births are from the interview date, the less exactly the woman remembers the dates of birth, and if the dates of birth are obtained in the order in which they occurred (ie starting from the first), then the date given by the woman for the subsequent births is influenced by the date she has given about the first or previous births. Thus, the model assumes that women report their births, particularly those most distant from the survey date, in intervals between births and that dates of birth may be brought forward because of the reporting of a later date for the first child or the exaggeration of birth intervals. Potter (1977b) later confirmed the assumptions of the model with data from fertility surveys of Bangladesh and El Salvador.

Birth data from surveys are collected from respondents who are survivors of their birth cohorts. There is the implicit assumption in using the data that their fertility does not differ significantly from the fertility of those who have died. Should this assumption not be true, then estimates based on the data may be biased, particularly for periods more distant from the survey. If female (especially maternal) mortality is high and is positively related to the number of children ever born, then the obtained estimates may be biased downwards.

Dating of marriages

Most human reproduction takes place within stable sexual unions (marriages); consequently many fertility estimates are based on the duration of exposure to the risk of childbearing within marriages. The accuracy of such marital rates depends on the correct reporting of the beginning and end of such unions from which the woman-years of exposure are calculated. Just like age or date of birth of children, respondents often do not remember (or know) their exact date (or age) at marriage. Estimates of this date or age, especially when based on the reproductive history of the woman, are often wrong and often end in preferred digits, thus resulting in considerable heaping. Some respondents may also give the date of their second or subsequent marriage as the date of their first marriage. There also appears a tendency for women, who incorrectly report their age or date of birth, to misreport their dates of marriage (or marital duration). Such misreporting introduces errors in such important variables for fertility analyses as age at first marriage as well as marital fertility rates.

2.4 OMISSION OF BIRTHS AND MARRIAGES

While the dates of some births and marriages may be misreported, some births or marriages may be completely omitted from the birth and marriage histories. Such omissions are assumed to be more characteristic of older than younger women and to involve births and marriages that occurred in the distant past from the survey date. However, more recent births may also be omitted, especially if they occurred in unstable (or outside) unions. Children who died in infancy or were living away from home are also often omitted. In societies with strong preferences for sons, female births may be omitted. When omissions are more frequent for births that occurred in the more distant past, the effect of such omissions is to bias the fertility in those periods

downwards, possibly resulting in a spurious increase in fertility at subsequent periods. Cumulative fertility for the older women may also be lower than is actually the case. When very young children at the time of survey are omitted, recent fertility levels may be underestimated.

In surveys where marriages are defined to include more or less stable sexual unions (as in the NFS), some stable non-legalized unions that occurred in the past may be erroneously omitted; also not-so-stable unions may be included. Since it is assumed in most surveys that a woman is typically not exposed to the continuous risk of childbearing between marriages, omission of marriages may lead to incorrect calculation of woman-years of exposure on which marital fertility rates are based. The net effect of such omissions are, however, low in societies where marriages are stable, remarriage frequent and soon after marital dissolution and where women spend most of their lives in the married state after their first marriage.

3 Age Reporting

3.1 SOURCES OF DATA ON AGE

There are two independent sources of data on age in the NFS. The first source is the household schedule in which the date of birth (where known) and the age of all listed household members were obtained. The second source is the individual questionnaire in which questions were asked on the respondent's date of birth and age.

Information on date of birth and age for all household members was given by the respondent to the household schedule who may be the household head, the respondent to the individual questionnaire or another adult household member. Thus, except for the particular respondent, information on date of birth and age in the household schedule is provided by proxy and probably more liable to error than if the members had each provided such information themselves. As it was expected that some respondents to the household schedule may not know the dates of birth (in month and year) of all household members, the interviewers were instructed to accept 'don't know' as a correct response for date of birth. They were, however, instructed to obtain age (in completed years) for all members listed; the response 'don't know' was not to be accepted and, in instances where the respondent did not know the age of a particular household member, some means had to be found to estimate the required age. Such means include asking the individual himself/herself, asking other household members, reference to some documents (such as baptismal certificates), or obtaining the age from non-household members or even neighbours. When age could not be obtained from any of the above means, then the interviewer herself was to estimate the age from whatever clues that were available to her including, as a final resort, the physical appearance of the person. Unfortunately, no provision was made in the household schedule for indicating who the respondent was, who provided answers to the questions on date of birth or age, whether age was estimated by the interviewer and on what the estimate was based.

In the individual questionnaire, information on date of birth and age was obtained directly from each respondent through the following questions:

Date of birth
'Do you know your date of birth?'
If known: 'In what month and year were you born?'
Age
'How old are you?'

All respondents were first asked whether they knew their dates of birth; if they did, the month and the year were obtained. If they did not know, then they were asked the question on age. As a check, respondents who knew and had provided their dates of birth were also asked their age. The interviewers were instructed to check and correct any inconsistency between age and date of birth. If the respondent was the person who had supplied information on her date of birth and age during the household interview, her answers could be copied in response to the above questions. If the respondent did not know both her date of birth and age, then the interviewer was to indicate how age was estimated. The interviewer would then enter the date of birth or age on the date event chart provided in the questionnaire.

In the computation of the age of respondents to the individual questionnaire as presented here and in the principal report, greater emphasis was placed on date of birth if there were any inconsistencies between that and reported age. Where only age was given, the date of birth was imputed with the month randomly assigned within the calendar year. All data presented here (and in the principal report) are based on age as entered in the household schedule; age was not recalculated from the date of birth.

Data on age as obtained above are subject to various types of error as already described in section 2 of this report. Such errors could arise through misstatement of age out of ignorance, preference for certain digits or deliberate falsification of age either by the respondents or the interviewers. Proxy age reporting is particularly prone to error since the proxy is not likely to know the exact ages or dates of birth of all household members especially in large and extended households. Estimation of age by the interviewer, especially when based on events in the respondent's life, eg number of children born, age of the eldest child, or on physical appearance, is particularly prone to error and to the obtained age ending in a preferred digit. Even when such estimates are based on documentary evidence, they are still likely to be biased because documents obtained later in life are themselves based on estimates of age. Such errors often result in the concentration of the sample population at particular ages or digits (heaping) or in age transference. The extent of these errors is investigated below for age data as obtained from the household schedule and from the individual questionnaire. The WFS experience in many African countries shows that data on age, and, in fact, on the dating of vital events are almost always poor (see, for instance, Rutstein 1984). Where possible, comparisons with age data from the 1963 Nigerian census are made.

3.2 AGE IN THE HOUSEHOLD SCHEDULE

Errors in age data from censuses and surveys are often manifest in the age and sex distribution of the enumerated population. Theoretically, the distribution of a

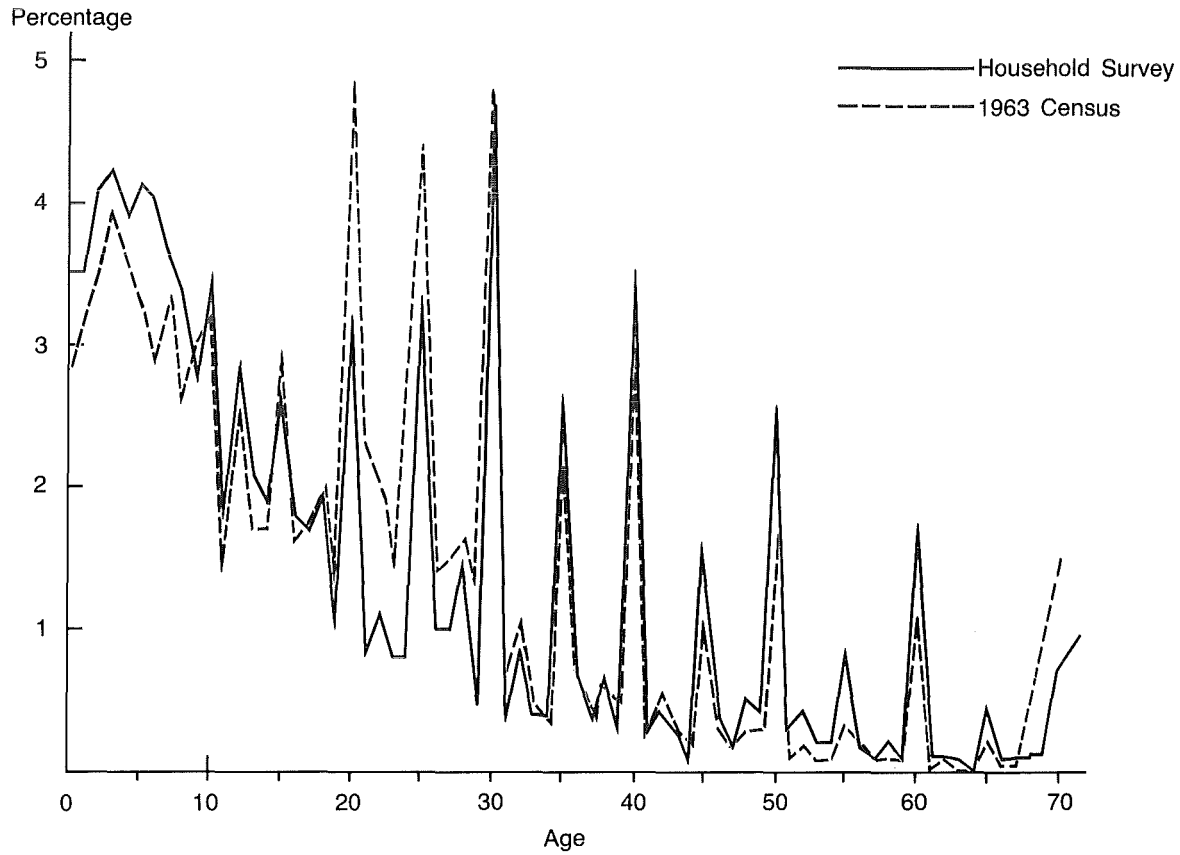

Figure 1 Percentage distribution of the *de facto* population enumerated in the household survey and the 1963 census, by single years of age

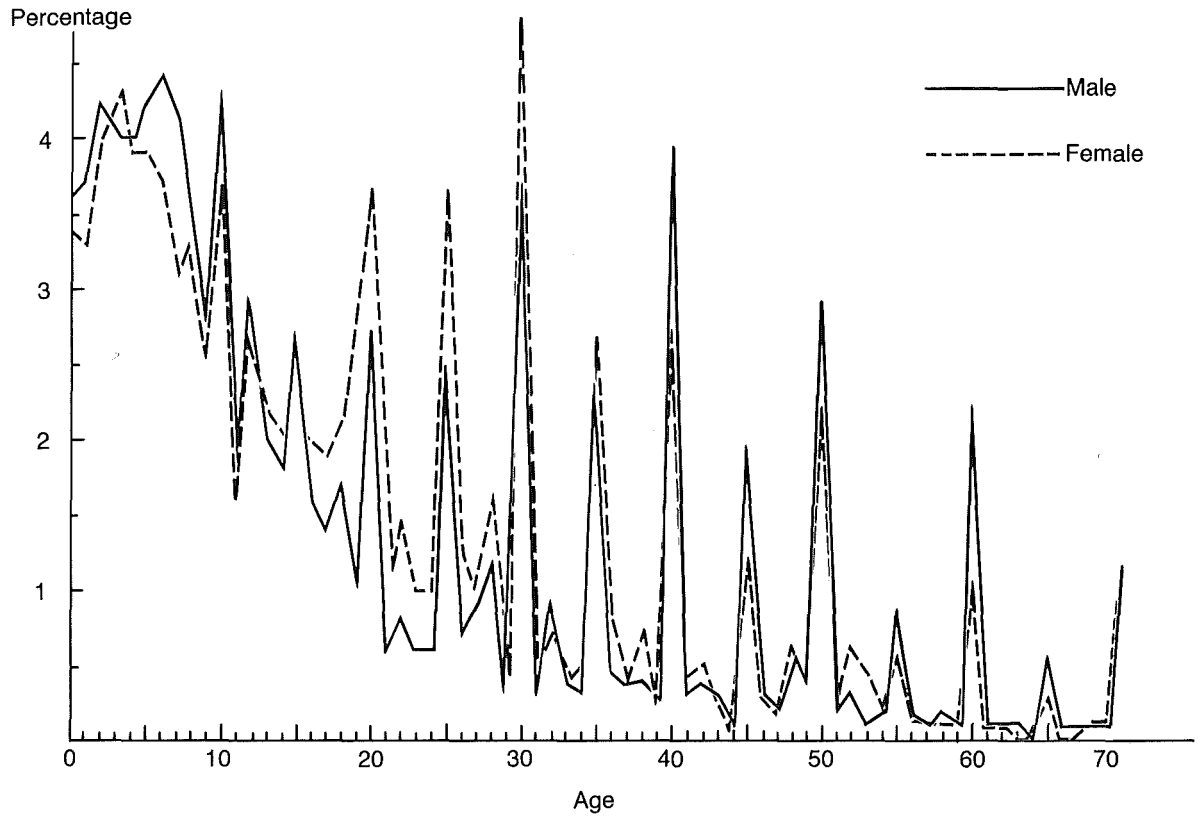

Figure 2 Percentage distribution of the *de facto* population enumerated in the household survey and the 1963 census, by sex

population by single years of age should be fairly smooth and, at any given single year of age, there should be slightly fewer persons than in the preceding year. Although the distribution of any population will hardly approximate this because of selective immigration or emigration, changing fertility and mortality and wars, huge unexplained deviations are often found in the obtained age data. This is the case if heaping at ages ending in particular digits is observed without any accompanying evidence of cyclical variations in births corresponding exactly with these digits.

The distribution of the *de facto* population enumerated in the NFS household survey and in the 1963 census in single years of age is shown in figures 1 and 2 for the total population by sex. Instead of a relatively smooth distribution, the familiar saw-edged pattern indicative of heaping is evident. This is particularly the case from age 10 onwards, with considerable heaping at digits ending in 0 and 5, and to a much lesser extent, on 8 and 2. Ages ending in digits 1, 9 and 3 show sizeable troughs indicating avoidance of those digits. Figure 2 shows that heaping is more acute for females under age 30 and greater for males thereafter. The distribution of the rural and urban populations in single years by sex, and shown in figures 3 and 4 respectively, show a similar pattern of heaping at ages ending in 0, 5 and avoidance of ages ending in 9 and 1. The heaping also appears greater in the rural areas where it is greater for females under 40 years compared with males; in the urban areas, it appears greater for females under 35 years compared with males. Such a pattern of peaks and troughs in the distribution is indicative of faulty age reporting.

The preference or avoidance of particular digits which is the primary cause of the type of age distribution obtained above can be measured by certain indices. The Myers' index, which is used here, yields an index of preference or avoidance of each terminal digit, representing the deviation from 10 per cent of the proportion of the total population reporting on the given digit. Half of the sum of the absolute deviations is taken as the summary index of preference for all terminal digits and is interpreted as an estimate of the minimum proportion of persons in the population for whom age with an incorrect final digit has been reported. The index ranges from 0, indicating no preference or heaping to 90 if all ages had been heaped on a single digit.

Myers' indices for the rural, urban and total populations enumerated in the household survey by sex are presented in table 1. The indices further confirm the extensive heaping at digits 0 and 5 for both sexes and in all populations. The summary indices show that heaping is greater for males than females and greater in rural

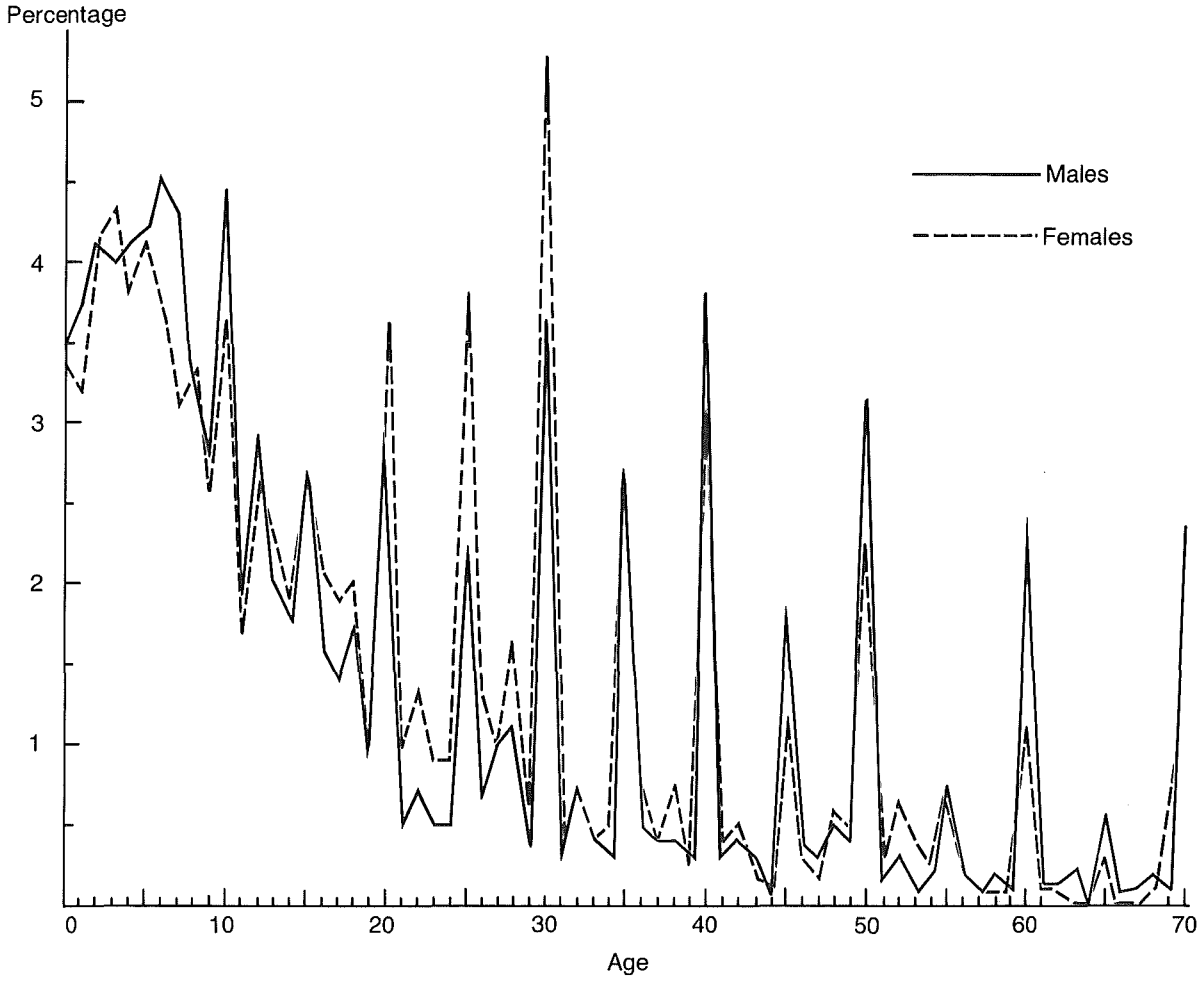

Figure 3 Percentage distribution of the *de facto* population enumerated in the household survey and the 1963 census, by sex, rural areas

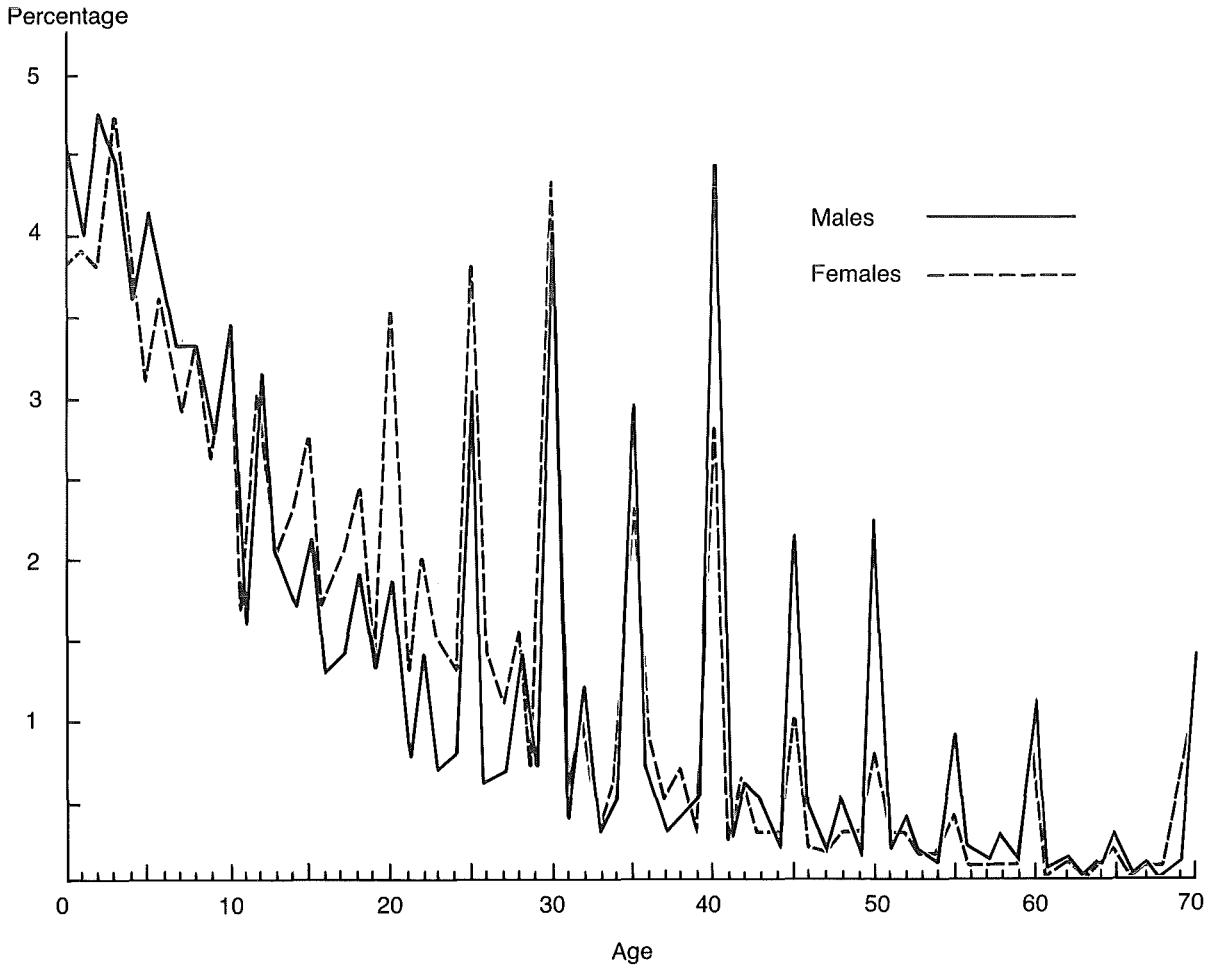

Figure 4 Percentage distribution of the *de facto* population enumerated in the household survey and the 1963 census, by sex, urban areas

Table 1 Myers' indices[a] of digit preference for the population enumerated in the household survey, by sex and rural-urban status

Digit	Rural			Urban			Total		
	Males	Females	All	Males	Females	All	Male	Female	All
0	26.1	20.3	23.1	19.6	15.7	17.6	25.1	19.6	22.2
1	−6.4	−5.3	−5.8	−5.8	−4.5	−5.1	−6.3	−5.2	−5.7
2	−3.6	−2.6	−3.0	−0.7	−0.5	−0.6	−3.1	−2.3	−2.7
3	−5.5	−4.6	−5.0	−4.8	−4.1	−4.4	−5.4	−4.5	−4.9
4	−5.8	−4.9	−5.3	−4.8	−3.3	−4.0	−5.7	−4.6	5.1
5	10.4	9.6	9.9	11.1	8.4	9.7	10.5	9.4	9.9
6	−3.6	−2.3	−2.9	−4.2	−2.7	−3.4	−3.7	−2.4	−3.0
7	−3.8	−4.0	−3.9	−5.0	−3.3	−4.1	−4.0	−3.9	−3.9
8	−2.1	−0.8	−1.4	−1.5	−0.8	−1.1	−2.0	−0.8	−1.4
9	−5.7	−5.4	−5.6	−3.8	−4.9	−4.4	−5.4	−5.3	−5.4
Summary index[b]	36.5	29.9	33.0	30.7	24.1	27.2	35.6	29.0	32.1

[a]Deviations from 10 per cent.
[b]Calculated as half of the sum of the indices.

than in urban areas. Age ending in an incorrect digit was reported for at least a third (32.1 per cent) of the enumerated population of both sexes (35.6 per cent for males and 29.0 per cent for females).

A further check on the quality of age reporting, this time for age groups, can be done by examining the distribution of the 5-year age groups as shown in figures 5a to 5c and table 2, and by the use of age ratio and

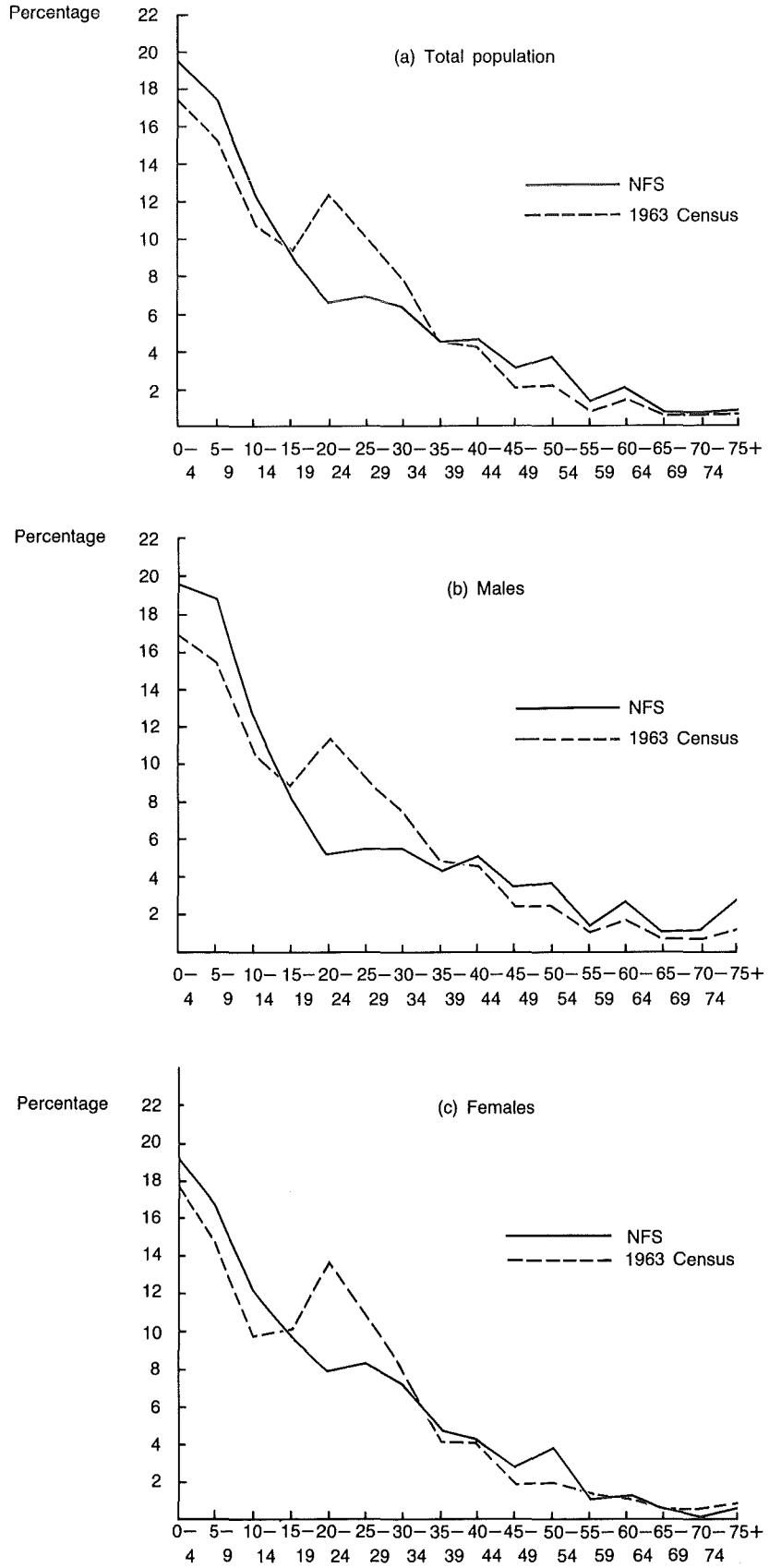

Figure 5 Percentage distribution of the *de facto* population enumerated in the household survey and the 1963 census, total and by sex

Table 2 Percentage distribution of the *de facto* population enumerated in the household survey and in the 1963 census, by five-year age groups and sex

Age group	Males		Females		Total	
	NFS	1963 Census	NFS	1963 Census	NFS	1963 Census
0–4	19.5	16.8	19.0	17.6	19.3	17.2
5–9	18.9	15.5	16.6	14.8	17.7	15.2
10–14	12.7	11.6	12.1	9.7	12.4	10.7
15–19	8.3	8.9	9.7	10.0	9.0	9.4
20–24	5.2	11.2	7.9	13.7	6.6	12.4
25–29	5.5	9.3	8.2	10.8	6.9	10.0
30–34	5.5	7.5	7.1	8.0	6.3	7.8
35–39	4.3	4.8	4.8	4.1	4.5	4.6
40–44	5.0	4.6	4.2	4.0	4.6	4.3
45–49	3.4	2.4	2.7	1.8	3.1	2.1
50–54	3.6	2.4	3.7	1.9	3.7	2.7
55–59	1.4	1.0	1.1	0.7	1.3	0.8
60–64	2.6	1.6	1.3	1.2	2.0	1.4
65–69	1.0	0.6	0.5	0.4	0.7	0.5
70–74	1.1	0.6	0.1	0.5	0.7	0.6
75–79	0.3	0.3	0.2	0.2	0.2	0.2
80–84	0.5	0.4	0.1	0.3	0.4	0.3
85+	1.0	0.5	0.3	0.4	0.7	0.4
All	100.0	100.0	100.0	100.0	100.0	100.0

the age accuracy index. The age ratio is simply the ratio of the population in a given age group to one third of the sum of the population in that age group and the preceding and following age groups. Ideally the age ratio should approximate 100 except in cases of extreme fluctuation in demographic processes in the past or historical events that would distort the age composition of the population. The ratio serves as a measure of net age misreporting. The average of the absolute deviations of the age ratios (from 100) for all age groups yields the

Table 3 Age ratios and age accuracy index for the population enumerated in the NFS household survey and 1963 census, by sex

Age group	Age ratios			
	NFS		1963 Census	
	Male	Female	Male	Female
0–4	–	–	–	–
5–9	111.0	104.6	106.1	105.5
10–14	95.3	94.3	96.5	84.6
15–19	95.0	98.0	132.7	89.6
20–24	82.4	91.9	114.5	119.2
25–29	102.3	106.2	99.3	99.4
30–34	107.3	105.9	104.5	105.2
35–39	87.1	89.3	84.5	76.7
40–44	118.5	107.7	117.8	121.3
45–49	84.6	76.9	76.6	68.7
50–54	129.5	146.6	124.7	132.9
55–59	54.2	55.2	59.1	52.7
60–64	158.6	136.2	151.4	159.7
65–69	63.4	65.2	61.3	57.3
70–74	–	–	–	–
Sum of absolute deviations from 100	265.2	235.8	274.4	314.8
Age accuracy index[a]	20.4	18.1	21.1	24.2

[a]Calculated as the average of the absolute deviations of the ratios from 100.

age accuracy index which is a measure of the overall accuracy of the age data. The lower the age accuracy index, the more adequate are the age data. These ratios and indices for the total enumerated population by sex, are given in table 3. Most age ratios, especially at older ages, show very substantial deviations from 100. The age accuracy indices of 20.4 and 18.1 for males and females further confirm the greater misreporting of age for males than females. The average of the two values (19.3) is the measure of the overall accuracy of the reported age data. Comparison with age ratios from the 1963 census, also shown in table 2, shows a slight improvement in the NFS household data on male age, and a more substantial improvement for female age reporting. On the other hand, the curve of the age distribution of the population in 5-year age groups is smoother for the NFS than for the 1963 census. The huge concentration of persons in the 20–24 and 25–29 age groups in the census is also evident.

Age misreporting can also be assessed by examining the sex ratios for each age group. Theoretically, and in the absence of historical upheavals or drastic changes in demographic processes in the past, the population of a country should have a slight excess of males at the younger age groups, about equal numbers in the middle age groups and an excess of females in the older ages. The sex ratios for the rural, urban and total populations enumerated in the household survey and in the 1963 census are presented in table 4; the ratios for both total populations are illustrated in figure 6. It is observed that the ratios for the household survey show an almost equal number of males and females at age 0–4, an excess of males at 5–9, then a very considerable deficit of males in the 15–39 age range which there is an excess of males for all age groups (except 50–54). A similar pattern is observed for the 1963 census except that the deficit of males is less, restricted to the 15–34 age range and excess of males over 40 higher. That the underenumeration of males in these age ranges is greater in the survey than in the census is typical of many countries which have carried out WFS surveys (see Rutstein 1984: 14) and can be explained by the fact that the household survey did not cover institutional households (educational, military and penal institutions) in most of which men of those ages are found, or that interviewers might have regarded the listing of males not as important as the listing of females since the primary objective of the household schedule was to locate women eligible for individual interview. The substantial excess in the number of males at ages over 40 could be attributed to the tendency to report higher age for adult males. The low sex ratio for ages 50–54 is a common phenomenon in fertility surveys where age 49 is the upper limit and could have arisen if the interviewers deliberately recorded higher ages for some women 45–49 in order to exclude them from the individual survey and thus minimize their workload. That this actually happened (although to a lesser extent), is seen in table 2 where the proportion of women aged 50–54 is 3.7 per cent compared with 2.7 per cent for women aged 45–49.

3.3 AGE IN THE INDIVIDUAL QUESTIONNAIRE

As stated earlier, information on the age of most household members was probably provided by proxy respondents who were not likely to have the correct information for all members. In the individual questionnaire, however, information on age was provided directly by the respondent or she provided the information from which her age was estimated. If the age had to be estimated by the interviewers, provision was made in the questionnaire for the interviewer to state the criteria

Table 4 Sex ratios of the *de facto* population enumerated in the household survey and in the 1963 census, by rural-urban status

Age group	Rural		Urban		Total	
	1963 Census	NFS	1963 Census	NFS	1963 Census	NFS
0–4	97.0	99.9	99.1	102.8	97.3	100.4
5–9	108.0	111.3	100.2	109.3	106.9	111.0
10–14	121.6	103.8	119.7	94.0	121.3	102.3
15–19	86.3	84.2	118.0	78.4	91.0	83.3
20–24	77.0	66.2	118.6	56.9	83.7	64.5
25–29	82.0	64.3	116.5	73.1	87.9	65.7
30–34	91.3	72.2	117.2	93.3	95.3	75.2
35–39	113.4	85.6	139.2	99.2	117.8	87.6
40–44	116.8	113.7	129.7	135.1	118.8	117.0
45–49	139.4	116.3	146.0	178.2	140.5	123.0
50–54	128.5	93.9	123.6	126.8	127.7	97.1
55–59	150.9	111.1	138.9	201.2	148.9	120.2
60–64	130.4	201.7	118.1	129.4	132.0	192.8
65–69	150.2	205.3	121.8	120.4	145.6	193.2
70+	144.7	269.8	115.9	234.0	140.8	262.2
All	99.7	97.3	114.9	98.5	102.0	97.5

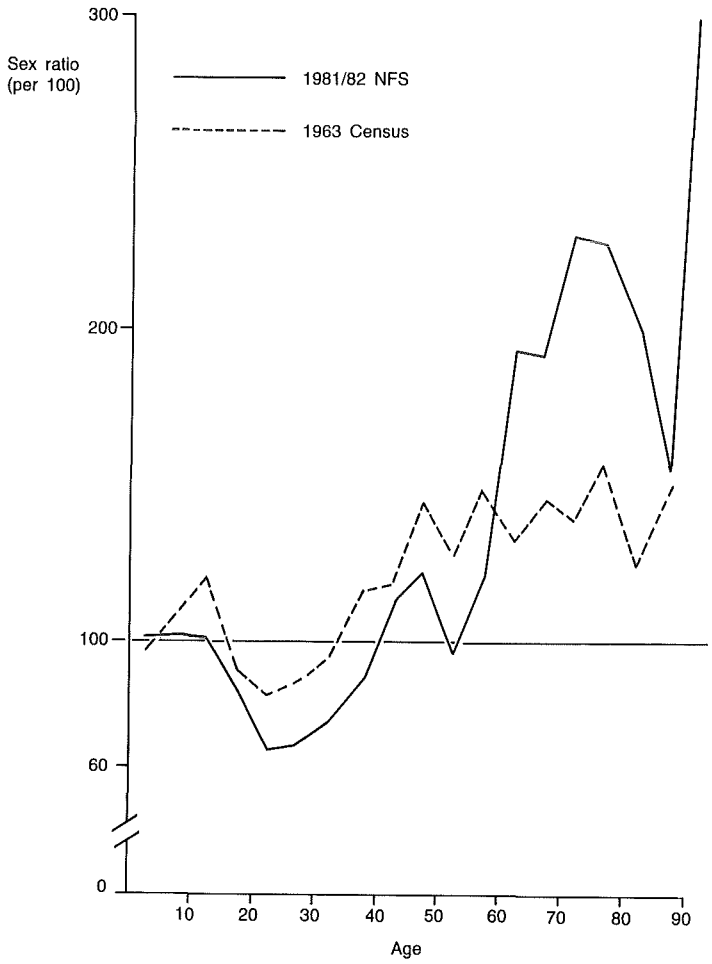

Figure 6 Sex ratios for the *de facto* population enumerated in the household survey and the 1963 census

used in the estimation. In all instances, age as given or estimated was to be entered in the date event chart. It might thus be illuminating to examine the quality of age data as contained in the individual questionnaire and later to ascertain the consistency of age data from both the household schedule and individual questionnaire.

The format in which the information on the date of birth of respondents was obtained is likely to affect the data on age. Three formats are possible from the questionnaire design: exact date of birth in calendar month and year, year of birth only and years ago of birth. Age obtained from the exact date is likely to be free from error; if any errors do occur, they will probably arise from the preference to report the year of birth in rounded numbers. This is also likely to be the case if only the year of birth is given; here the month of birth would be randomly imputed by the program to obtain age in completed years. A problem arises here when the respondents first think of their age (often in rounded numbers) and subtract this from the date of interview to obtain calendar year of birth (with or without the help of the interviewer). When this happens, there will be the tendency for year of birth to be given in rounded durations from the date of interview. Information on date of birth based on years ago is likely to result in rounded ages; the respondents are likely to be those for whom age had to be estimated or who estimated their own ages themselves. Such ages are likely to be in rounded numbers and obtained dates of birth will tend again to be concentrated on rounded durations from interview date. Exact date of birth was obtained for only 16.4 per cent of the respondents; 26.5 per cent gave information on only the calendar year of their birth while for a majority (57.1 per cent), years ago of birth was obtained. Thus the age of over three-quarters (83.6 per cent) of the interviewed women is likely to have been estimated either by the respondents themselves or by the interviewers. The proportion for whom age had to be estimated increases with reported age as evidenced by the fact that the exact date of birth was available for only 5.2 per cent of respondents aged 45–49 compared with 28.1 per cent for those aged 15–19 (table 5).

The distribution of the respondents by single years of age is shown in figure 7. It is observed that the heaping at ages ending in 0 is particularly acute, especially at age 30. Heaping also occurs at ages ending in 5 but to a lesser extent. The value of Myer's summary index for the total sample is 25.5, indicating that age for at least a quarter has been reported in an incorrect final digit. The same distribution in single years but by format of the respondents' date of birth is shown in figure 8. More severe heaping, and in digits ending in 0 and 5, are observed for all formats, but particularly when years ago of birth was given. The figure also shows substantial heaping even

Table 5 Percentage distribution of respondents to the individual questionnaire according to format of date of birth, by current age

Current age	Format of date of birth				Number of women
	Exact date	Calendar year only	Years ago	Total	
15–19	28.1	27.3	44.6	100.0	1201
20–24	21.0	27.5	51.5	100.0	1710
25–29	15.7	25.3	59.0	100.0	1766
30–34	10.1	27.7	62.3	100.0	1547
35–39	9.9	27.7	63.4	100.0	1110
40–44	7.7	25.9	66.4	100.0	904
45–49	5.2	22.3	72.5	100.0	591
All	16.4	26.5	57.1	100.0	9729

when the exact date of birth has been given. Myers' summary indices for the formats are 15.8, 19.5 and 34.6 when exact date, calendar year only and years ago of birth respectively were given. Thus, though the extent of heaping is least when the exact date was obtained, that format is still not free from errors.

The extent of age misreporting also varies considerably by the socio-economic characteristics of the respondents as shown in table 6 and figure 9. It is slightly better in urban than in rural areas, and virtually the same whether the respondents co-operated well with the interviewers or not. It is worse for respondents in the north-west and north-east regions (particularly from age 20 onwards) compared with respondents in the south-west and south-east (Myers' indices: 36.8, 30.0, 23.7 and 21.0 respectively). The largest variations are observed for level of education. Heaping is least pronounced for respondents with secondary or higher education and extreme for respondents with no formal education or with only Koranic education (Myers' indices: 15.0, 31.0 and 36.6 respectively). Respondents with complete or incomplete primary education occupy intermediate categories. Consequently, heaping is less among the women who can read (ie literate) than those who cannot (Myers' indices: 17.0 and 31.1, respectively).

Since most demographic data by age are presented for

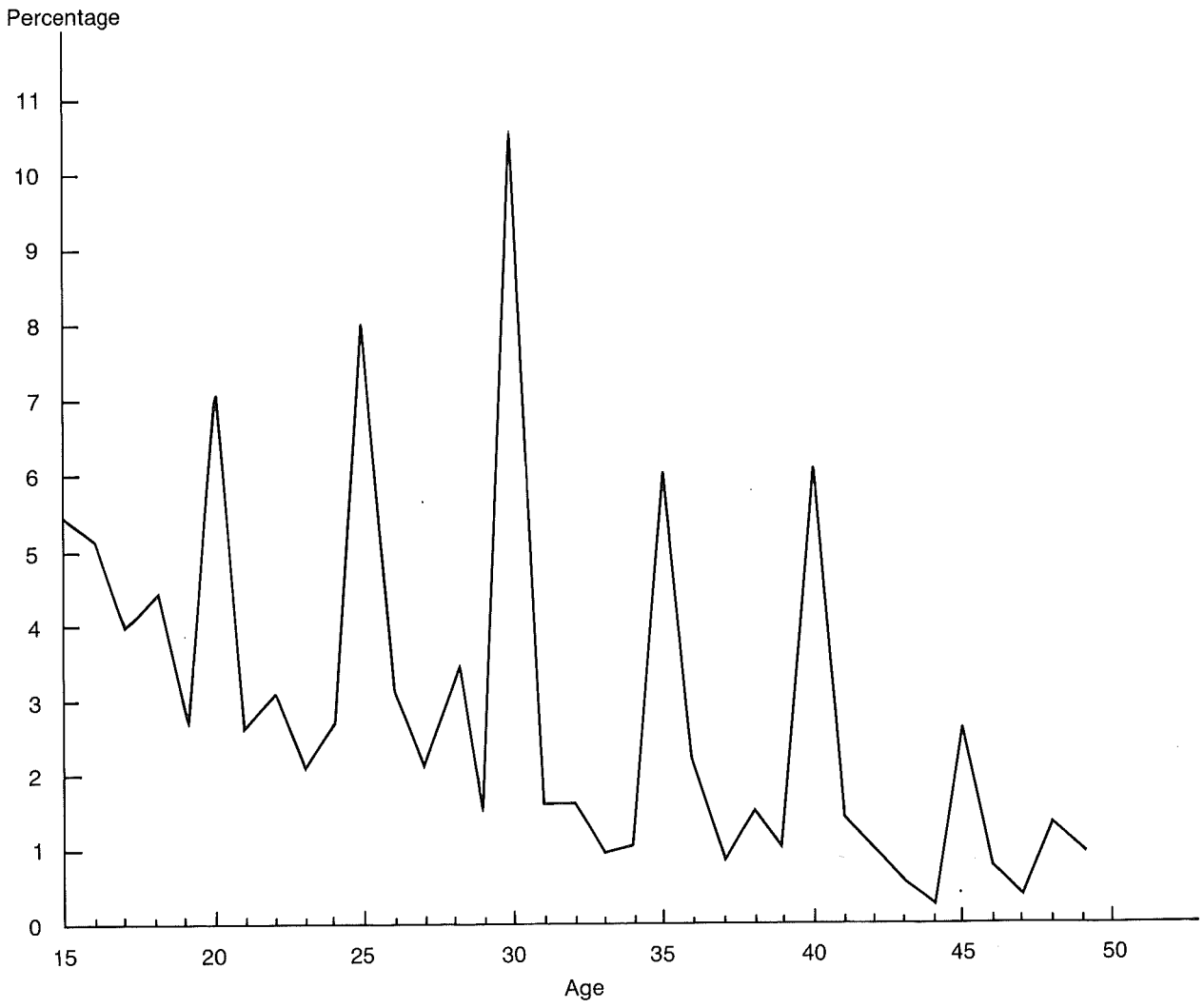

Figure 7 Percentage distribution of surveyed women, by age in completed years

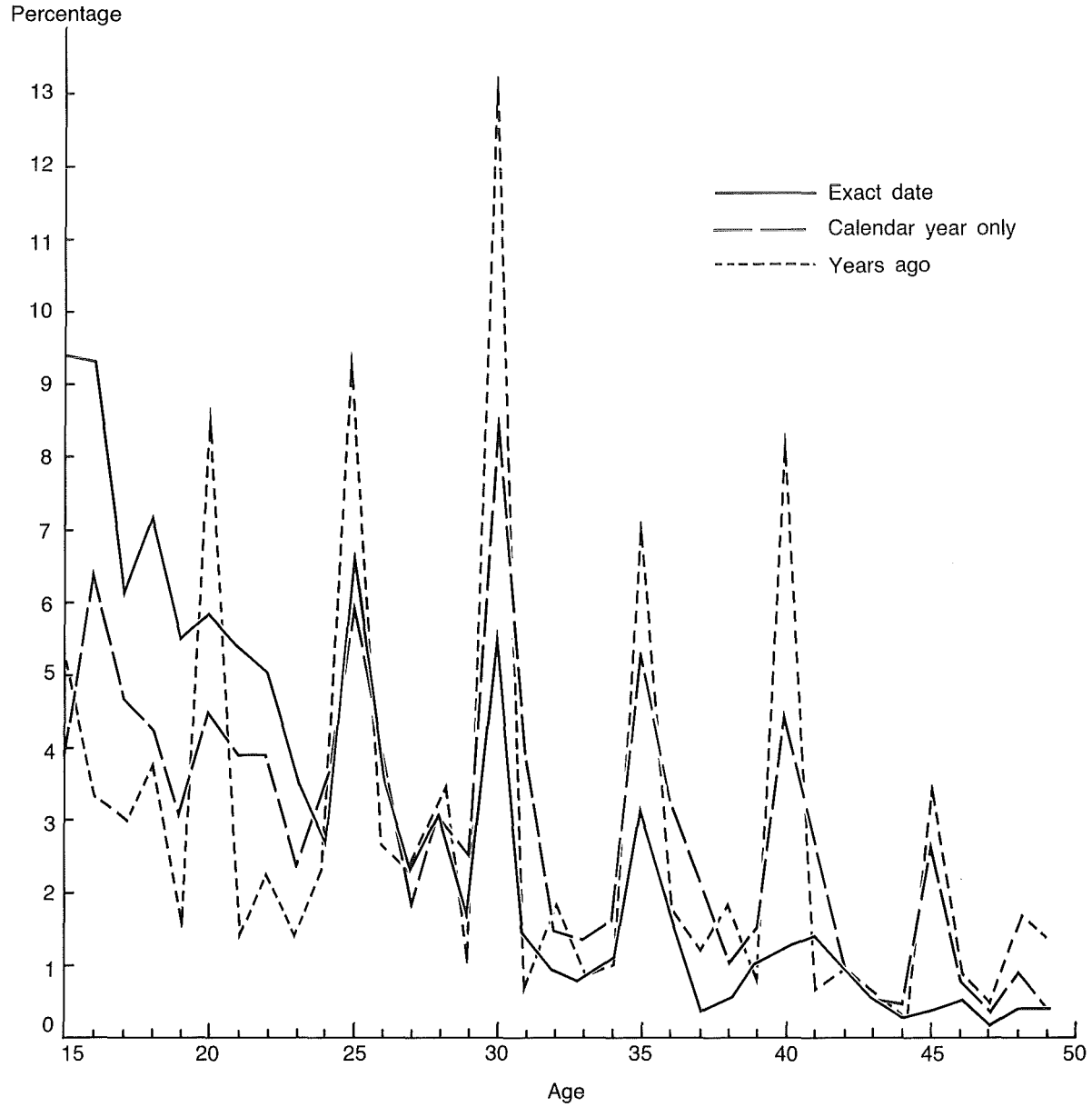

Figure 8 Percentage distribution of the surveyed women, by single years of age and format of respondent's date of birth

5-year age groups, the bias introduced by heaping at particular ages is severe only when it results in the displacement of respondents from one age group to another. Thus, it is necessary also to examine the age distribution of the respondents in 5-year age groups as shown in table 7. In a developing country like Nigeria, the proportions of the age groups should be expected to decrease with increasing age, if age had been correctly reported. In this respect, the obtained distribution is anomalous in that the proportion of the women aged 25–29 is higher than for those aged 20–24 (18.1 and 17.6 per cent respectively). The age ratios also show that there are excesses of respondents in the 25–29, 30–34 and 40–44 age groups at the expense of the other age groups. This may be primarily due to the excessive heaping of respondents at ages 25, 30 and 40 as observed earlier. The sum of the absolute deviations of the age ratios from 100 for the 20–24 to 40–44 age groups is 29.0, yielding an age accuracy index of 5.8. This value is not too bad and is indicative that, though some gross distortion exists in the data, the extent is acceptable if not perfect. Figure 10 shows that the extent of distortion again varies markedly by some background characteristics of the respondents. It is relatively better for respondents in urban than rural areas, for those living in the south-east are literate, have primary or higher education and gave their exact date of birth. It is particularly bad for illiterate women or those that have only Koranic or no education and those that live in the north-east or north-west.

Table 6 Myers' summary index of age preference by selected background characteristics of the survey sample

Background characteristics	Summary index[a]
A Type of place of residence	
Rural	26.6
Urban	21.8
B Type of place of childhood residence	
Rural	25.8
Urban	24.8
C Region of residence	
North-east	30.0
North-west	36.8
South-east	21.0
South-west	23.7
D Level of education	
None	31.0
Koranic	36.6
Primary incomplete	23.0
Primary completed	18.0
Secondary and above	15.0
E Literacy	
Can read	17.0
Cannot read	31.1
F Format of respondents birth date	
Exact date given	16.8
Calendar years only	19.5
Years ago	34.6
G Degree of co-operation	
Poor or fair	26.3
Good or very good	25.7
H All	25.5

[a]Calculated as half of the absolute deviations.

Table 7 Percentage distribution of respondents to the individual questionnaire in five-year age groups

Age group	Number	Per cent	Age ratio
15–19	2101	21.6	–
20–24	1710	17.6	92.0
25–29	1766	18.1	105.5
30–34	1547	15.9	104.9
35–39	1110	11.4	93.5
40–44	904	9.3	104.1
45–49	591	6.1	–
All	9729	100.0	29.0

3.4 CONSISTENCY IN AGE REPORTING

Information on the age of survey respondents from both the household and individual surveys can be matched to determine the extent of consistency of age data from both sources. This may be used as a further internal check on the quality of the data. The distribution of all respondents to the individual questionnaire according to difference in reported ages (in years) from both sources is presented in table 8. Age is the same in both sources for 83.7 per cent of the women. Age is higher for 9.8 per cent and lower for 6.5 per cent in the individual than in the household survey. Most of the age difference is one year higher (6.5 per cent) or lower (4.0 per cent); age differences exceed one year for only 5.8 per cent of the respondents. The difference in both directions is very similar for the younger women (15–24); for women 25 and over (except the oldest age group), proportionately

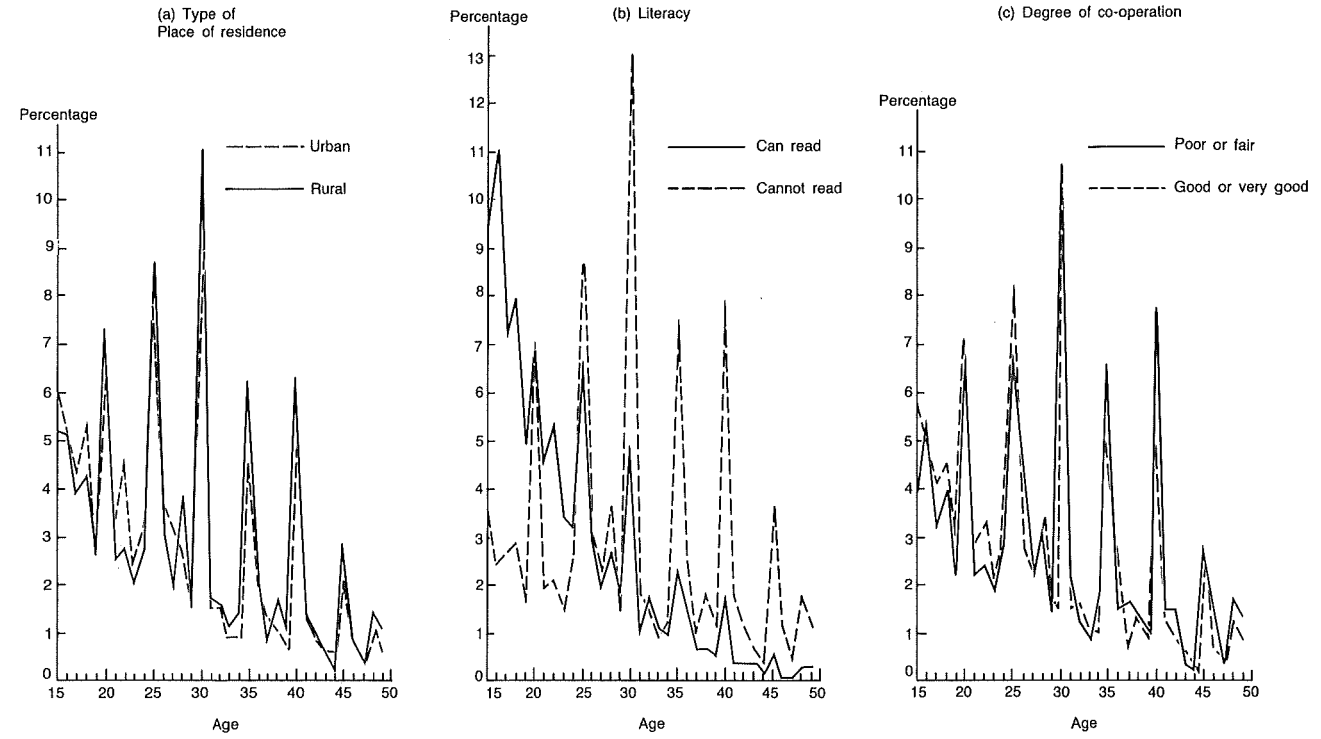

Figure 9 Percentage distribution of the surveyed women, by single years of age and selected characteristics (*continued next page*)

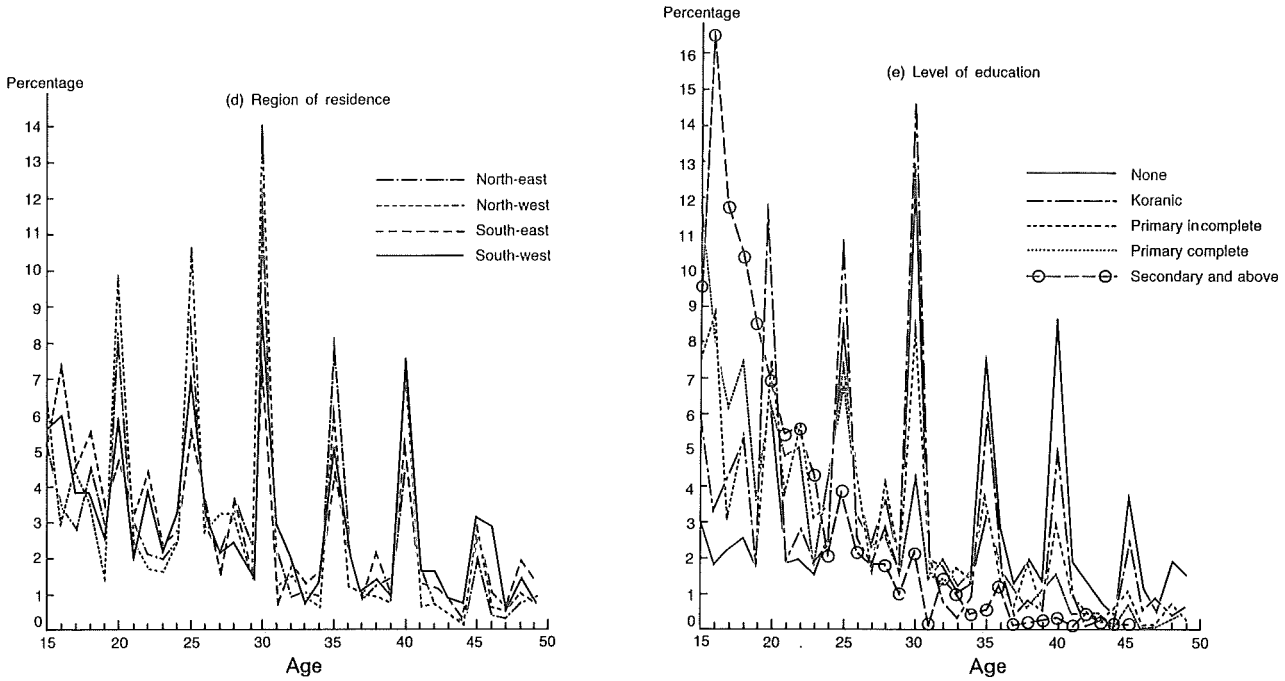

Figure 9 (*continued from previous page*) Percentage distribution of the surveyed women, by single years of age and selected characteristics

more report higher ages in the individual than was reported for them in the household survey.

The extent to which the difference in reported ages from both sources has resulted in respondents being classified under different age groups is shown in tables 9 and 10. Despite the observed differences, 94.0 per cent are in the same age group in both sources; 3.0 per cent each have been transferred to either higher or lower age groups. There is no consistent pattern of transference by age. There are, however, some minor variations by background characteristics as is shown in table 11. Literacy, higher education and to a lesser extent residence in the south-east or south-west regions and in urban areas are associated with greater consistency in age from both sources. Consistency, however, does not indicate that data from either or both sources are necessarily correct.

3.5 SUMMARY

Age reporting in the NFS was found to be subject to a substantially high degree of heaping because of the preference of some terminal digits, particularly 0 and 5 and avoidance of others, particularly 9 and 1. The Myers' index shows that age with an incorrect final digit was reported for at least a third of the population enumerated in the household survey and a quarter of respondents to the individual survey. The age accuracy index for the population enumerated in the household survey is 19.3. Heaping was found to be higher for males than females and in rural than in urban areas.

Only about a sixth of the women could give their exact dates of birth. Heaping was considerably higher among women for whom information on birth date was obtained in 'years ago', for those with no formal or only Koranic education and for those not literate. Some distortion was also observed in the distribution of the respondents in 5-year groups. The extent of such distortions was, however, not too bad as shown by an age accuracy index of 5.8. Data on age from the household and individual surveys were relatively consistent. Age is the same in both sources for 83.7 per cent of the respondents to the individual questionnaire and as many as 94.0 per cent of them are in the same quinquennial age group in both sources. Observed differences were more in the form of age or age group being higher in the individual than in the household survey. This consistency does not, however, mean that age data in either or both sources are necessarily correct.

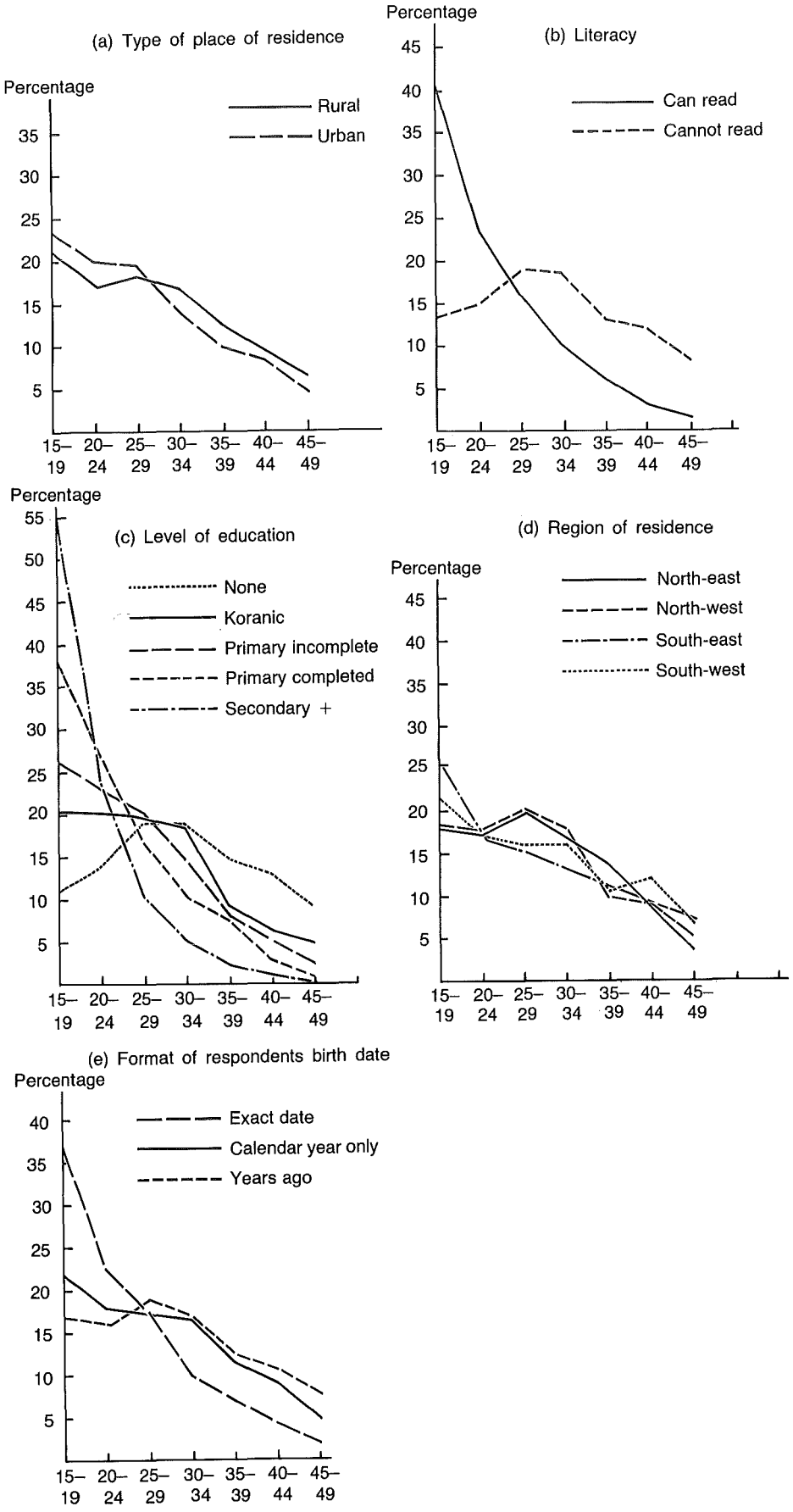

Figure 10 Percentage distribution of the surveyed women in five-year age groups, by selected characteristics

Table 8 Percentage distribution of respondents to the individual questionnaire according to difference in reported ages (in years) between the household and individual surveys

Age difference in years[a]	Current age								Number of cases
	15–19	20–24	25–29	30–34	35–39	40–44	45–49	Total	
−3 or more	0.2	1.9	2.6	3.0	7.2	2.5	5.5	2.7	258
−2	0.1	0.6	0.9	0.7	0.8	0.8	0.2	0.6	57
−1	5.5	6.8	7.2	7.6	6.7	7.6	2.8	6.5	635
0 (the same)	86.7	82.5	82.6	83.1	79.8	84.2	88.0	83.7	8145
+1	5.2	5.6	4.0	3.0	3.4	2.2	2.1	4.0	392
+2	0.7	0.5	0.3	0.7	0.4	0.7	1.4	0.6	57
+3 or more	1.6	2.7	2.3	1.8	1.7	2.0	0.0	1.9	187
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

[a](−) Age is higher in individual survey.
(+) Age is higher in household survey.

Table 9 Percentage distribution of respondents to the individual questionnaire according to difference in reported age groups between the household and individual surveys

Age group in household survey	Age group in individual survey								Number of cases
	15–19	20–24	25–29	30–34	35–39	40–44	45–49	Total	
15–19	97.5	1.7	0.5	0.1	0.1	0.1	0.1	100.0	2091
20–24	3.1	92.8	2.7	1.1	0.0	0.2	0.0	100.0	1712
25–29	0.5	4.1	91.9	1.9	1.5	0.1	0.0	100.0	1799
30–34	0.1	0.4	2.4	93.1	3.4	0.5	0.1	100.0	1559
35–39	0.0	0.2	1.2	2.9	94.4	1.0	0.3	100.0	1058
40–44	0.0	0.5	0.4	1.0	2.5	93.1	2.6	100.0	918
45–49	0.0	0.0	0.4	0.0	0.9	4.0	94.7	100.0	591
Number of cases	2101	1710	1766	1547	1110	904	591		9729

Table 10 Percentage distribution of respondents to the individual questionnaire according to differences in age group recorded in the individual and household surveys

Difference in age group[a]	Age groups in individual survey								Number of cases
	15–19	20–24	25–29	30–34	35–39	40–44	45–49	Total	
−3 or more	0.0	0.0	0.0	0.1	0.2	0.8	0.8	0.2	15
−2	0.0	0.0	0.6	1.3	2.4	0.9	0.6	0.7	69
−1	0.0	2.1	2.6	2.2	4.8	1.2	4.0	2.1	204
0 (the same)	97.0	92.9	93.6	93.8	90.0	94.5	94.7	94.0	9146
+1	2.5	4.3	2.1	2.0	2.0	2.6	0.0	2.5	241
+2	0.4	0.3	0.7	0.6	0.5	0.0	0.0	0.4	41
+3	0.1	0.4	0.3	0.0	0.0	0.0	0.0	0.1	14
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	9729

[a](−) Age group is higher in individual survey.
(+) Age group is higher in household survey.

Table 11 Percentage distribution of respondents to the individual questionnaire according to differences in age groups recorded in the individual and household surveys, by selected background characteristics

Background characteristics	Difference in age group[a]								Number of cases
	−3+	−2	−1	0	+1	+2	+3+	Total	
A Childhood place of residence									
Rural	0.1	0.9	2.1	93.6	2.7	0.5	0.2	100.0	7279
Urban	0.2	0.2	2.2	95.3	1.7	0.4	0.0	100.0	2450
B Type of place of residence									
Rural	0.2	0.8	2.1	93.7	2.6	0.4	0.2	100.0	7528
Urban	0.1	0.2	2.1	95.1	2.0	0.4	0.0	100.0	2202
C Region of residence									
North-east	0.0	1.0	3.4	90.7	3.7	0.7	0.4	100.0	2343
North-west	0.3	1.2	2.1	93.5	2.1	0.7	0.2	100.0	2221
South-east	0.1	0.3	1.2	96.7	1.5	0.1	0.0	100.0	3139
South-west	0.2	0.4	2.0	94.1	2.9	0.3	0.0	100.0	2027
D Level of education									
None	0.1	0.9	2.5	92.9	2.9	0.4	0.2	100.0	5728
Koranic	0.5	0.7	2.4	93.9	1.2	1.2	0.3	100.0	872
Primary incomplete	0.0	0.5	1.8	95.1	2.3	0.3	0.0	100.0	1095
Primary complete	0.1	0.1	1.2	96.4	2.1	0.1	0.0	100.0	965
Secondary and above	0.2	0.3	0.8	96.9	1.6	0.2	0.0	100.0	1106
E Literacy									
Can read	0.1	0.3	1.2	96.4	1.9	0.2	0.0	100.0	2903
Cannot read	0.2	0.9	2.5	93.0	2.7	0.5	0.2	100.0	6826

[a](−) Age higher in individual survey.
(+) Age higher in household survey.

4 Nuptiality

4.1 INTRODUCTION

In most human societies, reproduction of the population takes place within relatively stable sexual unions, often referred to as marriages. Consequently, the nuptiality patterns in any society are important factors affecting its level of fertility. Reliable and accurate assessment of the effects of nuptiality, however, depends on the quality of data collected. Errors in nuptiality data from sample surveys are mostly in the form of under-coverage of women in marital unions, the determination of current marital status, misreporting of dates of start and end of marriages, omission of marriages and errors from other sources (eg reporting of the age of respondents). The objective of this section is to examinine the quality of nuptiality data collected in the NFS with respect to the above possible errors.

The NFS collected information on three aspects of nuptiality namely: current marital status, retrospective marital history and the presence of co-wives. Each of these is prone to particular types of error which affect the analysis of the data. Marriage was defined to include all legally, religiously or traditionally contracted unions as well as other stable cohabitations. Since what constitutes 'stable cohabitations' can be differentially interpreted, some women may be wrongly included or excluded in the category of married women. Then there is the common problem of formerly married women reporting themselves as single or never married which would lead to their exclusion from the ever-married category. Marital histories, just like other retrospective data collected in surveys, are particularly prone to misreporting dates of start or end of marriages (particularly the first marriage) and omission of marriages, especially stable co-habitations that precede formal marriages. In the latter instance, (older) women may under-report previous short but relatively stable unions before their first formalized union. Misreporting of dates of start or end of unions, or of age at first or subsequent marriages is prevalent in societies with a low level of literacy and manifests itself in the heaping of respondents at particular dates or ages at marriage. Some respondents, especially older ones, may also give the date (or age) of subsequent marriages as the date (or age) of first marriage. Such misreporting seriously affects the obtained mean age at first marriage, durations of exposure to risk of childbearing within marriage; intervals between marriage and births and marital fertility rates based on woman-years of exposure.

The incidence of these potential errors in the nuptiality data collected in the NFS are examined along the following lines: heaping, age at first marriage and number of times married. Such an examination will invariably be confined to internal comparisons of nuptiality data as collected in the individual questionnaire since information on nuptiality was not collected in the household survey, in the 1963 census or in any other nationwide survey in the country.

4.2 HEAPING IN NUPTIALITY DATA

For every respondent who has ever been married, information on the date of her first and subsequent marriages (if any) was collected in the NFS. If her date of marriage was not known, then her age at the time of the marriage was either obtained from her or estimated; her date of marriage was subsequently calculated from this age. The duration of marriage (ie years since first marriage) of each respondent was then calculated from the above information, in conjunction with the information on the date of interview.

Thus the nuptiality variables in which heaping may occur are the date (year) of marriage, the years since first marriage and age at first marriage. An important source of error that might lead to heaping on the above variables is the format of reporting of date of first marriage. When the exact date (in calendar month and year) is reported, then heaping should be minimal and, if it occurs, would be on rounded calendar years. When the date is, however, estimated or obtained from reported age at first marriage, then greater heaping may be expected and the heaping would be on digits of year of first marriage and years since first marriage corresponding with rounded durations from the survey date. The combination of these two probable heaping patterns would be likely to produce a pattern of heaping on year of first marriage or years since first marriage that may be on both the conventionally preferred digits and on those that are avoided (if the latter correspond with preferred durations from the survey date).

The exact date of first marriage in calendar month and year was obtained for only 18.1 per cent of the ever-married women; the date in calendar year only was obtained for 45.0 per cent and for 36.9 per cent, the date was calculated from reported age at first marriage. That the pattern of heaping on year of first marriage, years since first marriage and age at first marriage varies considerably by the format of date of first marriage is shown in figures 11, 12 and 13 respectively. Figure 11 shows that when the exact date of first marriage was reported, the heaping in year of first marriage occurs mostly in rounded numbers whether the interviews were conducted in 1981 or 1982; the exceptions are 1965 and 1970 where the heaping occurs one year later, corresponding to preferred durations of 15 and 10 years ago respectively from the survey date. If, on the other hand, only the calendar year of first marriage was reported, the pattern is more confused since heaping occurs at some

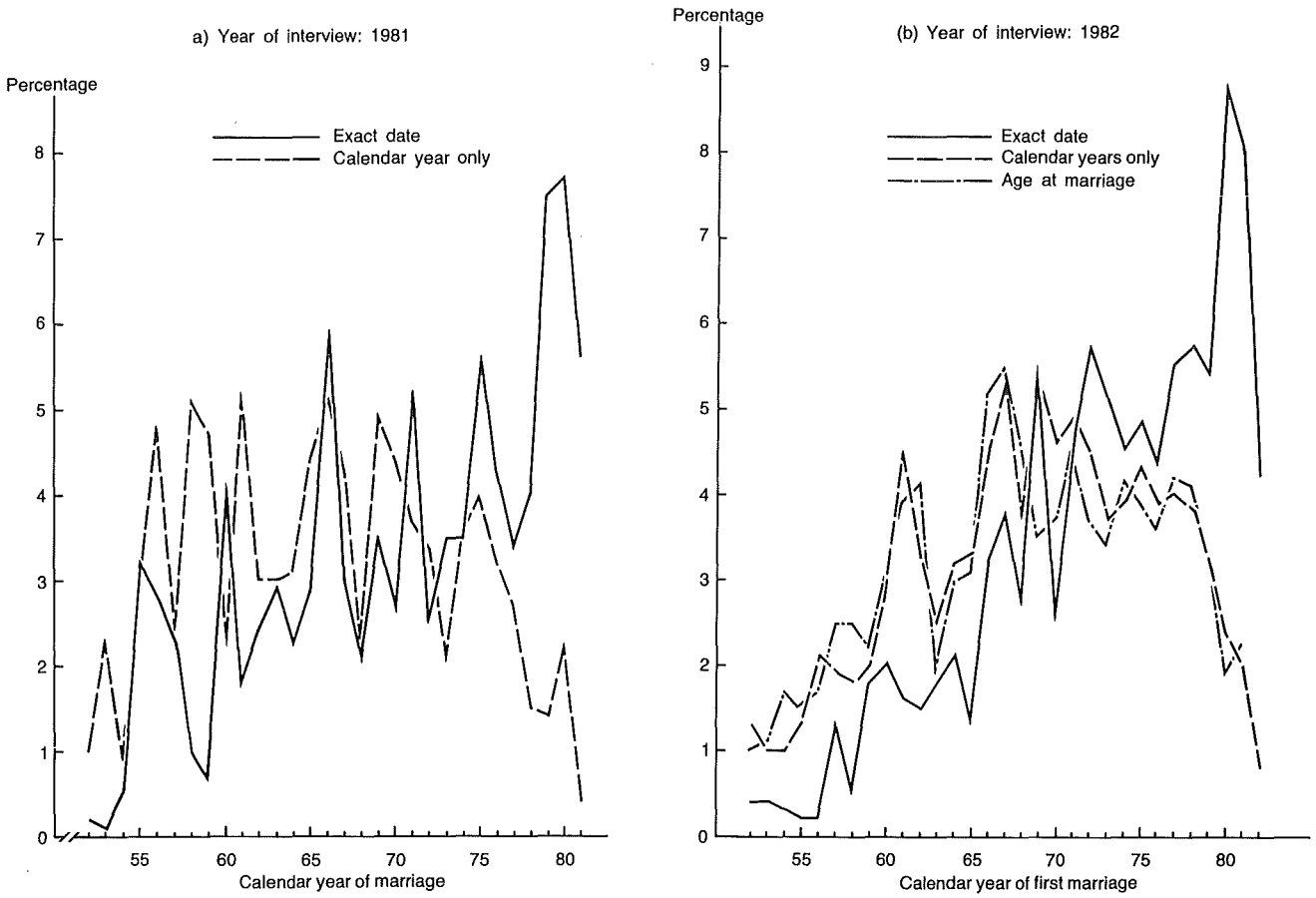

Figure 11 Percentage distribution of ever-married women according to year of first marriage and year of interview, by format

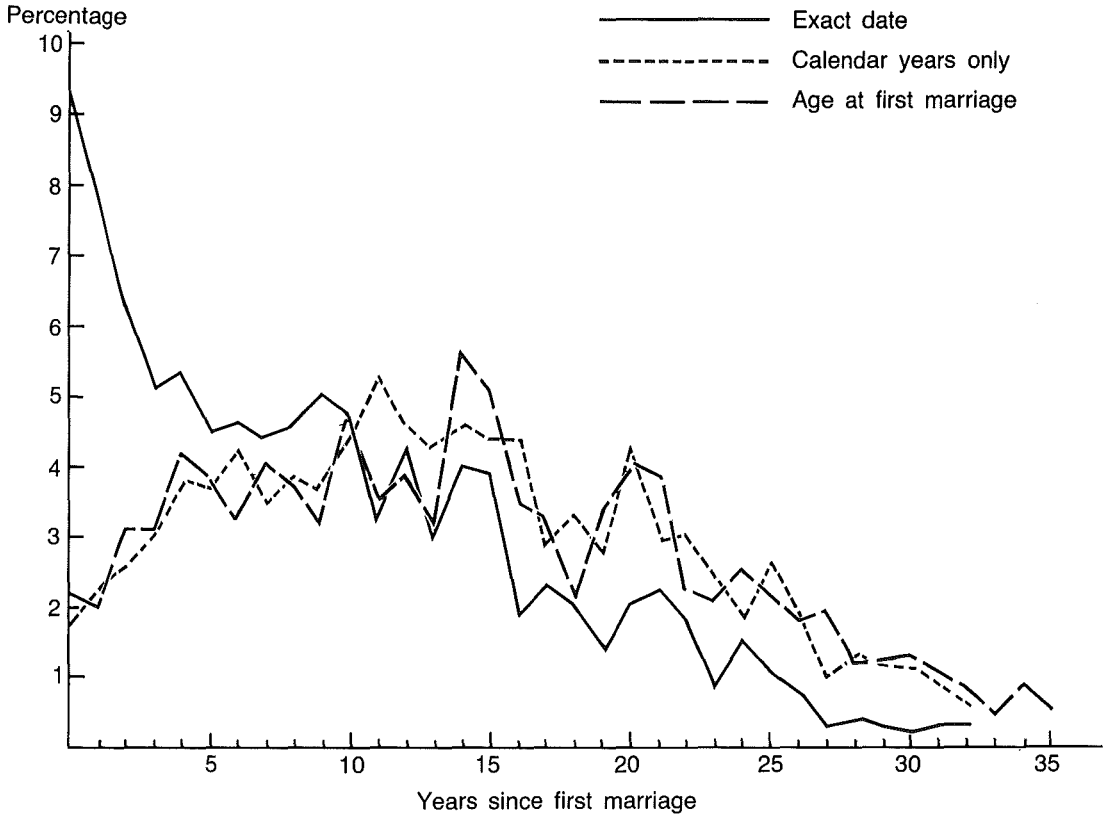

Figure 12 Percentage distribution of ever-married women according to years since first marriage, by format of date of first marriage

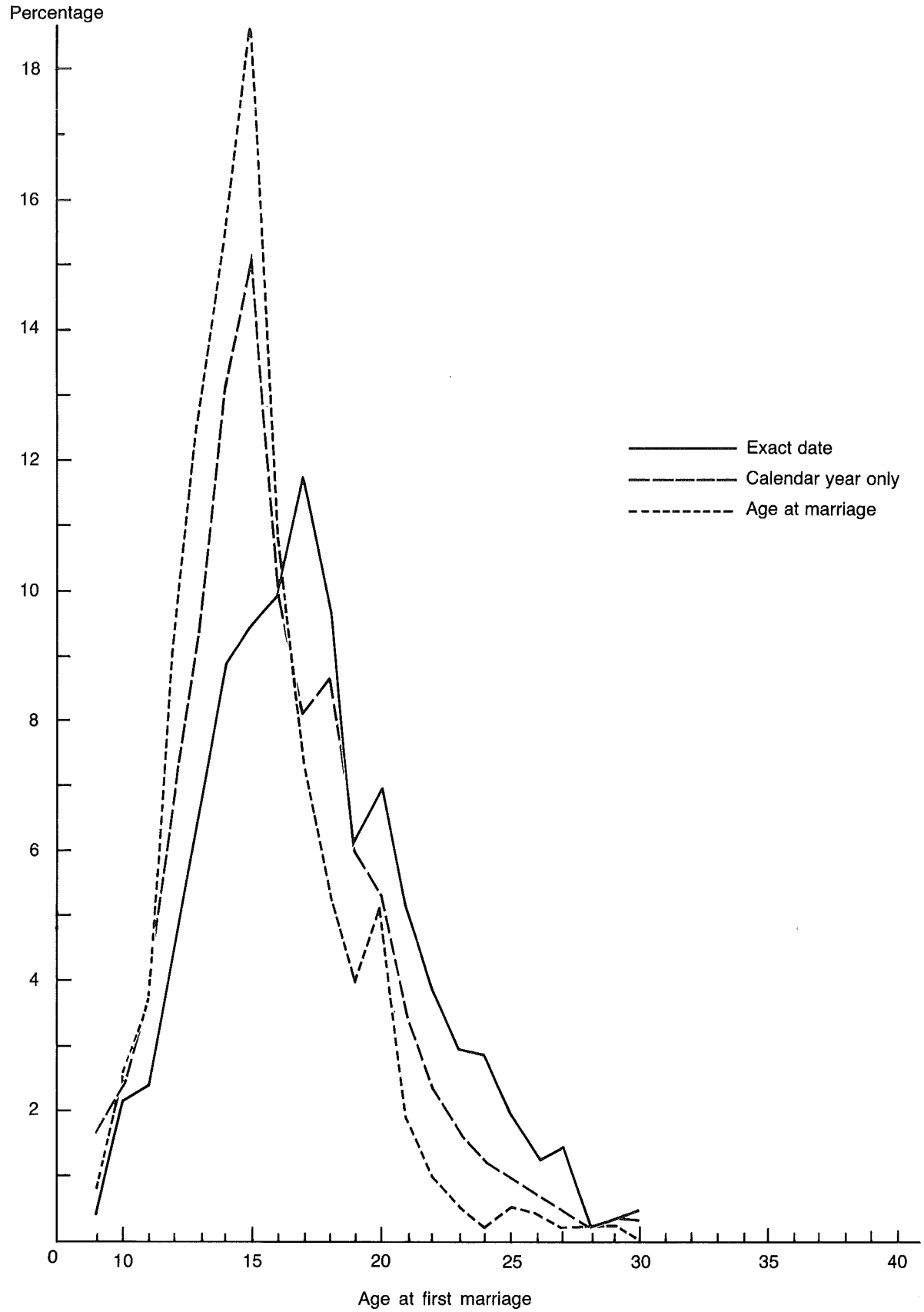

Figure 13 Percentage distribution of ever-married women according to age at first marriage, by format of date of first marriage

rounded years, some rounded durations from survey date and on digits corresponding to preferred ages at first marriage. A similar pattern is also observed when date of first marriage was calculated from age at first marriage. It would appear that in many instances where calendar year of first marriage was reported, the year was not exactly known but arrived at by the respondent (or interviewer) after calculating from her age at first marriage.

Years since first marriage was obtained by subtracting the date of first marriage from the date of interview. It should, thus, be very much affected by any heaping in date of first marriage. That it also varies by the format of that date is shown in figure 12 though no systematic pattern is observed partly because of its dependence on the reported date of first marriage and the huge concentration of marriages around age 15. The figure, however, shows that when date of first marriage was reported in calendar year only, heaping at marital durations under 20 years occurs at digits 1 and 6 corresponding to years of marriage ending in 0 or 5. Despite the usual concentration of marriages within a narrow age range, the distribution shown in figure 13 is most definitely heaped at the preferred ages around 15 years and also varies by the format of reporting of date at first marriage.

The effect of these variations is that the distributions of the ever-married women by year of first marriage, years since first marriage and age at first marriage either as a whole or by various socio-economic characteristics shows a pattern of heaping that is irregular and occurs at both preferred and avoided digits, the predominant pattern depending on the relative distribution of women reporting their date of first marriage in a particular format. Thus the distributions by year of first marriage show peaks that variously correspond with rounded years, rounded durations from survey date or preferred age at first marriage (figure 14). The same is also observed for years since first marriage and age at first marriage for all ever-married women and by residential and educational characteristics (figures 15(a) to 15(e) and 16(a) to 16(c) respectively). There is substantial heaping in NFS nuptiality data on those variables. Another important observation is the decline in the proportion of ever-married women since the early 1970s in figure 14, which is reflected in the lower proportions of ever-married women of less than ten years since first marriage in figure 15(a).

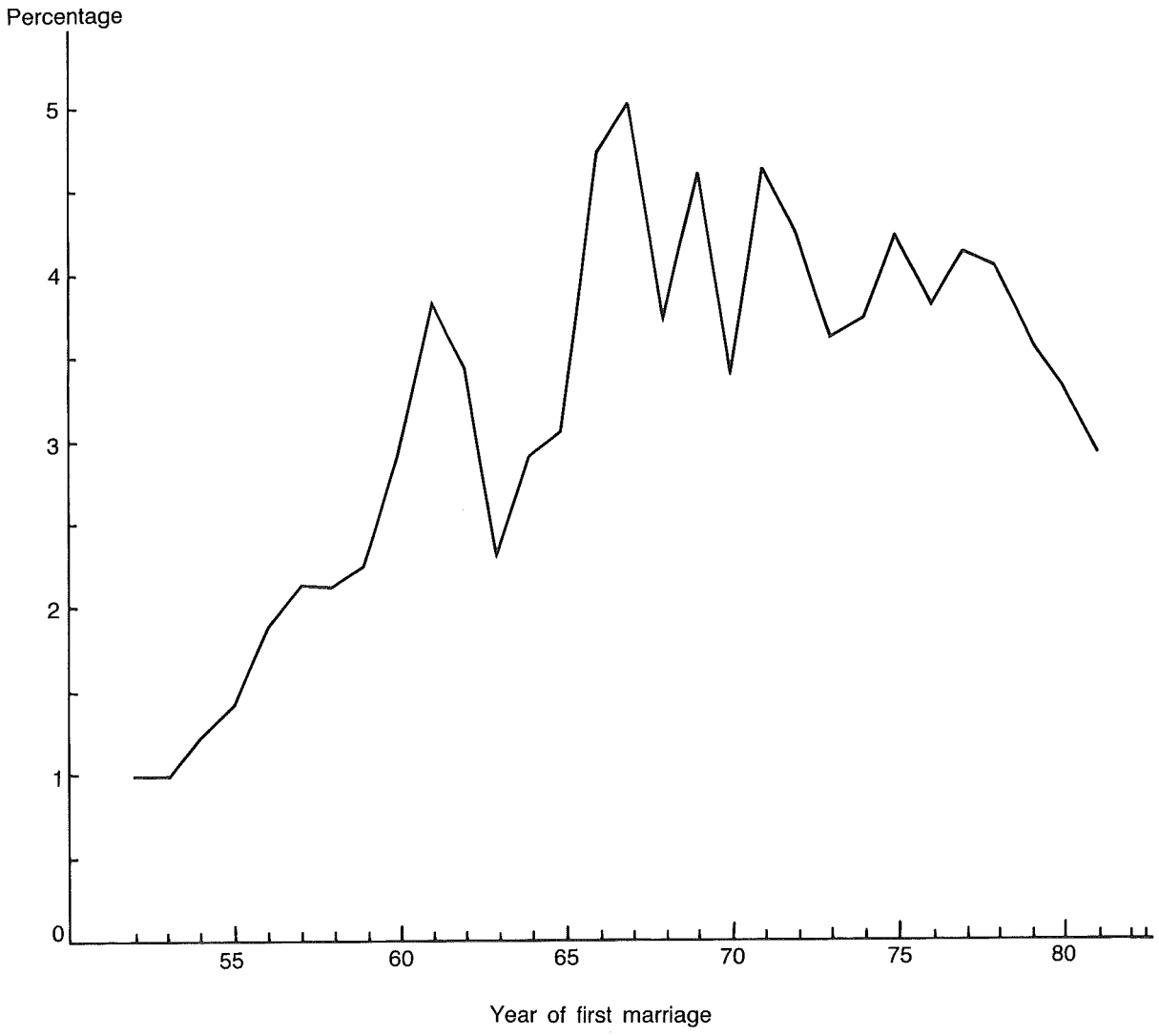

Figure 14 Percentage distribution of all ever-married women according to year of first marriage

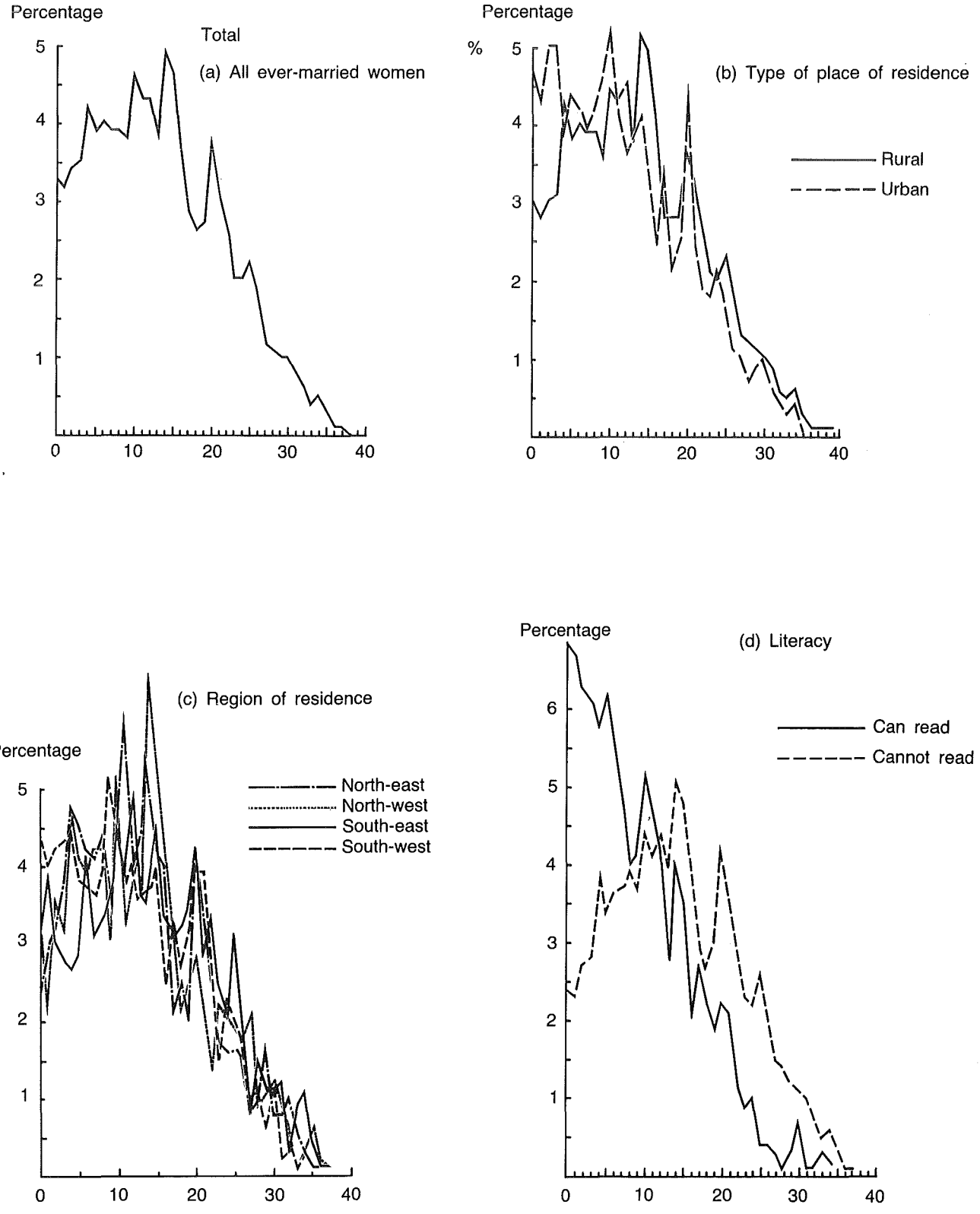

Figure 15 Percentage distribution of ever-married women according to years since first marriage, by selected background characteristics (*continued next page*)

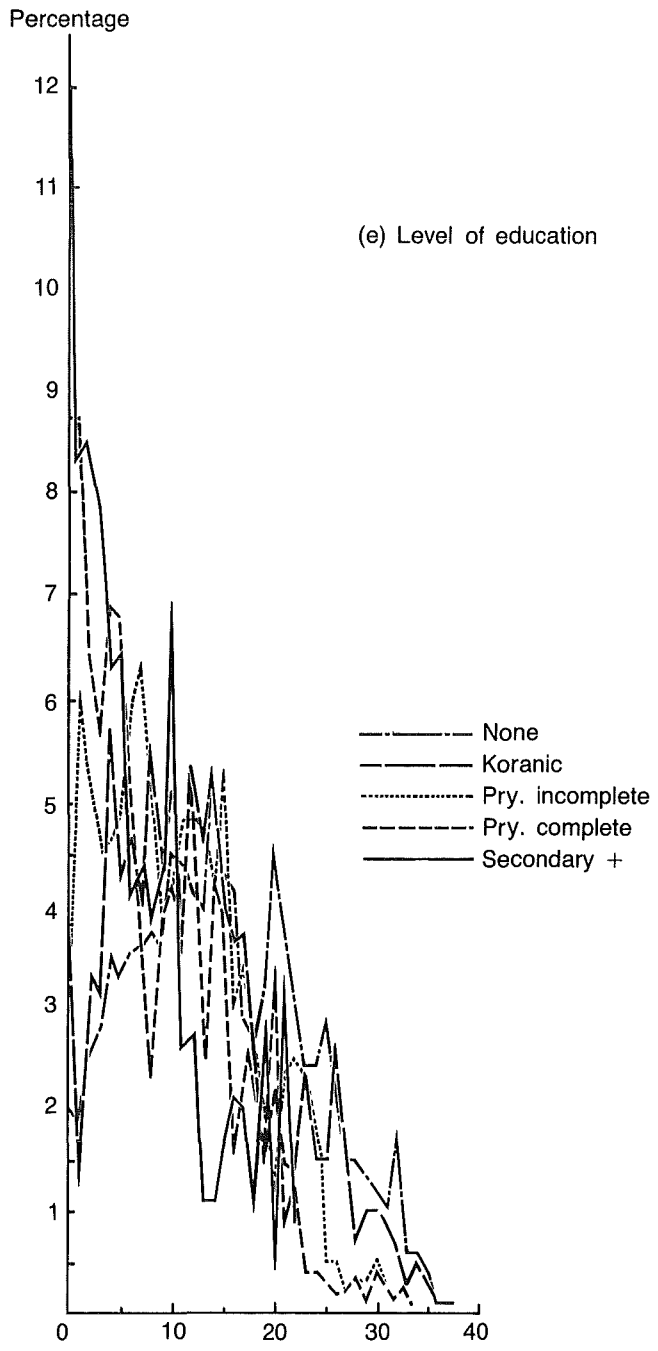

Figure 15 (*continued from previous page*) Percentage distribution of ever-married women according to years since first marriage, by selected background characteristics

4.3 MEAN AGE AT FIRST MARRIAGE

Age at first marriage is one of the most important nuptiality variables affecting fertility performance. Typically, in societies where marriages are relatively stable and the use of voluntary contraception is low, the earlier the age at first marriage, the longer the continuous exposure to childbearing and the higher the completed fertility. It is thus important to examine the quality of data on age at first marriage.

Age at first marriage was calculated by subtracting the date of birth from the date of first marriage, except where the latter was not known and exact age at marriage obtained. It follows that where either or both dates were misreported, then the obtained age at first marriage would be wrong. It has been shown in the previous sections that there had been substantial misreporting of age (or date of birth) and date of first marriage. Considerable heaping of age at first marriage at around 15 years had also been observed.

One method of examining the data is to look at the proportion of women ever married by single years of age for each age cohort. This method not only shows trends in the proportion ever-married over time but can also be used to detect misstatement of age at first marriage or omission of early unions. In the latter context, a consistent increase in the proportion ever-married or a stable proportion by increasing age could indicate better reporting of age at first marriage. A consistent decline or a decline followed by an increase is probably indicative of reporting errors. The relevant information is presented in table 12 and illustrated in figure 17. Entry into union for each cohort has been truncated before the youngest age since some women in each cohort have not been equally exposed to the risk of marriage. The concentration of first marriages around 15 years is once more evident from these proportions. The cumulative proportions also show remarkable differences between women aged under 35 and those older. For the former, age at first marriage appears to have been increasing as shown by the smaller proportions married at each age for younger cohorts. On the other hand, an erratic pattern of lower or higher proportions is observed for women over 35 years of age, especially for marriages before the age of 20. Such fluctuating proportions are probably indicative of misreporting of ages at first marriage by those older women. Though the cumulative proportions of ever-married increase with each single year of age as it should, it would appear that these older women brought forward their age (or date) at first marriage closer to the interview date (ie forward telescoping) thereby reporting lower proportions ever-married at each age compared with the younger cohorts. The smaller proportions may also be due to these older women giving their ages (or dates) at subsequent marriages as their age (or date) of first marriage. Another explanation could be, as shown in section 3, that their actual age has been misreported to a greater extent than for the younger women.

The above evidence of possible misreporting of age at first marriage especially for the older cohorts is once more evident in table 13 showing the proportion of women ever married by cohort for 5-year periods before the survey. The diagonal (from left to right) shows the proportion of each cohort married while in the same age group. It is observed that, for cohorts under 35, the proportions decrease as age decreases (suggesting an increase in the age at first marriage); however, the proportions always fluctuate for the older cohorts. For each period and at each age group, proportionately fewer ever-married women are recorded for the 40–44 cohort than for the immediately younger or older cohort.

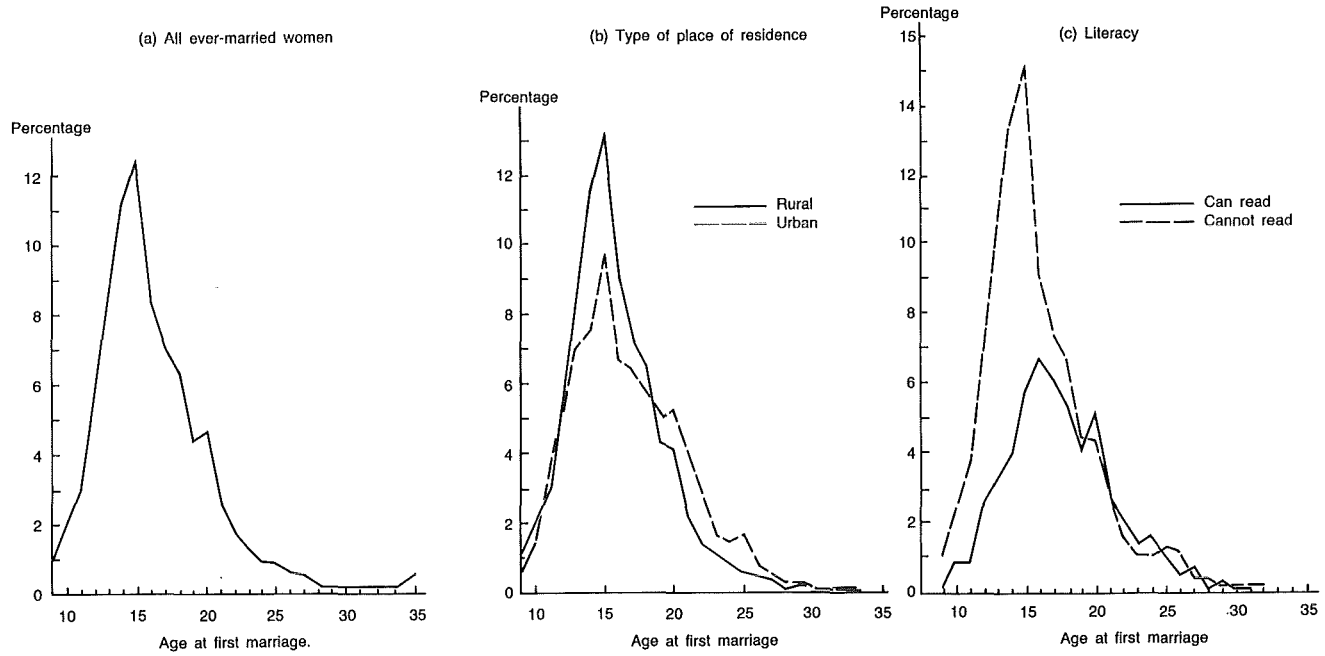

Figure 16 Percentage distribution of ever-married women according to age at first marriage in single years, by selected background characteristics

Table 12 Cumulative percentage of women entering marriage at specified age, by current age

Age at first marriage	Current age							
	15–19	20–24	25–29	30–34	35–39	40–44	45–49	All
9	0.4	0.5	1.1	2.0	1.2	0.6	0.5	0.9
10	1.6	2.7	3.9	4.3	3.5	2.4	1.4	2.9
11	3.4	6.0	7.6	8.9	6.9	3.8	2.8	5.9
12	8.4	13.1	14.8	14.4	12.5	9.4	8.3	11.9
13	14.6	22.3	24.4	23.0	21.7	16.4	19.0	20.3
14	22.1	33.8	37.9	35.6	32.7	23.9	27.6	30.9
15		47.5	50.5	50.7	47.8	35.5	40.0	43.3
16		57.4	59.9	60.7	58.1	44.7	47.1	51.8
17		65.2	67.9	69.5	66.1	52.4	58.4	58.9
18		71.5	75.6	77.6	73.2	61.1	63.8	65.2
19		75.7	79.7	83.3	79.3	70.0	70.5	69.6
20			85.2	88.3	86.6	79.1	76.5	74.2
21			88.0	91.7	89.8	84.0	81.7	76.8
22			90.3	94.4	92.4	85.8	84.9	78.5
23			92.3	95.3	94.0	87.8	88.1	79.1
24			94.3	96.3	94.7	89.8	90.2	80.6
25				96.9	95.6	92.3	92.5	81.4
26				97.7	96.6	94.0	93.5	82.0
27				98.0	97.2	95.5	95.3	82.5
28				98.4	97.3	96.5	95.4	82.7
29				98.8	97.9	97.2	95.6	82.9
30					98.1	97.6	97.0	83.1
31					98.2	97.7	97.5	83.2
32					98.2	98.0	98.1	83.3
33					98.5	98.2	98.6	83.4
34					98.7	98.2	98.6	83.4
35						98.3	99.0	83.5
36						98.5	99.0	83.5
37						98.6	99.1	83.5
38						99.0	99.2	83.5
Never-married	59.7	15.2	3.2	1.0	0.9	1.0	0.6	15.5

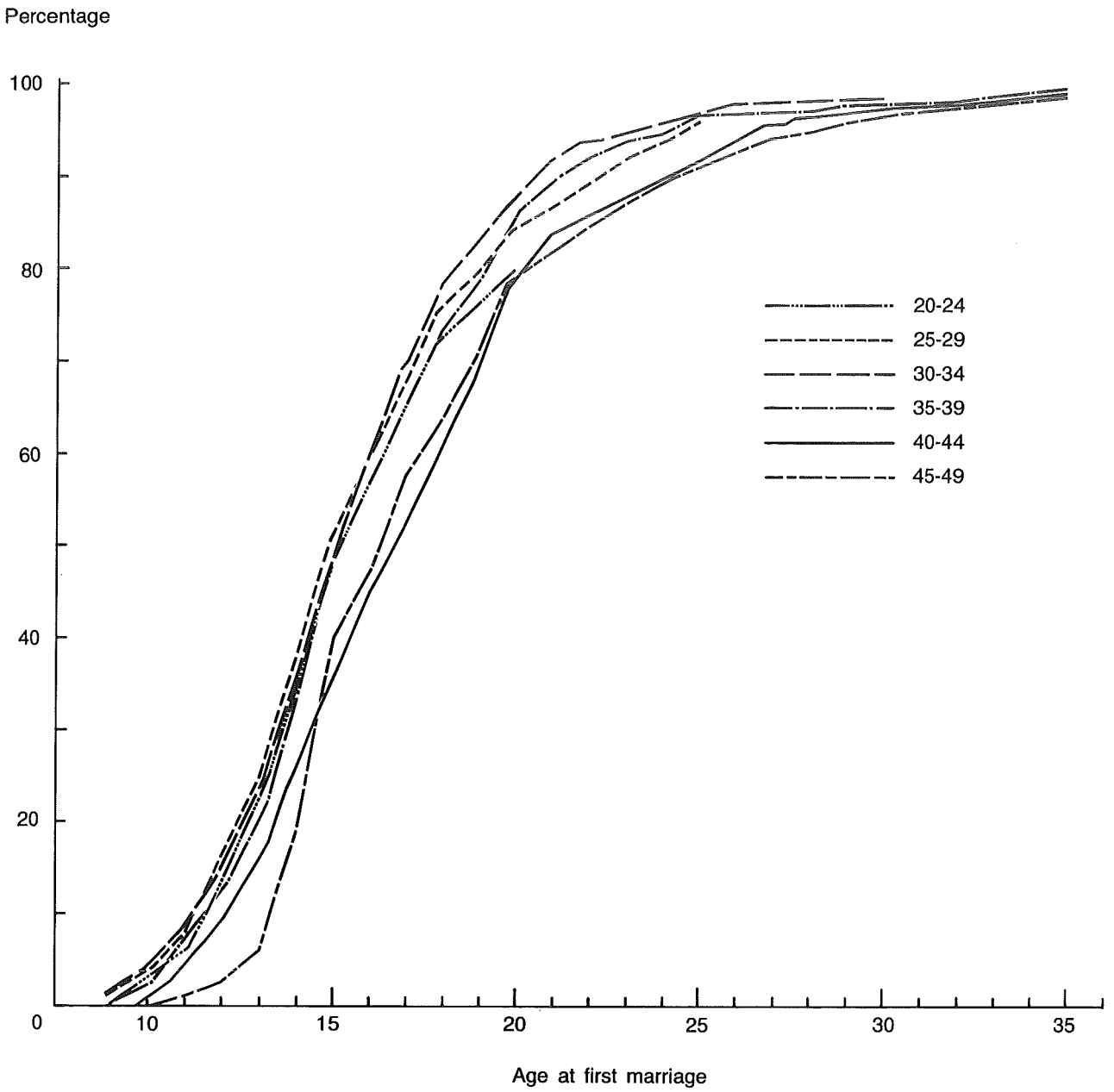

Figure 17 Cumulative percentage of ever-married women according to age at first marriage single years of age, by age cohort

Table 13 Proportion of women ever married, by age group at five-year intervals before the survey

Current age	Years before the survey							
	0	5	10	15	20	25	30	35
15–19	40.1	6.9						
20–24	84.7	54.1	8.0					
25–29	96.8	86.6	58.1	8.7				
30–34	99.0	97.4	89.2	49.9	6.3			
35–39	99.1	98.4	96.0	87.0	59.5	7.2		
40–44	99.0	98.1	97.8	91.7	78.9	35.1	3.1	
45–49	99.4	99.4	98.9	96.8	93.4	81.2	48.7	6.9
All	83.4	68.6	52.1	33.8	20.3	9.1	3.3	0.4

4.4 INTERVAL BETWEEN FIRST MARRIAGE AND FIRST BIRTH

An unsatisfactory but often used method of evaluating the accuracy of the information on age at first marriage is to examine the interval between age at first marriage and age at the birth of the first child. When that interval is considerably less or more than one year, it is regarded as probably being indicative of misreporting of the date of either or both events, but especially the date (or age) of first marriage. The logic is that, for most women, the birth of the first child is a more significant event than marriage and thus its timing is less likely to be forgotten. Very short or negative intervals would tend to indicate greater tendency to report higher age at first marriage than was actually the case, while very long intervals would indicate greater tendency to report lower age at first marriage. The plausibility of this argument in societies where not much importance is attached to dates (of any events) is, however, debatable.

The average interval between the mean ages at first marriage and at first birth for ever-married women who have had at least one birth is approximately two and a half years (30.3 months, as indicated in table 14). This interval is definitely long and would indicate a substantial misreporting of either or both events. The interval is approximately 18 months for women aged 15–24 but then increases to 27.2 and 30.0 months for the next two older age groups and to over three years for the women aged 35–39 and 40–44 (38.1 and 40.3 months respectively) and over four years for women aged 45–49. The intervals for the older women, especially those 35 years or more, are definitely too long and, if the argument applies, would further indicate a tendency to have reported lower ages at first marriage as was observed earlier. They might also have misreported the dates of their first birth to a greater extent than the younger women.

4.5 NUMBER OF TIMES MARRIED

The tendency for respondents to omit marriages, especially earlier marriages of very short durations, can be assessed by examining the number of times the women have been married by age. Since the older women have been exposed to the risk of marital dissolution and remarriage to a greater extent than younger women, one would expect that the number of times married would increase with age. Any decline, especially for women under 35 beyond which age remarriage may become difficult, would indicate omission of marriages.

The data presented in table 15 show that coverage of marriages appears complete or that any omissions are equally distributed by age. The number of marriages for the ever-married women increases with age as expected. Minor fluctuations are, however, observed for some residential or socio-economic categories of women and

Table 14 Mean interval (in months) between age at first marriage and at birth of the first child for ever-married women who have had at least one birth, by current age

Current age	Mean interval (in months)
15–19	18.0
20–24	18.3
25–29	27.2
30–34	30.0
35–39	38.1
40–44	40.3
45–49	48.7
All	30.3

Table 15 Mean number of times married for ever-married women, by age and selected background characteristics

Background characteristics	Age						
	15–19	20–24	25–29	30–34	35–39	40–44	45–49
A Type of place of residence							
Rural	1.04	1.07	1.09	1.14	1.15	1.18	1.25
Urban	1.05	1.07	1.12	1.12	1.28	1.26	1.19
B Region of residence							
North-east	1.05	1.08	1.12	1.17	1.28	1.31	1.36
North-west	1.04	1.11	1.12	1.18	1.18	1.26	1.29
South-east	1.04	1.05	1.06	1.11	1.12	1.13	1.21
South-west	1.05	1.04	1.08	1.08	1.14	1.15	1.16
C Level of education							
None	1.03	1.07	1.11	1.13	1.19	1.20	1.25
Koranic	1.09	1.03	1.09	1.25	1.24	1.23	1.25
Primary incomplete	1.06	1.06	1.07	1.09	1.21	1.19	1.37
Primary complete	1.01	1.05	1.04	1.17	1.03	1.08	1.04
Secondary and above	1.01	1.02	1.04	1.02	1.00	1.09	1.00
D Literacy							
Can read	1.04	1.05	1.05	1.10	1.11	1.13	1.04
Cannot read	1.04	1.08	1.10	1.14	1.20	1.21	1.26
E All	1.04	1.07	1.10	1.13	1.18	1.20	1.24

are likely to be due to sampling variations since the number of women involved is almost always small in those instances.

4.6 SUMMARY

The data presented in this section show that there is substantial heaping in the NFS nuptiality data involving dates. The pattern of heaping varied with the format of the information on the date of first marriage of the respondents which could be in calendar month and year, calendar year only and age at first marriage. It also varied according to year of interview (1981 or 1982). The tendency was for heaping to occur in rounded years, in rounded durations from date of interview, or both. The combination of these tendencies produced a very irregular pattern of heaping at both preferred and usually avoided digits for year of first marriage and years since first marriage not only for the entire sample but for various subgroups of women. It was also observed that there has been a recent decline in the proportion of women ever married.

The analysis on age at first marriage showed a very substantial concentration at around the age of 15, more concentration than would have been explained simply by the cultural patterns of the timing of first marriage. It was also observed that the older respondents tended to have brought forward their date of first marriage closer to the date of interview thereby reporting higher ages at first marriage and lower proportions married at each age. The average interval between first marriage and first birth was found to be excessively long especially for the older women further confirming that they were more likely than the younger women to have misreported their dates (or ages) at first marriage. There was, however, no evidence to indicate that there was an undercoverage of marriages resulting from omission of marriages; if such omissions occurred, the incidence was likely to have been equally distributed for the age groups.

5 Fertility

5.1 INTRODUCTION

Information on fertility and mortality was collected in the NFS by means of a comprehensive maternity history supplemented by additional questions. The number of children a woman has given birth to, by sex and current place of residence (home or away from home), was obtained. Then the number of children who had died, by sex, was ascertained. The total from these two sources was checked with the respondent for correctness. Then followed a comprehensive listing of all the children, dead or alive, in the birth history section. For each live birth, listed in chronological order from the first, the following information was collected: birth order, name, sex, date (or years ago) of birth, survival status and if dead, age at death (in months and years). For each interval between births, questions were asked as to whether any pregnancies had occurred, the duration of such pregnancies, if seven months or more whether they terminated in live births or still births and if less than seven months, how they terminated.

All fertility data in the NFS were derived from the information as gathered from the above. But it is generally known that fertility data based on birth histories are subject to various types of error which may affect the derived estimates. Such errors mostly arise from incorrect reporting of the dates of births, the omission of births or the misreporting of the ages of the women themselves. Experience has shown that, in most surveys, respondents do not often report all births, especially those that occurred in the remote past or in instances where the children died soon after birth. Such omissions, the extent of which may vary by periods before the survey, may distort estimates of fertility levels at different periods thus giving a spurious indication of increase or decrease. Incorrect reporting of dates of births results in the displacement of those births. Such incorrect reporting is known to be more characteristic of births which occurred long before the date of interview or of births to older women; the latter is more often in the form of forward telescoping, ie a tendency to bring the reported dates closer to the date of interview. Again such displacements may result in spurious increases or decreases in fertility. Misreporting of the ages of the respondents themselves leads to similar problems; understatement or overstatement of age may lead to lower or higher estimates for the age cohort or particular period. These sources of error often interact to produce a pattern of current or past levels of fertility that are incorrect and distorted.

The attempt is made in this section to identify errors from the above sources in the NFS data on fertility, and the effects of those errors on the obtained fertility rates. Since there are no external sources with which the data can be compared, the assessment would, again, be based solely on internal checks of consistency and conformity with logical or theoretical patterns. As a preliminary step, the assessment will start with the form of dating of the live births as obtained from the birth histories.

5.2 DATING OF LIVE BIRTHS

An examination of the form of dating of the live births may give an indication of the quality of the obtained data. For instance, births for which the exact date in calendar month and year was obtained are less likely to have been displaced compared with those for which only the age of the child at interview was available. Any heaping observed in the distribution of the births would also be expected to vary, most probably in rounded years when exact date or calendar year only was given and in rounded durations from date of interview if date of birth was calculated from age (or years ago of the birth) at interview. Differences in the form of dating by age or region may also indicate, and possibly help explain, differentials in quality of other fertility data.

The format of dating of the first, penultimate and last births is shown in table 16. It is immediately evident, when considering the proportion of births for which the exact date was available, that the quality of dating is better for the last than for the first or penultimate births, generally better as the age of the women decreases (except for the youngest age group 15–19), better in the urban than in the rural areas and in the south-east and south-west than in the north-west and north-east. On the whole, the exact date was provided for only 26.3 per cent of the first births, 25.4 per cent of the penultimate births and 34.2 per cent of the last births. The proportion of births for which calendar year only was given is seemingly high at about a third of each birth (almost a half in the south-east). It would appear that in some instances, where only the years ago of the birth had been supplied by the respondent, the interviewers had calculated and recorded calendar year of birth after entering the birth in the date-event chart provided in the questionnaire. Dating of births was particularly bad in the north-west region.

Despite the above, not much heaping is observed in the distribution of all live births according to year of birth as shown in figure 18. No heaping in rounded years is evident, though there are minor heaps a year before or after such rounded years – evidently arising from ages given in rounded durations at interview. An important observation is the apparent decline in the proportion of births from 1978 onwards, that would indicate some omission of births under three or four years of age.

Table 16 Percentage distribution of first, penultimate and last live births according to format of reporting birth dates and selected characteristics

Characteristics	First birth				Penultimate birth				Last birth			
	Exact date	Calendar year only	Years ago	Total	Exact date	Calendar year only	Years ago	Total	Exact date	Calendar year only	Years ago	Total
A Current age												
15–19	38.9	25.5	35.6	100.0	29.9	27.7	42.4	100.0	41.9	25.2	32.8	100.0
20–24	37.5	31.8	30.7	100.0	33.8	34.0	32.2	100.0	44.6	28.6	26.8	100.0
25–29	28.5	31.7	39.8	100.0	28.2	31.5	40.3	100.0	38.9	27.9	33.1	100.0
30–34	23.0	37.0	40.0	100.0	26.2	34.2	39.6	100.0	33.9	31.7	34.4	100.0
35–39	20.5	39.9	39.7	100.0	20.9	39.9	39.2	100.0	27.4	36.4	36.1	100.0
40–44	17.9	38.9	43.3	100.0	20.2	40.1	39.8	100.0	23.6	38.4	38.0	100.0
45–49	13.9	37.3	48.7	100.0	16.8	36.6	46.6	100.0	18.5	34.4	47.1	100.0
B Type of place of residence												
Rural	23.4	36.9	39.7	100.0	23.1	37.4	39.5	100.0	31.6	33.3	35.0	100.0
Urban	36.8	26.9	36.3	100.0	34.4	26.9	38.7	100.0	43.6	24.8	31.6	100.0
C Region of residence												
North-east	22.8	36.2	41.0	100.0	22.9	35.4	41.7	100.0	31.6	33.7	34.7	100.0
North-west	3.7	22.4	73.9	100.0	2.7	21.3	76.0	100.0	7.0	22.7	70.3	100.0
South-east	39.2	52.4	8.4	100.0	38.1	54.1	7.8	100.0	48.7	45.4	5.9	100.0
South-west	39.4	23.2	37.4	100.0	35.7	23.2	41.1	100.0	49.5	19.8	30.7	100.0
D All	26.3	34.8	39.0	100.0	25.4	35.2	39.3	100.0	34.2	31.5	34.3	100.0

5.3 OMISSION OF BIRTHS

Children ever born

In the absence of recent increases in fertility, the number of children ever born should increase with increasing age of the women. So a simple method of assessing the extent of omission of births is to examine the mean number of children ever born by the age of the women at the time of the survey as presented in tables 17 and 18 and illustrated in figure 19. It is observed that there are no substantial fluctuations in the data except at older ages (35 years and above). For these older women the mean number of children ever born fluctuates especially around rounded ages; such fluctuations could be due to sampling variability, omission of births or distortions due to the heaping on the age of respondents. The latter could have arisen if women had reported higher ages, thereby depressing the obtained mean number of children. This should actually be the explanation for the trough around ages 30 to 43 and some ages ending in 0 or 5. Though the pattern of distribution in figure 19 seems characteristic of most developing countries, it is safe to conclude that there has been some omission of births to older women. Such omission, however, does not affect the mean number of children ever born when the data are presented for 5-year age groups as in table 18.

Sex ratios of births

The sex ratio of births can be used as more evidence of (selective) omission of births. Typically, there are more male than female births in any society and sex ratios at birth usually range between 103 and 107 male births per 100 female births. Substantial variations in these ratios can be taken as indicative of selective omission of births by sex. Values substantially above the range would indicate omission of female births while values substantially below would indicate omission of male births.

The overall sex ratio of births recorded in the NFS is 113. This is higher than the expected range of sampling variation and shows some omission of female births. Such selective omission is almost characteristic of all age groups of the women (except for those aged 45–49 among whom omission of both sexes appears common) and greater in the rural than in the urban areas and in the north-west and north-east than south-east and south-west (table 19).

When the sex ratios by birth order are considered, it is observed that under-reporting of female births is greater for the first, second or third than for higher order births. On the whole, these higher order births seem to have been completely covered as evidenced by a sex ratio of 104. Sex ratios by 5-year periods before the survey show adequate coverage for the 0–4-year period (103) but systematic omission of females for subsequent periods. Cross-classification of the ratios by 5-year periods and age cohort of mother shows almost a systematic omission of female births among all cohorts and periods for the past 20 years except for a few, eg omission of males 0–4 years before the survey for women aged 25–29 (tables 20 and 21).

Thus the preponderance of the evidence from the analysis of data on sex ratios shows that there has been a systematic omission of some female births by almost all age cohorts especially at lower order parities and for all five-year periods before the survey except the last. Such omissions would tend to lower the obtained fertility estimates.

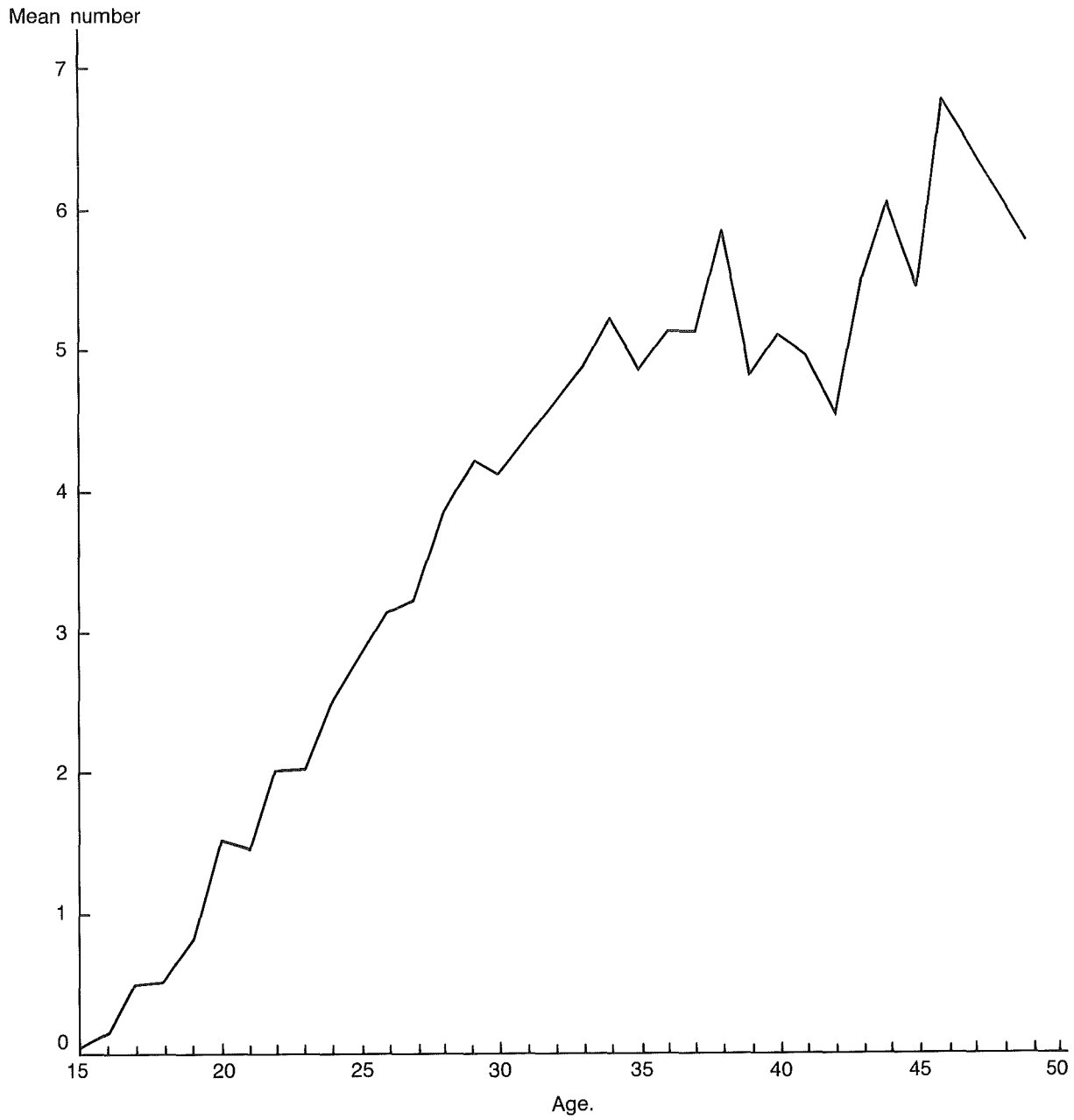

Figure 18 Mean number of children ever born for all women, by single years of age

Proportion of children who died

It is commonly assumed that any omission of births in a birth history would involve children who had died, especially those that died at younger ages. This should particularly be the case in cultures where women are most reluctant to discuss their dead children. In the absence of evidence for increasing infant or child mortality, the proportions of dead children should increase with the age of the mother since older women are likely to have more children and these children have been exposed for greater periods to the risk of death. Moreover, probabilities of infant or child death are higher for women in older ages. Thus a further test for omission of births would be the examination of the proportion of children ever born who had died by the time of interview by current age of the women as presented in table 22.

The data show that, for all deaths, the proportion of dead children generally increases with age of respondents as would be expected; the only exceptions are women aged 30–34 who report proportionately fewer deaths than women in the next younger age group. Further examination of the data by sex of child shows that the proportion of dead females again increases by age; for males, however, the proportion for women aged 30–34 is again lower than for women in the next younger age group. It would, thus, appear that there was no systematic omission of dead children except among women aged 30–34 who tended to have omitted some of their dead boys. It should, however, be borne in mind that this age group is the one mostly affected by the

Table 17 Mean number of children ever born to all women, by age in single years

Age	Children ever born	Number of women	Age	Children ever born	Number of women
15	0.07	531	33	4.87	99
16	0.15	492	34	5.23	111
17	0.48	386	35	4.87	570
18	0.50	430	36	5.17	209
19	0.83	261	37	5.22	89
20	1.49	680	38	5.84	143
21	1.44	265	39	4.82	99
22	1.99	302	40	5.18	587
23	2.05	200	41	4.95	132
24	2.50	264	42	4.52	95
25	2.78	778	43	5.46	59
26	3.12	306	44	6.02	31
27	3.22	212	45	5.45	255
28	3.84	318	46	6.74	78
29	4.22	152	47	6.35	39
30	4.12	1026	48	5.99	126
31	4.36	159	49	5.76	93
32	4.63	151	All	3.07	9729

Table 18 Mean number of children ever born to all and ever-married women, by current age in five-year groups

Current age	Number of children ever born	
	All women	Ever-married women
15–19	0.35	0.83
20–24	1.79	2.10
25–29	3.21	3.31
30–34	4.32	4.36
35–39	5.07	5.12
40–44	5.13	5.15
45–49	5.84	5.87
All	3.07	3.67

tremendous heaping of respondents at age 30, a heaping shown to be more likely due to younger women being moved up in age. The effects of such heaping can be reduced if broader age groups are used.

Proportion of surviving children living away from home

A crude indicator for omission of children is the proportion of surviving children living away from home at the time of interview. The underlying assumption of this indicator is that the children of older respondents are likely to be older than the chidren of younger respondents and are thus more likely to have been married away, if daughters, to have married and set up their new homes, if sons or to have moved out to set up by themselves, especially if they are old enough to work.

Table 19 Sex ratio of all births for all women, by age and by selected background characteristics

Characteristic	Sex ratio (per 100)
A Current age	
15–19	125
20–24	113
25–29	107
30–34	117
35–39	120
40–44	114
45–49	103
B Type of place of residence	
Rural	114
Urban	109
C Region of residence	
North-east	116
North-west	119
South-east	109
South-west	112
D All	113

This indicator is, however, affected by the various cultural or economic patterns that make women allow their children to live with other people and thus away from home. The available information (in table 23) shows that, for women aged 20 and over, the proportion of surviving children living at home decreases as the age of the respondent increases. The pattern is the same both for sons and daughters and the proportion living away from home is over 40 per cent for women aged 45–49. The exceptions are the youngest women aged 15–19 who have lower proportions living away from home for both sexes than the next older age group.

Number of births

In any growing population, the absolute number of births should be increasing, except for very sudden and drastic declines in fertility levels. In the absence of such declines, any substantial decreases in the number of children at a given age should be indicative either of omission of births or incorrect reporting of age for the children. The distribution of the actual number of children born by the respondents by years before the survey is shown in figure 20. The numbers fluctuate quite substantially showing the familiar heaping at years ending in 0 and 5 (also at 12) and the troughs at years ending in 1 and 9 (and at age 4). The distribution also indicates that recent births might have been omitted as indicated by the dip in the curve from age 3 to 0. This would correspond to the decline in the proportionate distribution of live births by year of birth as shown in figure 19. If more recent births were actually omitted, then the recent fertility levels for the 5-year period before the survey would be underestimated. Omission of recent births seems to be one of the more characteristic features of demographic surveys in the developing countries.

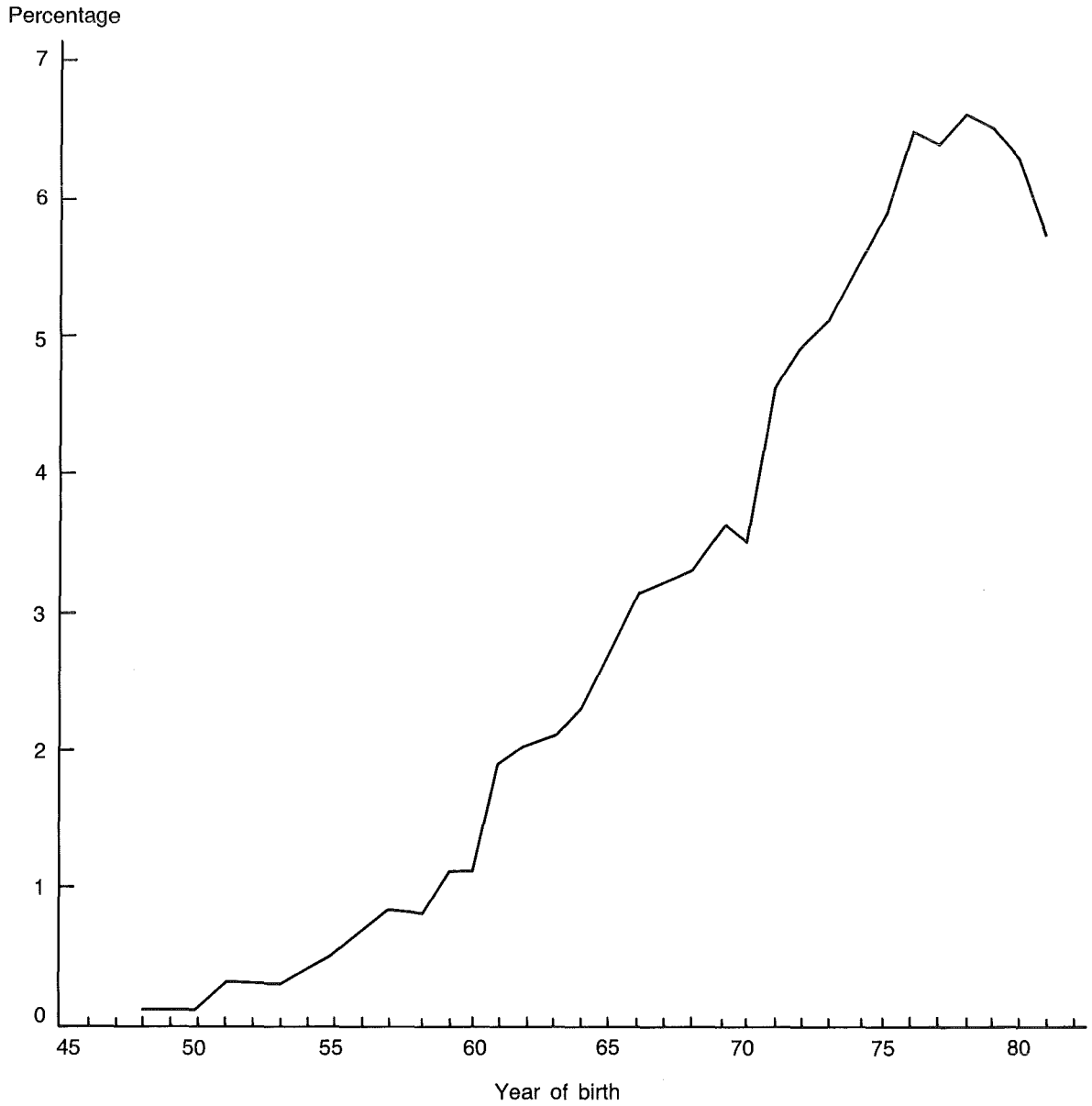

Figure 19 Percentage distribution of all live births according to year of birth

Table 20 Sex ratios by birth order for five-year periods before the survey

Period before the survey	Birth order			
	First birth	Second or third births	Fourth and above	All births
0–4 years	121	106	96	103
5–9	121	124	116	120
10–14	131	117	107	117
15–19	125	111	91	109
20–24	144	103	129	123
25–29	145	132	(134)	139
30–34	(168)	(176)	–	(172)
All	128	114	104	113

5.4 MISREPORTING DATE OF BIRTH OR AGE OF CHILDREN

Evidence from several surveys has shown that though respondents may be able to give a complete record of their maternity history, they may, however, misreport the dates of the birth of their children or their ages. Typically, such misreporting is more characteristic of the older women (as evidenced by the greater proportion of respondents in older ages who could not give the exact date of their births in table 16) and the less educated, and often involves births that occurred further in the past from the survey date. Such misreporting does not follow a consistent pattern though the experience in developing countries shows a tendency on the part of the older women to report their early births as occurring closer to the survey date than is actually the case. This results in a forward displacement of the births and their occurrence

Table 21 Sex ratio at birth by period and age cohort of mother

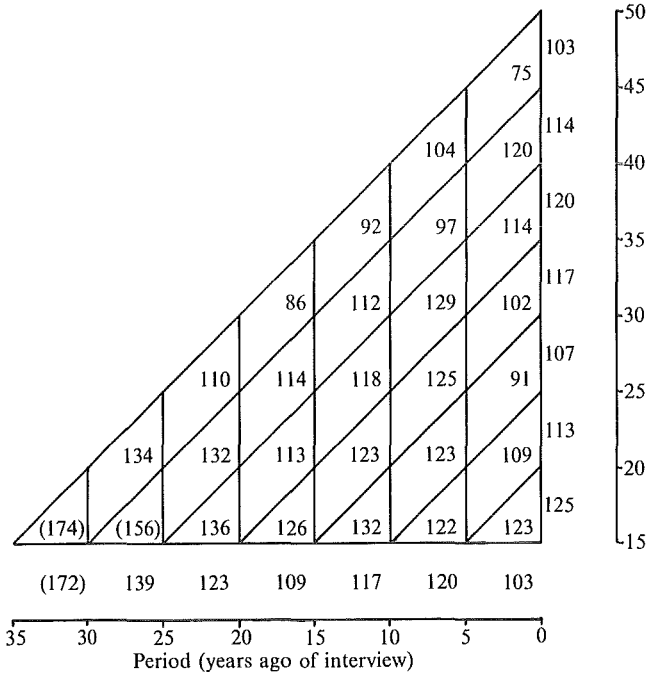

Table 22 Proportion of children who died, by sex and current age of mother

Current age	Proportion of children dead		
	Male	Female	Total
15–19	13.8	12.3	13.1
20–24	16.2	13.5	14.9
25–29	16.5	14.1	15.3
30–34	14.9	15.2	15.0
35–39	18.6	17.2	17.9
40–44	22.7	21.1	21.9
45–49	25.6	21.5	23.6
All	18.4	16.8	17.6

Table 23 Proportion of surviving children living at home at the time of interview, by sex and age of respondents

Current age	Proportion of children living at home		
	Sons	Daughters	Both sexes
15–19	89.5	90.3	88.7
20–24	90.2	88.7	91.8
25–29	85.2	85.6	86.2
30–34	82.4	81.1	83.2
35–39	78.5	71.7	75.8
40–44	72.8	66.9	70.6
45–49	57.9	56.2	57.6
All	79.3	76.4	77.9

at older ages of the women. More recent births are typically not subjected to the same extent of forward displacement (if at all) and the net effect is a concentration of births in an intermediate period, eg 5 to 15 years before the survey. When this occurs a spurious trend of rising fertility followed by a decline in more recent periods may be observed. If fertility was already declining, then the extent of decline would be exaggerated (Potter 1977a; Goldman 1984). The problem is compounded when the displacement of births (whether forwards or backwards) is associated with misreporting of the age of the respondents themselves. If the age of the woman is overestimated and her births displaced forwards, then the births would be seen as having occurred to her at very high ages resulting in an exaggerated mean/median age at birth. If, on the other hand, her age was underestimated and the births displaced forwards, the reverse will be the case.

Whether such displacement has occurred can be investigated by examining the age at first birth for the various age cohorts of the women. This is done here by means of the (cumulative) proportion of women in each age cohort who are reported to have given birth to their first child in single years of age as shown in table 24. The figures reveal marked differences between women aged from 20–24 to 30–34 and the older women. Among the latter group, the proportion having their first child at each single year of age is consistently lower when compared with the younger women; also for this latter group, the proportions decrease with increasing age group. For instance, while about a quarter of the women aged between 20 and 34 years had their first child around the age of 15, women aged above 40 did so just before the age of 17. Similarly, while the median age at first birth for women 20–34 is around 18, it is slightly over 20 for women aged 40–44 and 21 for women aged 45–49. This pattern is rather consistent and the difference

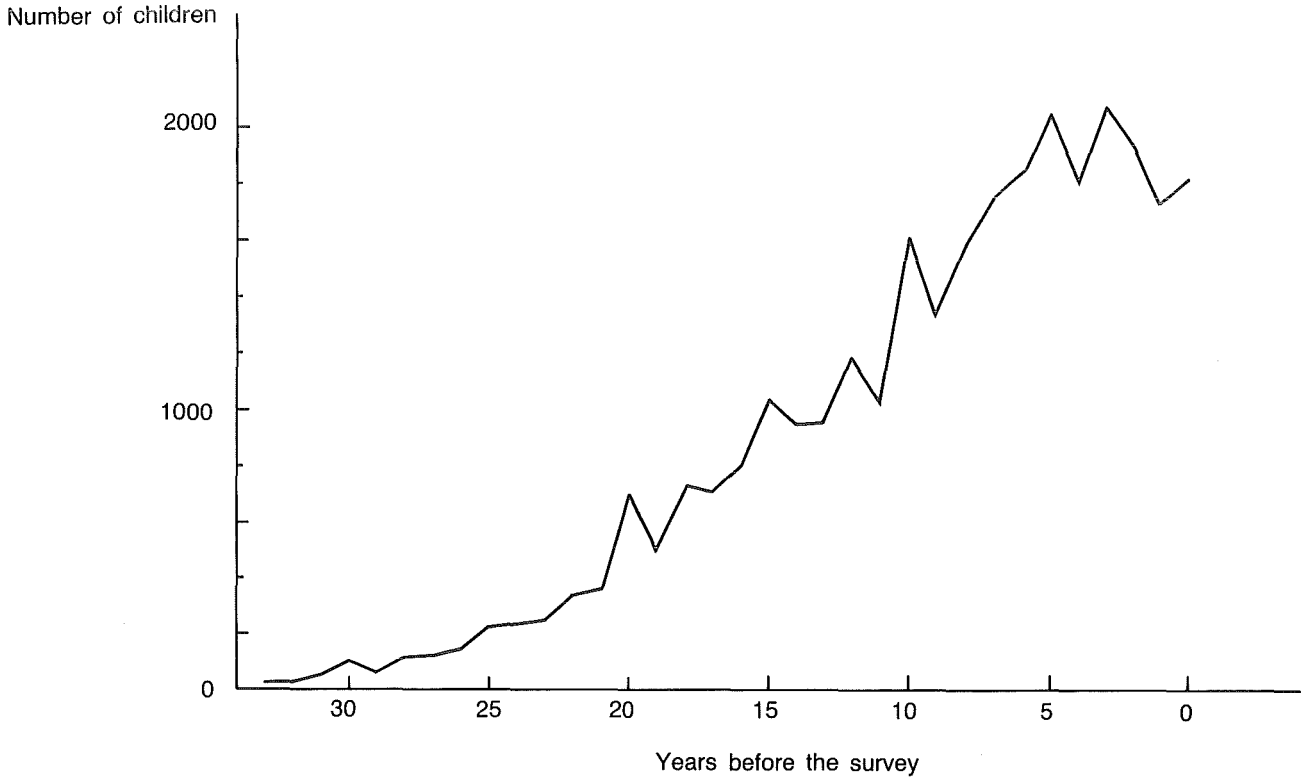

Figure 20 Number of children born, by single years before the survey

Table 24 Cumulative percentages of respondents to the individual questionnaire according to age at first birth in single years, by age cohort

Age at first birth	Age cohort							
	15–19	20–24	25–29	30–34	35–39	40–44	45–49	All
11	0.2	1.0	1.4	1.6	0.2	0.2	0.0	0.8
12	1.6	3.9	4.5	3.9	1.2	2.2	0.9	2.9
13	3.6	9.3	8.4	7.5	3.8	3.5	2.0	6.0
14	8.0	15.7	14.5	15.2	10.2	7.4	7.0	11.8
15		24.5	24.1	24.4	20.0	13.0	13.0	19.5
16		35.4	33.3	32.7	29.4	18.1	18.0	27.4
17		44.4	41.7	41.7	37.0	26.4	25.1	34.8
18		53.1	50.7	51.1	43.3	30.8	32.5	41.5
19		61.8	59.6	60.2	52.0	40.5	37.8	48.6
20			67.5	67.7	60.0	51.4	45.5	54.6
21			73.7	73.4	66.2	56.4	50.7	58.8
22			78.9	78.1	71.5	61.8	54.9	62.1
23			83.0	82.8	75.7	64.4	60.6	64.8
24			86.0	86.3	79.5	68.5	66.6	67.2
25				88.9	82.2	73.6	71.2	69.1
26				90.8	84.7	76.1	73.5	70.3
27				92.4	86.6	79.1	75.6	71.2
28				93.0	87.8	82.1	81.3	72.1
29				93.5	89.3	84.3	82.0	72.6
30					91.0	87.0	83.2	73.2
31					91.8	87.7	84.3	73.5
32					92.7	88.9	85.3	73.8
33					93.2	89.2	86.5	74.0
34					93.5	89.9	87.2	74.1
35						90.6	88.9	74.3
Not yet a mother	75.3	25.5	10.0	6.0	5.9	7.4	8.3	25.4

reaches over 5 years by the time three-quarters of the women have had their first birth.

In a situation where the age at first birth should be stable or increasing instead of decreasing (as in the case of the younger women), the observed pattern shows strongly the effect of greater misreporting of age of the older respondents, the forward displacement of the date of birth of their first child, greater omission of early births by the older women or all of the above. Despite the poor quality of the data on the age of the respondents themselves, there are strong indications that some omission of births to the older women has occurred and that there might have been some forward displacement of the date of first birth especially by the older women. This is bound to affect subsequent analysis of fertility trends.

5.5 EFFECTS OF THE OBSERVED ERRORS ON ESTIMATES OF RECENT AND PAST FERTILITY LEVELS

The effects of errors due to omission and displacement of births and to misreporting of the ages of the respondents on estimates of recent and past fertility can be assessed by examining the cohort-period fertility rates of the women. These rates would permit an analysis of the fertility experience of the women (by age) throughout their reproductive life by taking into account the number of births occurring in specific five-year periods and the age of the women at the time of interview, not at the birth of the children.

The necessary data for such analysis by age cohorts at the time of the survey are presented in table 25. Panel A of the table contains the cohort-period fertility rates for five-year periods before the survey, panels B and C the cumulative fertility rates for real and synthetic cohorts respectively and panel D the P/F ratios. For a proper interpretation of the data on rates (panels A, B and C), the figures read horizontally show fertility rates for each cohort at different five-year periods from the survey date. Diagonally (from top left to bottom right), the figures show the fertility rates experienced by the different cohorts while at the same age. Finally, the figures read vertically show the fertility of the different cohorts at the same five-year period. The cumulative fertility rates in panel B are obtained by summing up the cohort-period fertility rates horizontally for the periods (and multiplying by five) for each cohort and represent the

Table 25 Cohort-period fertility rates, cumulative rates for real (P) and synthetic (F) cohorts and P/F ratios

Age cohort at survey	Number of women	Five-year periods before the survey						
		0–4	5–9	10–14	15–19	20–24	25–29	30–34
A Cohort-period fertility rates								
15–19	2101	0.065	0.005					
20–24	1710	0.245	0.107	0.006				
25–29	1766	0.293	0.251	0.088	0.009			
30–34	1546	0.257	0.304	0.221	0.078	0.004		
35–39	1110	0.199	0.267	0.260	0.212	0.074	0.004	
40–44	904	0.128	0.219	0.238	0.241	0.162	0.037	0.001
45–49	591	0.086	0.144	0.215	0.256	0.238	0.164	0.065
B Cumulative fertility rates for real cohorts (P)								
15–19		0.350	0.027					
20–24		1.792	0.564	0.028				
25–29		3.209	1.742	0.485	0.045			
30–34		4.322	3.034	1.516	0.410	0.018		
35–39		5.075	4.080	2.744	1.446	0.388	0.018	
40–44		5.126	4.484	3.390	2.201	0.999	0.190	0.003
45–49		5.844	5.415	4.697	3.620	2.342	1.153	0.334
C Cumulative fertility rates for synthetic cohorts (F)								
15–19		0.350	0.027					
20–24		1.577	0.564	0.028				
25–29		3.044	1.821	0.468	0.045			
30–34		4.331	3.338	1.575	0.410	0.018		
35–39		5.326	4.674	2.872	1.446	0.388	0.018	
40–44		5.968	5.769	4.061	2.201	1.197	0.205	0.003
45–49		6.397	6.487	5.138	3.620	2.386	1.025	0.325
D P/F ratios								
20–24		1.136	1.002					
25–29		1.054	0.957	1.037				
30–34		0.998	0.909	0.963	0.938			
35–39		0.953	0.873	0.955	0.968	1.001		
40–44		0.859	0.777	0.835	0.816	0.835	0.925	
45–49		0.914	0.835	0.914	0.911	0.982	1.125	1.025

entire reproductive performance of an age cohort at specific five-year periods before the survey. Similarly, the cumulative fertility rates for synthetic cohorts are obtained by summing up the cohort-period rates vertically (and multiplying by five) for the successive age cohorts. Dividing the cumulative fertility of real cohorts (P) by that of synthetic cohorts (F) yields the P/F ratio. Since this ratio compares parity reported as of the survey date (by women of a particular age) with cumulative fertility rates (up to the same age), it is an important indicator of both changes in fertility levels and errors in the fertility data. If fertility had remained constant during the period of interest, and if there are no errors of omission or displacement of births, the P/F ratios should equal unity (1.00). Ratios greater than unity typically reflect fertility decline while ratios less than unity indicate fertility increase. Declining P/F ratios by age are typically indicative of omission of births by the older women. The reverse indicates fertility decline (or omission of births by younger women). Where the ratios show the effects of errors rather than genuine changes in fertility, they can be used as adjustment factors for reported fertility rates or the total fertility rate TFR (Goldman 1984).

Panel A of table 25 shows that (except for the oldest and the youngest cohorts), the fertility rates experienced by each cohort while at the same age (ie at each central age) increases systematically from the more distant 5-year periods since the survey, reaches a peak at the 5–9-year period and then declines in the most recent (0–4-year) period. Consequently, total fertility rates calculated for these rates increase from 5.24 for the

Table 26 P/F ratios for the three five-year periods before the survey, by age cohorts and selected background characteristics

Background characteristics and age cohorts	Five-year periods before the survey		
	0–4	5–9	10–14
A Type of place of residence			
Rural			
20–24	1.146	0.998	
25–29	1.066	0.965	1.038
30–34	1.007	0.905	0.954
35–39	0.953	0.856	0.942
40–44	0.876	0.770	0.830
45–49	0.930	0.828	0.915
Urban			
20–24	1.106	1.019	
25–29	1.009	0.919	1.026
30–34	0.941	0.896	0.968
35–39	0.937	0.930	0.998
40–44	0.783	0.802	0.847
45–49	0.823	0.837	0.888
B Region of residence			
North-east			
20–24	1.043	0.937	
25–29	0.991	0.926	1.044
30–34	0.922	0.842	0.894
35–39	0.807	0.727	0.818
40–44	0.708	0.661	0.787
45–49	0.838	0.790	0.955
North-west			
20–24	1.022	0.995	
25–29	0.934	0.886	0.983
30–34	0.838	0.800	0.931
35–39	0.882	0.862	1.001
40–44	0.697	0.673	0.739
45–49	0.632	0.593	0.657
South-east			
20–24	1.266	1.017	
25–29	1.184	1.023	1.086
30–34	1.207	1.038	1.027
35–39	1.110	0.955	0.995
40–44	1.062	0.882	0.912
45–49	1.099	0.925	0.978
South-west			
20–24	1.137	1.045	
25–29	1.011	0.963	0.992
30–34	0.957	0.945	1.023
35–39	0.937	0.931	0.999
40–44	0.854	0.829	0.833
45–49	0.906	0.871	0.887
C Education			
None			
20–24	1.035	0.973	
25–29	0.940	0.883	1.041
30–34	0.905	0.850	0.941
35–39	0.869	0.825	0.948
40–44	0.825	0.771	0.858
45–49	0.869	0.814	0.913
Koranic			
20–24	1.035	0.975	
25–29	0.990	0.914	0.979
30–34	0.826	0.785	0.936
35–39	0.924	0.870	1.062
40–44	0.661	0.673	0.801
45–49	0.689	0.681	0.790
Primary incomplete			
20–24	1.175	1.015	
25–29	1.025	0.992	1.038
30–34	1.006	0.962	0.947
35–39	0.973	0.902	0.853
40–44	0.787	0.676	0.665
45–49	1.038	0.946	1.048
Primary complete			
20–24	1.057	1.015	
25–29	1.039	1.033	0.974
30–34	1.096	1.023	0.997
35–39	0.955	0.910	0.968
40–44	0.727	0.657	0.639
45–49	0.936	0.990	1.089
Secondary and above			
20–24	1.109	1.000	
25–29	1.107	0.975	1.000
30–34	1.225	1.128	1.288
35–39	1.237	0.960	0.840
40–44	1.070	0.857	0.807
45–49	0.898	0.780	0.769

15–19-year period, to 5.97 and 6.79 for the 10–14- and 5–9-year periods respectively, and decline to 6.37 in the most recent period. A similar pattern is shown in panel C giving the cumulative fertility rates for synthetic cohorts; for each diagonal, they increase systematically up to the 5–9-year period and then decline (except for the oldest and youngest cohorts).

An examination of the P/F ratios for the most recent three 5-year periods since the survey show that the ratios decline as the age of the women increases; the only distortion in this pattern is for women aged 40–44 for whom the ratios are always the lowest. This pattern is typically indicative of omissions or errors in the reference period in the reporting of births by older women. The P/F ratios are slightly greater than unity for the two younger age groups (20–24 and 25–29) thus indicating an incipient decline in fertility for such women.

To verify whether the above patterns are peculiar to particular subgroups of the women, P/F ratios for the three 5-year periods before the survey for selected subgroups of the women are presented in table 26. Except for a few minor fluctuations, the ratios generally

Table 27 Percentage change in cohort-period (cumulative) fertility rates for the three most recent five-year periods before the survey[a]

Age at end of each period	Characteristics and periods					
	Total		Rural		Urban	
	5–9 to 0–4	10–14 to 5–9	5–9 to 0–4	10–14 to 5–9	5–9 to 0–4	10–14 to 5–9
15–19	−37.9	+16.3	−39.2	+12.1	−34.0	+44.5
20–24	+2.9	+14.9	+1.8	+16.5	+8.7	+14.3
25–29	+5.8	+10.6	+5.6	+13.1	+10.5	−3.8
30–34	+5.9	+12.2	+6.8	+19.6	+2.4	+22.6
35–39	+13.2	+20.3	+10.8	+4.7	+23.2	−1.5
40–44	−5.3	–	−5.2	–	−3.8	–
	No education		Koranic		Primary incomplete	
	5–9 to 0–4	10–14 to 5–9	5–9 to 0–4	10–14 to 5–9	5–9 to 0–4	10–14 to 5–9
15–19	−9.7	+43.1	−9.1	+26.7	−53.7	+4.2
20–24	+15.6	+19.7	+5.2	+36.7	+12.0	+6.7
25–29	+7.9	+11.5	+27.6	−4.9	+1.6	+12.7
30–34	+7.5	+14.8	−8.4	+41.3	+4.0	+50.9
35–39	+7.8	−2.5	+45.3	+4.1	+28.2	(−28.1)
40–44	−4.1	–	−0.5	–	(−24.2)	–
	Primary complete		Secondary+		North-east	
	5–9 to 0–4	10–14 to 5–9	5–9 to 0–4	10–14 to 5–9	5–9 to 0–4	10–14 to 5–9
15–19	−27.8	−13.2	−65.0	+27.0	−12.1	+21.6
20–24	−2.0	−1.6	−19.2	−29.2	+5.9	+27.4
25–29	−9.5	+17.4	−17.7	+21.6	+11.0	+33.9
30–34	+16.2	+59.0	−9.1	+18.1	+20.0	+20.5
35–39	+43.0	−33.5	(+14.8)	(+10.4)	+19.2	−13.3
40–44	−21.9)	–	(+19.2)	–	(−14.4)	–
	North-west		South-east		South-west	
	5–9 to 0–4	10–14 to 5–9	5–9 to 0–4	10–14 to 5–9	5–9 to 0–4	10–14 to 5–9
15–19	−5.9	+43.5	−68.8	−6.1	−55.6	+25.0
20–24	+13.2	+30.1	−10.3	−5.0	−0.6	+8.6
25–29	+19.3	−1.4	−9.7	+10.8	+7.8	+5.0
30–34	−2.1	+42.8	+6.4	+12.1	+4.1	+19.6
35–39	+32.8	+23.9	+3.8	−3.9	+12.3	−2.5
40–44	+18.2	–	+3.7	–	−4.9	–

[a] −decline. +increase.

tend to decrease as the age of the women increases for the three periods. The ratios are conspicuously low among the least educated, older women and in the north-west region. Evidence of declining fertility for the youngest women in the last 5-year period before the survey still persists among all subgroups. That evidence holds true for more age groups as the level of education increases and for all age groups in the south-east. In the latter area, it would appear that the decline started earlier, in the 5–9-year period before the survey. The magnitude of the proportional fertility change between 5–9 and 0–4 and 10–14 and 5–9 periods for the total sample and selected subgroups is shown in table 27. A substantial increase from the period 10–14 to the period 5–9 is generally observed except for the younger age groups, the more educated women and those that live in the south-east region. This increase is, however, accompanied by a decline in the 0–4-year period again for the youngest age group (15–19) in all background categories and for more age groups as the level of education increases and in the south-east. Women aged 40–44, in almost all background categories, registered a decline while the immediately younger age groups were registering increases.

The immediate question is whether the overall decline in fertility between 0–4 and 5–9 periods before the survey is genuine or a function of the displacement of births as postulated by Potter (1977a). The postulation holds that such displacement is most typical for older women who tend to displace their earlier births forwards towards the date of the survey, resulting in a concentration of births in the 5–9- and, to some extent, in the 10–14-year periods before the survey. The data do not, however, tend to show this. Focussing attention on fertility rates for the two oldest cohorts (45–49 and 40–44) at the central age of 20, and comparing these rates with those of younger cohorts at the same age, we

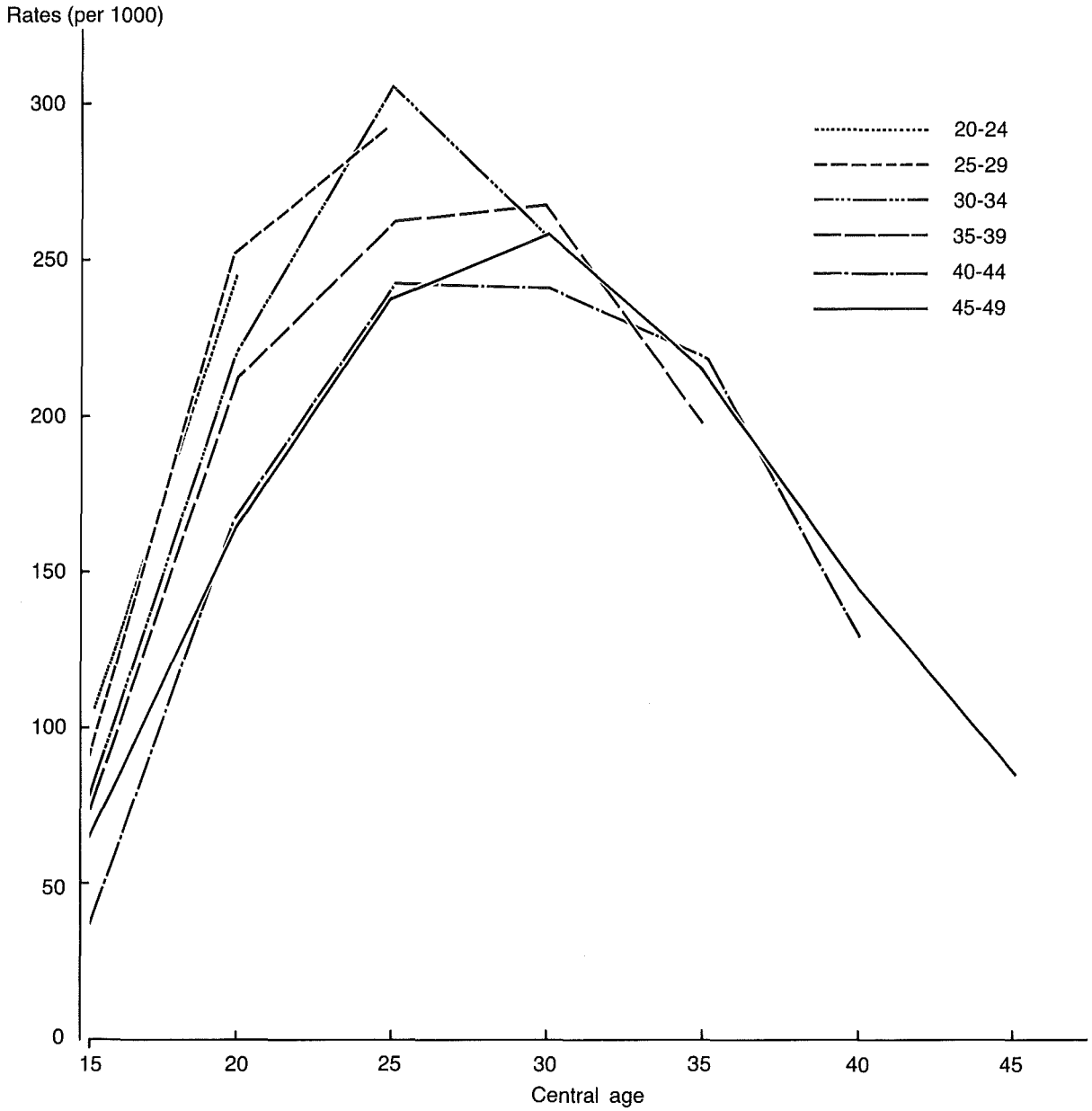

Figure 21 Age cohort-period fertility rates

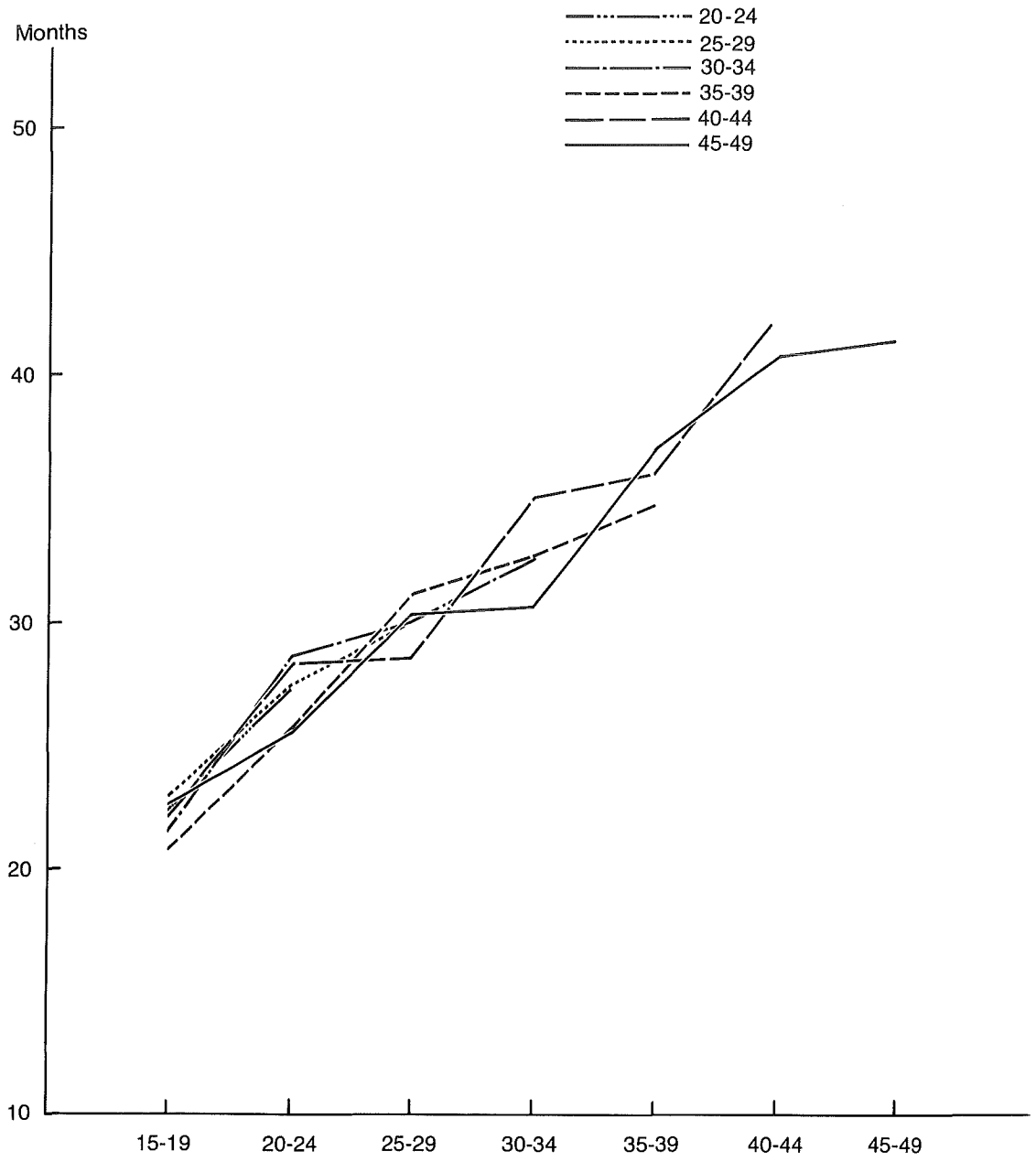

Figure 22 Mean length of closed birth intervals (in months), by age cohorts

observe that the rates for the older cohorts (164 and 162 respectively) are lower than those for the immediately younger cohorts (212 and 221 for the 35–39 and 30–34 cohorts). If these lower rates were due to displacement, we would expect higher rates for later periods and, thus, at higher central ages. But this is not the case; rates at central ages 25 and 30 are still lower for these older cohorts than for the immediately younger ones (figure 21). Moreover, an examination of the mean length of the closed birth intervals does not seem to confirm the displacement of the births; should such a displacement have occurred a dip in the curve of the intervals shown in figure 22 (indicating concentration of births and, thus, shorter intervals) would have occurred at the central ages corresponding to the periods. Thus, it is evident that omission but not much displacement of births had occurred among the older women and that would account for the decline among them. The decline among the younger women may, however, be genuine despite some omission of recent births and probably attributable to increasing age at first marriage and lower proportions getting married especially among the more educated women.

5.6 SUMMARY

Assessment of the quality of fertility data collected in the NFS shows that knowledge of exact dates of births of children ever born is low. The dating of most of the births had to be estimated though this has not led to much heaping at particular years of birth. Analysis of the mean number of children ever born and sex ratios of births by the age of the women show indications of

omission of births, particularly female births, among the older age cohorts. Under-reporting of females also seems more characteristic of lower order than the higher parities and for durations more distant from the survey date. Examination of the number of births shows substantial heaping at preferred digits and the fact that very recent births (of under three years) might have been omitted. There does not seem to be any systematic tendency to under-report dead children or those not living at home at the time of the survey.

Examination of cohort-period fertility rates and associated P/F ratios for the entire sample and selected subgroups shows that fertility levels increased systematically reaching a peak in the 5–9-year period before the survey and declined thereafter not only for the entire sample but for most subgroups. The P/F ratios also reflected the greater omission of births by the older cohorts. The decline in fertility in the most recent 5-year (0–4) period was found to be mostly confined to the youngest age group (15–19) and to other younger age groups (up to 25–29) among the more educated women and those that live in the south-east region. Further analyses showed the decline among these subgroups (higher education, residence in the south-east and aged 15–19) to be genuine, though partly enhanced by the omission of recent births. Observed decline among older cohorts was due solely to omission of births and not to the tendency for them to have transferred their earlier births closer to the date of the survey.

6 Infant and Child Mortality

6.1 INTRODUCTION

Information on the survival status of each live birth recorded in the respondent's maternity history was collected in the NFS; if the child had died, the age at death (in months and years) was ascertained. It was, thus, possible to derive estimates of infant and child mortality directly from these data in conjunction with the information on the date of birth of the children. The obtained estimates of infant and child mortality, like estimates of fertility and nuptiality, may, however, be biased in so far as there are errors in reporting the ages of the children at death, their dates of birth or in omitting completely some children who had died. Both the birth and death of some children may also be omitted. Such omission of dead children, if it occurred, generally biases the obtained estimates downwards. If, as is generally observed, such omission is selective say of age of respondent, sex of child, parity or duration from interview date, then not only will overall levels and trends, but also differentials in obtained rates, be distorted.

Misreporting of the age at death of the dead children often results from ignorance, recall lapse and preference for certain digits. It often leads to concentration of deaths at particular intervals (usually six-month intervals) and seriously affects estimates of neo-natal or infant mortality. It may also affect overall estimates of child mortality though this is not usually the case since most child deaths occur well below the exact age of five. The misreporting of the birth dates of the children, whether dead or alive, would also distort the obtained estimates of mortality. If the date of birth of the children were brought forward (forward displacement) then mortality rates would be overestimated for the periods to which they were moved and an exaggerated decline in mortality may be observed for the period from which they were moved. If the date of birth of the children had been shifted backwards (backward displacement), then the reverse would be the case (Rutstein 1984).

In addition to omission and misreporting, there are also other sources of error that might introduce bias in the obtained estimates. These include censoring, truncation and selection biases. Since the available information was collected from respondents who are themselves survivors from their birth cohorts, the obtained estimates may be biased in so far as mortality to the children of these women differs from the mortality to the children of the non-survivors of their cohorts. This may actually be the case since it is commonly believed that the children of dead mothers are more likely to die than the children of surviving mothers and would result in an underestimation of mortality. Furthermore, as one goes back from the survey date, the average age of the respondents at the time of birth of their children becomes younger and the information obtained for the past is restricted to younger women. Since there is a close relationship between age of mother at birth and the probabilities of survival of the child, this truncation of data would lead to distortions in obtained rates for periods further back from the survey date. Censoring results from the incomplete or unequal exposure of all children to the risk of death in a particular period and, thus, to a lower count of deaths. This problem can, however, be obviated by ignoring the censored cohorts.

The best way of evaluating the effects of most of the above sources of error in the obtained estimates of infant and child mortality is by comparison with results from the vital registration system. Where this is not possible, comparisons with indirect estimates may be helpful though such indirect estimates assume certain patterns of mortality that may not be applicable in a particular context. For the present exercise, the evaluation of the quality of infant and child mortality data as collected in the NFS will be confined to internal consistency checks on the data and the obtained estimates. In doing so, attention will be focussed on the quality of reporting of age at death, the evidence for possible omission of deaths and the plausibility of obtained rates by age of mother at birth, sex, parity, periods and selected characteristics. The fact that the obtained patterns are plausible does not, however, provide concrete evidence that the data are free from error.

6.2 AGE AT DEATH

Some idea about the quality of the age at death data can be obtained by examining the proportion of deaths that are concentrated (ie heaped) on certain ages. Table 28 shows that as many as 9.0, 6.9, 4.1 and 2.4 per cent of all deaths were recorded as occurring at the age of 12, 24, 6 and 18 months respectively. This pattern of heaping at 6-month intervals is characteristic of some misreporting of exact age at death and should be higher if the deaths are restricted to those occurring before the age of 5 years, the upper limit of child mortality. The heaping at the age of 12 months will definitely bias estimates of infant mortality since some of the deaths are likely to have occurred before that age. Should half of the deaths of infants exactly 12 months old be regarded as occurring before that age to compensate for the heaping, then the obtained infant mortality rates would increase by about 8.0 per cent while mortality to children between 1 and 4 years would decrease by 10.0 per cent. Mortality to all children under 5 years will, however, decrease by only 1.0 per cent. It is also observed that 25.0 per cent of all the deaths were reported as occurring before the children

Table 28 Percentage of deaths occurring at selected specific ages (in months), by birth order of child

Birth order	Percentage of deaths				
	0 month	6 months	12 months	18 months	24 months
1	28.7	3.1	8.8	2.4	7.8
2	23.7	4.1	8.6	2.3	5.4
3	24.1	3.2	8.6	1.6	7.1
4	22.2	4.0	9.6	2.8	7.5
5	21.9	6.3	10.6	3.1	6.8
All births	25.0	4.1	9.0	2.4	6.9

were one month old. This proportion is high and would reflect the extent of neo-natal mortality; it is, however, likely that deaths to children just over a month might have been included in this category. Heaping at particular ages varies to some extent by birth order though not in a systematic manner.

6.3 MORTALITY BY AGE OF MOTHER AT BIRTH

Studies of mortality patterns by age of mother at birth show that infant and child mortality levels are usually high for very young mothers (usually under 20 years), low for mothers between 20 and 29 and high again for mothers above that age range. The pattern of relationship between infant/child mortality and age of mother at birth is, thus, in the form of a 'U' or 'J'; omission of deaths often introduces distortions in this pattern.

The data in table 29 show that, for the first two five-year periods before the survey, the relationship between infant mortality and age of mother at birth conforms to this U pattern; this is not the case, however, for other periods further removed from the interview date. The same observation can also be made for childhood mortality rates; they conform to the expected pattern in the 5–9-year period before the survey and not thereafter. The rates for these periods in the more distant past either

Table 29 Probabilities of infant and child death by five-year periods before the survey and age of mother at the time of child's birth

Age of mother at birth	Total	Periods before the survey					
		0–4	5–9	10–14	15–19	20–24	25–29
A Infant mortality ($_1q_0$)							
10–14	0.122	0.084	0.159	0.096	0.090	0.169	(0.161)
15–19	0.114	0.101	0.107	0.115	0.099	0.112	0.211
20–24	0.088	0.079	0.064	0.105	0.120	0.102	
25–29	0.080	0.081	0.065	0.094	0.093		
30–34	0.092	0.077	0.107	0.094			
35–39	0.102	0.112	0.109				
40–44	0.084	0.081					
B Children 1–4 years ($_4q_1$)							
10–14	0.086	–	0.102	0.080	0.082	0.104	0.059
15–19	0.093	–	0.079	0.082	0.101	0.121	0.124
20–24	0.085	–	0.071	0.083	0.102	0.114	
25–29	0.081	–	0.070	0.086	0.080		
30–34	0.076	–	0.056	0.086			
35–39	0.071	–	0.071				
40–44	0.083	–					
C Children 0–5 years ($_5q_0$)							
10–14	0.201	–	0.245	0.168	0.164	0.255	0.210
15–19	0.199	–	0.177	0.188	0.190	0.219	0.309
20–24	0.168	–	0.130	0.179	0.209	0.204	
25–29	0.154	–	0.130	0.171	0.166		
30–34	0.164	–	0.158	0.173			
35–39	0.159	–	0.172				
40–44	0.163	–					

NOTE: Figure in brackets is based on less than 100 cases.

Table 30 Probabilities of infant and child death, by birth order and five-year periods before the survey

Birth order	Periods before the survey				
	0–4	5–9	10–14	15–19	20–24
A Infant mortality ($_1q_0$)					
First births	0.079	0.100	0.102	0.085	0.091
Second and third births	0.084	0.076	0.091	0.115	0.126
Fourth to sixth births	0.089	0.082	0.134	0.115	(0.151)
B Children 1–4 ($_4q_1$)					
First births	0.058	0.062	0.090	0.076	(0.115)
Second and third births	0.079	0.073	0.097	0.132	(0.112)
Fourth to sixth births	0.096	0.095	0.136	(0.143)	(0.160)
C Children under 5 ($_5q_0$)					
First births	0.133	0.156	0.183	0.154	(0.196)
Second and third births	0.157	0.143	0.179	0.282	(0.224)
Fourth to sixth births	0.177	0.169	0.252	(0.240)	(0.287)

NOTE: Figures in brackets are based on less than 100 cases.

decline systematically or fluctuate. Thus it appears that substantial deaths had been omitted in these periods, especially among older women.

6.4 MORTALITY BY BIRTH ORDER

Mortality by birth order is assumed to follow the same pattern as by age of mother at birth, ie it is high for first births, low for middle parities and high again thereafter. The data in table 30, however, show this to be the case in only a few instances: infant mortality and mortality to children under 5 for the 5–9 and 10–14 periods for the survey. For the 5-year period just before the survey, all the rates increase as parity increases; the same is the case for periods over 10 years before the survey though the number of cases is so small in some instances to warrant conclusive statements. It would, thus, appear that deaths of first births in the 5-year period just before the survey and in periods after 14 years were omitted more often than deaths to children of higher parities.

6.5 MORTALITY BY SEX

It is universally known that just as more males than females are born at any given period, there are also more male than female deaths especially during infancy. It is, thus, conventional to expect that infant and child mortality rates would be higher for males than females. Furthermore, in societies where one sex is preferred over the other, greater omission of the less-preferred sex

Table 31 Probabilities of infant and child death, by five-year periods before the survey and sex of child

Sex of dead child	Periods before the survey						
	0–4	5–9	10–14	15–19	20–24	25–29	30–34
A Infant mortality ($_1q_0$)							
Males	0.098	0.092	0.107	0.113	0.107	0.181	0.216
Females	0.078	0.081	0.095	0.089	0.103	0.138	(0.128)
Both sexes	0.088	0.086	0.102	0.102	0.105	0.163	0.183
B Children 1–4 years ($_4q_1$)							
Males	–	0.073	0.086	0.102	0.125	0.079	(0.045)
Females	–	0.071	0.080	0.092	0.108	0.134	(0.186)
Both sexes	–	0.072	0.083	0.097	0.117	0.103	0.101
C Children under 5 ($_5q_0$)							
Males	–	0.158	0.184	0.203	0.218	0.246	0.252
Females	–	0.146	0.168	0.174	0.200	0.254	(0.290)
Both sexes	–	0.152	0.177	0.190	0.210	0.249	0.266
D Ratios of male/female probabilities							
$_1q_0$	1.26	1.14	1.13	1.27	1.04	1.31	(1.68)
$_4q_1$	–	1.03	1.08	1.11	1.16	0.59	(0.24)
$_5q_0$	–	1.08	1.10	1.17	1.09	0.97	(0.87)

NOTE: Figures in brackets are based on less than 100 cases.

Table 32 Probabilities[a] of infant and child death, by calendar years, 1956–80

Year	Single year probabilities			Three-year moving averages		
	$_1q_0$	$_4q_1$	$_5q_0$	$_1q_0$	$_4q_1$	$_5q_0$
1956	0.109	0.099	0.197	–	–	–
1957	0.114	0.102	0.205	0.115	0.117	0.219
1958	0.123	0.151	0.255	0.118	0.127	0.230
1959	0.117	0.128	0.230	0.125	0.148	0.255
1960	0.136	0.166	0.279	0.110	0.127	0.222
1961	0.076	0.086	0.156	0.098	0.116	0.201
1962	0.082	0.095	0.169	0.088	0.093	0.173
1963	0.107	0.097	0.193	0.102	0.099	0.190
1964	0.117	0.105	0.209	0.110	0.102	0.201
1965	0.110	0.104	0.202	0.108	0.104	0.201
1966	0.098	0.103	0.191	0.103	0.096	0.189
1967	0.101	0.082	0.174	0.104	0.087	0.182
1968	0.113	0.077	0.181	0.106	0.087	0.184
1969	0.105	0.102	0.196	0.107	0.088	0.185
1970	0.102	0.086	0.179	0.097	0.085	0.173
1971	0.085	0.066	0.145	0.094	0.077	0.164
1972	0.096	0.080	0.168	0.092	0.075	0.160
1973	0.095	0.078	0.166	0.096	0.070	0.159
1974	0.097	0.051	0.143	0.092	0.065	0.151
1975	0.085	0.066	0.145	0.094	0.077	0.164
1976	0.073	0.076	0.143	0.081	0.079	0.152
1977	0.086	0.095	0.169	0.081	–	–
1978	0.084	–	–	0.089	–	–
1979	0.098	–	–	0.085	–	–
1980	0.074	–	–	–	–	–

[a] $_1q_0$ Probability of death between birth and first year of life.
$_4q_1$ Probability of death between first and fifth year of life.
$_5q_0$ Probability of death before the age of five years.

would be expected (though there may also be greater reluctance to report the death of children of the preferred sex).

Probabilities of infant and child death, by sex and 5-year periods before the survey are given in table 31. So also are the ratios of male to female probabilities of infant and child death. It is observed that for all periods less than 25 years before the survey, the probabilities of death are always higher for males than females. The ratios of male to female probabilities of death are greater than unity throughout those periods as expected; however, they are, in some instances, not high enough to compensate fully for the ratio at birth. This would suggest some selectivity in the omission of female deaths, essentially between the ages of 1 and 4.

6.6 MORTALITY TRENDS

Probabilities of infant and child death by calendar years since 1956 are presented in table 32. Since the probabilities are likely to be unstable because some are based on a small number of births and deaths, probabilities based on three-year moving averages are also presented in the table and in figure 23. Generally, the rates tend to decline though not in a smooth or linear pattern. The data show a pattern of increase from 1956 to 1959 followed by a decline to 1961, another increase to 1964/65 with an uneven decline thereafter. When the rates are calculated by five-calendar year periods (as shown in table 33), the pattern of decline is more definite except that the infant mortality rate is higher in the 1966–70 period (the period of the civil war) than in the preceding period. Also, mortality to children aged one to four years seemed to have increased in the most recent period. The fluctuations in rates for the single years may be attributable to the incorrect placing of death especially within the 5-year periods instead of omission which would have affected the rates by 5-year periods.

6.7 MORTALITY BY RESIDENCE AND EDUCATION

As a further check, the mortality rates by place of residence and education of the respondents may be examined since it is always assumed that the rates are typically higher in the rural than in the urban areas and among the less educated. The data in table 34 show that the probabilities of death are consistently higher among the non-literate respondents and those that live in rural areas. The only exceptions are mortality to children

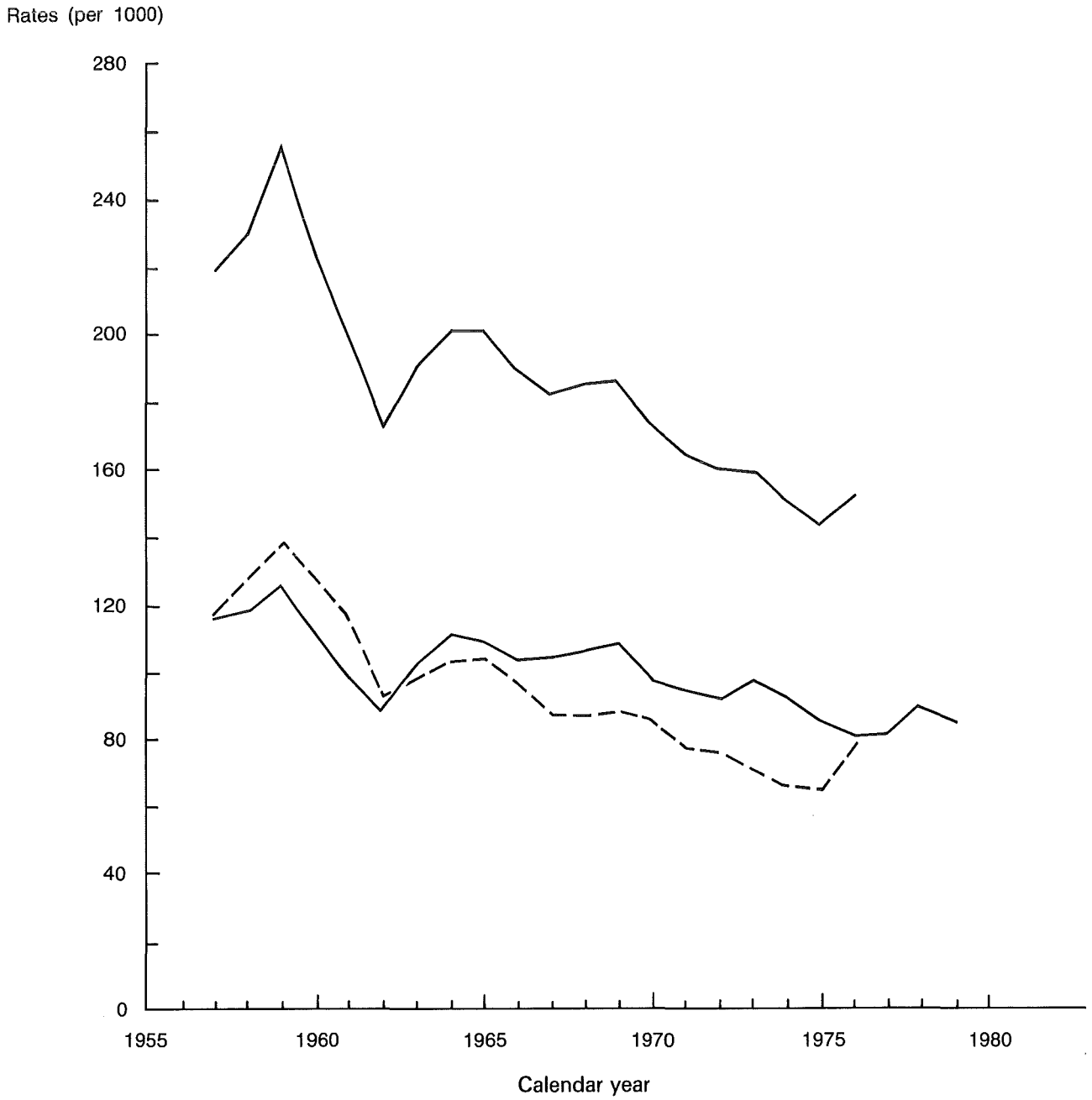

Figure 23 Infant and child mortality rates 1956–80 (three-year moving averages)

Table 33 Probabilities of infant and child death by calendar five-year groups, 1956–80

Five-year groups	Probabilities of death		
	$_1q_0$	$_4q_1$	$_5q_0$
1956–60	0.121	0.132	0.237
1961–65	0.099	0.098	0.188
1966–70	0.103	0.090	0.185
1971–75	0.091	0.068	0.153
1976–80	0.082	0.080	0.149

between one and four ($_4q_1$) in 1976–80 and 1956–60 where it is higher in rural than urban areas. The decline in the rates from 1961–65 period is almost always twice as much among the literate than illiterate and among the rural than urban residents (with the exception of IMR for literate women which increased between 1971–75 and 1976–80).

6.8 SUMMARY

In the absence of external sources of data for comparison, internal checks on the quality of NFS data on infant

Table 34 Probabilities of infant and child death, by five calendar-year periods from the survey and selected characteristics

Characteristics	Five-calendar-year periods				
	1976–80	1971–75	1966–70	1961–65	1956–60
A Literacy status[a]					
Infant mortality ($_1q_0$)					
Literate	0.068	0.065	0.074	0.073	0.106
Not literate	0.087	0.097	0.109	0.103	0.122
Children 1–4 ($_4q_1$)					
Literate	0.050	0.050	0.077	0.070	0.057
Not literate	0.087	0.072	0.093	0.101	0.139
Children under 5 ($_5q_0$)					
Literate	0.088	0.112	0.145	0.138	0.157
Not literate	0.161	0.162	0.192	0.194	0.245
B Rural-urban residence[a]					
Infant mortality ($_1q_0$)					
Rural	0.086	0.097	0.108	0.103	0.124
Urban	0.071	0.068	0.084	0.084	0.105
Children 1–4 ($_4q_1$)					
Rural	0.078	0.072	0.096	0.100	0.130
Urban	0.091	0.051	0.062	0.081	0.147
Children under 5 ($_5q_0$)					
Rural	0.153	0.162	0.194	0.193	0.238
Not literate	0.135	0.116	0.141	0.163	0.236

[a]As at the time of survey.

month intervals for age of death. Analyses of the estimated rates by age of mother at birth of the child, birth order and sex show that there might have been omission of deaths, most probably deaths to females and first births, in periods further removed (10 years or more) from the date of the survey. The data further show that infant and child mortality rates have been consistently declining since 1960, that they have been higher in the rural than urban areas, among the non-educated than educated women and that the decline is more than twice among the educated and urban residents than among the rural and non-educated.

7 Conclusions

An attempt has been made in this document to evaluate the quality of the data collected in the NFS, especially the data on age, nuptiality, fertility and infant/child mortality. The aim was to identify probable sources of errors and biases, assess their incidence and ascertain their effects on obtained parameters and estimates. Such an assessment was considered crucial in determining the extent of reliability of the estimates and of invaluable assistance to future analyses of the data. Since there were hardly any external sources for comparing the obtained estimates (except the age and sex distributions from the 1963 census), most of the assessment had to be based on internal consistency checks and comparisons of the data as obtained from different sections of the same survey and on conformity to logical or theoretical models.

The assessment of the data on age shows that the reporting of age was very poor both in the household and in the individual surveys. There was very substantial heaping at ages ending in 0 and 5 (especially at age 30) and avoidance of almost all other digits. This is because a large majority of the respondents did not know their exact ages which, consequently, had to be estimated. Age with an incorrect final digit was reported for at least a third of the enumerated household population as indicated by a Myers' index of 32.1. Because of such heaping, the age distribution appears as a saw-edged pattern which does not completely disappear even when the distribution is presented for 5-year groups. The extent of misreporting was found to be slightly higher in the rural than in the urban areas, and among males than females. Among the surveyed women, misreporting of age was more characteristic of the older and illiterate women. There was, however, a relatively high degree of consistency in age and age groups as reported in the household and in the individual surveys. The quality of age reporting was only marginally better than in the 1963 census.

There was also substantial heaping in all nuptiality data involving dates: age at first marriage, year of first marriage and years since first marriage. This again was because a majority of the women did not know the exact date of their marriage. The pattern of heaping was on both the conventionally preferred and avoided digits depending on the format in which the date of first marriage was obtained. There was an excessive concentration of first marriages at around 15 years of age and the older respondents tended to report dates of first marriage that were nearer to the survey date. This resulted in substantially higher mean ages at first marriage for these women and a spurious increase in age at marriage. The mean interval between first marriage and first birth was found to be excessively long, especially for these older women because of the misreporting of age at first marriage. There was, however, no evidence to show that there has been some significant omission of marriages. It was also observed that there has been a recent and genuine decline in the proportion of women married, especially among those under 25.

The data on fertility again showed that the dating of the live births was poor; the exact date of birth was reported for a third of the last live births and a quarter each of the penultimate and first live births. The distribution of the live births by year of birth, however, showed less heaping than was observed for the age of respondents or nuptiality variables involving dates. The data show indications of some selective omission of female births (especially among the older cohorts), at lower order parities and for durations more distant from the survey date. There was also some decline in the number of recent births especially since 1978, some of which was due to omission. No systematic tendency to omit dead children or those living away from home was observed. The cohort-period fertility rates and associated P/F ratios tend to suggest that fertility increased from the period 20–24 to 5–9 years before the survey and then declined in the last 5-year (0–4) period. Most of this decline, especially among the older women, is spurious because of the omission of births, the poor reporting of the ages of those women and some displacement of births over the years. However, some of the decline to the younger women and to the more educated women is genuine and due to increasing age at first marriage and corresponding declines in the proportions married among them.

The decline in infant and child mortality for the last decade is more genuine and universal. There was some heaping at six-month intervals of age at death. The analyses of infant and child deaths revealed some omission of deaths, particularly of females and low parities in the period over ten years from the survey date. The extent of omission was, however, considerably less for the two five-year periods just before the survey. The decline in both infant and child mortality was substantially higher in rural than in urban areas and among the educated compared with illiterate respondents.

In conclusion, the dating of vital events was found to be poor in the NFS. Coverage of events was, however, relatively adequate if not very complete. Because of this, most of the estimates of fertility, nuptiality and mortality are reliable to the extent that they are presented for either broad age groups or broad durations. Such broad groups tend to encompass the entire period during which the events have been displaced. In this context, more reliable estimates may be obtained when ten-year age groups or periods before the survey are used. This exercise has also shown the need for developing new techniques for estimating the dating of vital events. Though the problem of age reporting is one that is characteristic of most developing countries, it is one that would call for innovative techniques and probably better training of fieldstaff. The lack of external sources with which the NFS data could be compared further reinforces the often expressed need for more surveys of the NFS type in the future.

References

Brass, W. (1974). Métodos para estimar la fecundidad y la mortalidad en poblaciones con datos limitados. *Selección de trabajos de W. Brass*. CELADE, serie E, no 14. Santiago, Chile.

Brass, W. (1978). Screening Procedures for Detecting Errors in Maternity History Data. WFS Technical Paper no 810. See also *Asian Population Studies Series* no 44, 1978; and *WFS Occasional Papers* no 22, 1980.

Goldman, N. (1984). Assessment of the Quality of Fertility Data Collected in WFS Individual Surveys. Background Paper for WFS 1984 Symposium. Published with revisions in N. Goldman, S. O. Rutstein and S. Singh, Assessment of the Quality of Data in 41 WFS Surveys: a Comparative Approach *WFS Comparative Studies* no 44, 1985.

National Population Bureau (1984). *The Nigeria Fertility Survey 1981–82: Principal Report*. Lagos.

Potter, J. E. (1977a). Problems in Using Birth History Analysis to Estimate Trends in Fertility. *Population Studies 31(2)*: 335–64.

Potter J. E. (1977b). Methods of Detecting Errors in WFS Data: an Application to the Fiji Fertility Survey. *International Population Conference, Mexico City, 1977 1:* 101–22. IUSSP.

Rutstein, S. O. (1984). An Assessment of the Quality of Age Reports for Eligibility and Analysis. Background Paper for WFS 1984 Symposium. Published with revisions in N. Goldman, S. O. Rutstein and S. Singh, Assessment of the Quality of Data in 41 WFS Surveys: a Comparative Approach *WFS Comparative Studies* no 44, 1985.

World Fertility Survey (1984). The Nigeria Fertility Survey 1981–82: a Summary of Findings. *WFS Summaries of Findings* no 49.